



The
University
Of
Sheffield.

Hannah Gillian Barrett

Department of Geography

The University of Sheffield

El Niño Southern Oscillation from the pre-instrumental era: development of logbook-based reconstructions; and evaluation of multi-proxy reconstructions and climate model simulations

PhD Thesis

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

AUGUST 2017

Abstract

El Niño Southern Oscillation (ENSO) is a major mode of climatic variability which impacts weather patterns around the globe on inter-annual timescales. This thesis explores ENSO during the pre-instrumental period in order to gain a better understanding of its behaviour on longer time scales than available from the instrumental record. Firstly, the wind observations from ships' logbooks are used to reconstruct a new index of ENSO for the period, 1815-1854. New methods for assessing the robustness of these types of reconstructions which use limited data availability are presented. The logbook-based ENSO reconstructions are then compared to a range of existing ENSO reconstructions over this historical period. It is found that there is disagreement between the various ENSO reconstructions in the early-to-mid nineteenth century and that the majority of the reconstructions are biased towards identifying El Niño events compared to La Niña events, and towards identifying Eastern Pacific El Niño events compared to Central Pacific El Niño events. Historical reanalysis products are used to explore ENSO variability over the twentieth century and provide a data source for comparison with historical climate model simulations. The number and quality of observations assimilated into the historical reanalysis increases through the twentieth century, thus the agreement in ENSO indices is weaker earlier in this period. Historical- (1900-2004 AD) and paleo- (850-1848 AD) climate model simulations are used to assess ENSO further back in time, showing that the strength and character of ENSO-wind teleconnections vary over the Last Millennium. The results from this thesis suggest that there are still a number of uncertainties associated with the range of data types used: the observations from ships' logbooks, documentary records, multi-proxy reconstructions, reanalysis data and global climate model simulations. Only once these have been addressed can current and future ENSO behaviour be placed within its true historical context.

Acknowledgements

Thank you to my supervisors Dr. Julie Jones and Prof. Grant Bigg for their guidance, support and feedback throughout this research. I would also like to thank the Department of Geography at the University of Sheffield, specifically the Ice and Climate Research Group for funding this work and providing the funding to attend conferences and meet other researchers within this research area. Thank you to everyone in the research group and within the department who have supported me in my research during my time here.

I would like to thank all the friends I have made in Sheffield who have made the past three years a fun and enjoyable experience. In Geography, I would especially like to thank Kate Orgill and Tom Wilkes for all the advice, breaks and laughs. Also thanks to all the other PhD students I have spent time with in Sheffield and at conferences.

Thanks to Steven for supporting me over the past three years and providing a good distraction from research life. Finally, I would like to thank my family for always being there for me. Thanks Mam, Dad, Guy, Frances, Luke and Megan.

Table of Contents

Abstract	I
Acknowledgements	II
Table of Contents	III
List of Figures and Tables.....	VIII
List of Abbreviations.....	XXVI
1. Introduction	1
1.1 Introduction	1
1.2 Mean Pacific climate and El Niño Southern Oscillation.....	1
1.3 ENSO teleconnections and impacts	5
1.4 The instrumental record of ENSO.....	9
1.5 Ship logbook observations as a source of information for historical ENSO 11	
1.6 Data sources for historical- and paleo-ENSO	13
1.7 Aims of the thesis	15
1.8 Thesis structure.....	19
2. Reconstructing El Niño Southern Oscillation using wind observations from ships’ logbooks, 1815-1854.....	21
2.1 Introduction	21
2.2 Data	24
2.2.1 Reanalysis data	24
2.2.2 Instrumental ENSO data	25
2.2.3 Ships’ logbook data	25
2.3 Method.....	31
2.3.1 Selection of predictor grid boxes	31
2.3.2 Reconstruction and validation methods.....	34
2.3.3 Assessing reconstruction skill.....	35
2.3.4 Threshold tests	36
2.3.5 Data origin and wind force terms	39
2.4 Results	42
2.4.1 Reconstruction skill	42
2.4.2 Modern wind patterns during strong El Niño and La Niña events	43
2.4.3 Historical ENSO reconstructions.....	45

2.5 Discussion	49
2.5.1 Difference in wind force terminology	49
2.5.2 Difference in reconstruction methods: PCR and CPS	52
2.5.3 Threshold tests on number of observations per season	52
2.5.4 Comparison to Jakarta based records of the SOI.....	53
2.6 Summary	55
3. Comparison of ENSO reconstructions from ships' logbooks with existing multi-proxy and documentary-based ENSO reconstructions	57
3.1 Introduction	57
3.2 Data	59
3.2.1 Logbook-based DJF SOI reconstructions.....	59
3.2.2 Multi-Proxy ENSO reconstructions.....	61
3.2.3 Documentary ENSO reconstructions.....	67
3.3 Methods.....	69
3.4 Results	72
3.4.1 Fitting period events analysis, 1979-2013	72
3.4.2 Multi-proxy comparison to instrumental DJF SOI, 1876-1977	78
3.4.3 Documentary-based ENSO chronology comparison to instrumental DJF SOI, 1876-1977	81
3.4.4 Multi-proxy comparison over logbook-based reconstruction period, 1815-1854	82
3.4.5 Documentary chronology comparisons over the logbook-based reconstruction period, 1815-1854.....	87
3.4.6 Combined multi-proxy and documentary reconstruction, 1815-1854.....	88
3.5 Discussion	90
3.5.1 Skill of reconstructions in representing El Niño and La Niña events	90
3.5.2 Influence of non-ENSO factors on reconstruction: Case study 1818-1820 La Niña	91
3.5.3 Disagreement with documentary chronologies	93
3.6 Summary	95
4. ENSO diversity in historical and paleo-reconstructions: The detection of Eastern Pacific and Central Pacific El Niño events from ships' logbooks, documentary chronologies and multi-proxy reconstructions.....	97
4.1 Introduction	97
4.2. Data	101

4.2.1 Modern records of EP and CP El Niño events.....	101
4.2.2 Existing ENSO reconstructions	102
4.2.3 Teleconnection data	103
4.3 Methods	103
4.4 Results	104
4.4.1 ENSO diversity in the logbook-based reconstruction methods and predictor grid boxes, 1979-2013	104
4.4.2 ENSO diversity within the logbook-based reconstructions, 1815-1854..	109
4.4.3 Documentary records: Event chronology comparison to EP and CP El Niño events, 1876-1977	114
4.4.4 Multi-proxy ENSO reconstructions	117
4.4.5 North American tree rings and teleconnections during EP and CP El Niño events	120
4.4.6 Pacific corals and teleconnections during EP and CP El Niño events.....	125
4.5 Summary	129
5. El Niño Southern Oscillation from instrumental records and historical reanalysis products: Mean climatology and ENSO indices	133
5.1 Introduction	133
5.2 Data	138
5.3 Methodology	140
5.4 Results	142
5.4.1 Mean Sea Level Pressure climatology.....	142
5.4.2 The Southern Oscillation Index	146
5.4.3 Surface temperature climatology	150
5.4.4 The Niño 3.4 index	151
5.4.5. Comparison of the DJF SOI and the DJF Niño 3.4 index	155
5.5 Summary	160
6. El Niño Southern Oscillation from historical climate model simulations: Model climatology, ENSO indices and ENSO diversity over the twentieth century.....	161
6.1 Introduction	161
6.1.1 Introduction to climate models	161
6.1.2 ENSO in climate models.....	163
6.1.3 The representation of El Niño diversity within climate models	165
6.1.4 Model evaluation: diagnostics and performance metrics	166

6.2. Data	168
6.3. Methodology	170
6.4. Results	172
6.4.1 Mean Sea Level Pressure climatology.....	173
6.4.2 The Southern Oscillation Index	177
6.4.3 Surface temperature climatology.....	186
6.4.4 The Niño 3.4 index	187
6.4.5 ENSO diversity in the models over the twentieth century	193
6.5 Summary of historical simulation results.....	201
7. The representation of the mean zonal wind and the stationarity of ENSO-wind teleconnection within historical reanalysis and climate model simulations	206
7.1 Introduction	206
7.1.1 Zonal winds within historical reanalysis	207
7.1.2 Historical and paleo-winds in proxies and climate models	210
7.1.3 Strength and stationarity of historical ENSO teleconnections	210
7.2 Data and methods	214
7.3 Results	215
7.3.1 Mean DJF zonal wind in logbook predictor grid boxes within reanalysis	215
7.3.2 Mean DJF zonal wind in logbook grid boxes within historical climate model simulations.....	220
7.3.3 Correlations between DJF zonal wind and the DJF SOI, 1979-2004.....	225
7.3.4 Correlation between zonal wind and the SOI over the twentieth century	230
7.3.5 Stability of the SOI-wind teleconnections over the twentieth century	235
7.4 Summary	238
8. ENSO-teleconnections and ENSO diversity over the Last Millennium, 850/51-1848/49 AD.....	242
8.1 Introduction	242
8.1.1 ENSO in the Last Millennium climate model simulations.....	243
8.1.2 ENSO-teleconnections in the LM simulations	245
8.1.3 ENSO variability and events, ENSO-wind teleconnection and ENSO diversity within two LM climate simulations.....	248
8.2 Methods.....	249
8.3. Results	250

8.3.1 The DJF SOI in the Last Millennium	250
8.3.2 The Stability of ENSO-wind teleconnections in the Last Millennium....	259
8.3.3 El Niño diversity over the Last Millennium	267
8.4 Summary	271
9. Discussion and Conclusion	275
9.1 Discussion	275
9.2 Reconstruction methods and data comparisons.....	275
9.2.1 Differences in logbook-based reconstruction methodologies.....	275
9.2.2 Threshold test methodology.....	276
9.2.3 Differences in wind terms.....	277
9.2.4 Comparisons of multiple data sources	277
9.2.5 Defining ENSO events from instrumental indices	279
9.2.6 Future expansion of ENSO reconstructions.....	280
9.3 Stationarity of the ENSO-wind teleconnection	282
9.3.1 The assumption of stationarity.....	282
9.3.2 Stationarity over the twentieth century	283
9.3.3 Stationarity over the Last Millennium	284
9.3.4 Implications of non-stationarity and future work	285
9.4 ENSO Diversity.....	286
9.4.1. Methods of identifying and classifying El Niño types	286
9.4.2. ENSO diversity in reconstructions	288
9.4.3. ENSO diversity in climate models.....	290
9.4.4. Future studies on ENSO diversity	291
9.5 Conclusion.....	292
Bibliography.....	295
Appendix A.....	311
Appendix B	316
Appendix C	321
Appendix D.....	328
Appendix E	341

List of Figures and Tables

Fig. 1.1 Typical Normal conditions over the equatorial Pacific (National Oceanic and Atmospheric Administration, NOAA, 2017a).	3
Fig. 1.2 Typical a) El Niño and b) La Niña conditions over the equatorial Pacific (NOAA 2017a).	5
Fig. 1.3 Typical El Niño climate impacts during DJF and JJA seasons (NOAA 2017b).	6
Fig. 1.4 Typical La Niña climate impacts during DJF and JJA seasons (NOAA 2017c).	8
Fig. 1.5 Southern Oscillation Index (SOI) since 1880 (National Institute of Water and Atmospheric Research 2017).	10
Fig. 1.6 The SST regions of the Niño indices. Niño 1+2 are the ‘classical’ El Niño regions. Niño 3.4 SST region commonly used in calibrating reconstructions of ENSO, while Niño 4 covers the central Pacific region. (NOAA 2017d).	11
Fig. 2.1 Francis Beaufort’s wind scale as originally proposed in the logbook of HMS <i>Woolwich</i> facing the page dated 13 th Jan 1806 (Courtesy of UK Met Office, within Wheeler and Wilkinson 2005).	27
Fig. 2.2 The un-adjusted (left) and adjusted (right) route of HMS <i>Surprise</i> (1750-1751). The route is adjusted to take into account the modern day zero meridian and with incremental adjustments made to get route over the ocean (Können and Koek 2005).	29
Fig. 2.3 Location of data entries from ships’ logbooks with both wind force and wind direction observations from the digitised archives of CLIWOC and EEIC, 1750-1854.	30
Fig. 2.4 Yearly distribution of all logbook data entries (grey) and those with wind vector observations (black), 1750-1854.	30
Fig. 2.5 Pearson’s correlation coefficient between de-trended mean local noon DJF zonal wind and DJF Southern Oscillation Index, 1979/80-2013/14. Only those significant at the 90% confidence level are shown. Grid boxes used in the reconstructions are outlined and are numbered 1 to 9 from West to East. Black dots indicate location of DJF logbook wind observations 1815/16-1853/54.	32
Fig. 2.6 Number of days with observations per DJF season in the nine predictor grid boxes, 1815/16-1853/54.	33
Table 2.1 Pearson’s correlation coefficient (r) and RE Skill score from fitting reconstructions (1979/80-2013/14) carried out using 1 to 10 days of observations per grid box and full season of observations, from both PCR and CPS methods. The r and RE values shown are the mean from 100 tests at each threshold.	38
Fig. 2.7 RE skill scores obtained from PCR (dark grey) and CPS (light grey) reconstructions carried out over the fitting period, replicating the logbook data availability from each of the DJF seasons, 1815-1854. 1829/30 and 1844/45 are missing data years. Dashed line indicated RE=0.30.	39

Fig. 2.8 The number of DJF wind observations from the regions used in the SOI reconstruction which come from Dutch and English East India Company during DJF seasons 1815/16 -1853/54.	40
Fig. 2.9 Frequency of Beaufort-equivalent wind force terms used by the Dutch and EEIC in the a) Tropics (Grid boxes 1 and 7-9) b) Trade wind region (5-6) and c) southern African grid boxes (2-4).	41
Fig. 2.10 Instrumental DJF Southern Oscillation Index and reconstructed DJF SOI using (a) Principal Component Regression and (b) Composite-Plus-Scale (normalised values), 1979/80-2013/14. Dashed lines represent ± 1 standard deviation, the threshold used to define the main ENSO events.	43
Fig. 2.11 DJF zonal wind anomalies (ms^{-1}) during El Niño events in 1982/83 (left) and 1997/98 (right), relative to the 1979-2013 climatology. Predictor grid boxes used in the reconstruction are outlined in black.	44
Fig. 2.12 DJF zonal wind anomalies (ms^{-1}) during La Niña events in 1998/99 (left) and 2010/11 (right), relative to the 1979-2013 climatology. Predictor grid boxes used in the reconstruction are outlined in black.	44
Fig. 2.13 Reconstructions of DJF SOI 1815/16-1853/54 using Principal Component Regression (PCR) and Composite-Plus-Scale (CPS) (normalised), with (a) the same mean throughout and (b) using 1815-1833 and 1834-1853 means. Thin lines show the 95% confidence intervals. Light grey bars indicate years with RE scores less than 0.3. No data for 1829/30 and 1844/45.	45
Table 2.2 Pearson's correlation coefficients between historical reconstructions of DJF SOI 1815/16-1853/54 using Principal Component Regression (PCR) and Composite-Plus-Scale (CPS) (normalised), with (a) the same mean throughout and (b) using 1815-1833 and 1834-1853 means. All significant at 99% confidence level.	46
Fig. 2.14 Frequency of El Niño (negative SOI events) and La Niña (positive SOI events) in logbook-based reconstructions, 1815/16 – 1853/54.	48
Fig. 2.15 La Niña, 1836/37, (left) and El Niño, 1851/52, (right) zonal wind anomalies (ms^{-1}), relative to the 1834/35 – 1853/54 DJF climatology.	48
Fig. 2.16 CPS B logbook-based DJF SOI (1830/31 – 1850/51) and Jakarta rain-day June-Nov SOI (1830 – 1850).	54
Fig. 2.17 CPS B logbook-based DJF SOI reconstruction and Können et al. (1998) Jakarta single-site DJF SOI (1841/42 – 1853/54). Blank years indicate years with missing data.	54
Table 3.1 Mean number of DJF logbook observations per grid box (1-9), 1815-1854 used in sub-sampled ERA-Interim data.	61
Fig. 3.1 Location of proxies used in ENSO reconstruction studies by ST98 (yellow triangles only), MANN00 (purple) and MCG10 (yellow, purple and red), with the exception of Cook et al. (2008) used in MCG10 as specific location data unavailable. Shapes indicate different types of proxy records (triangle = Tree-rings, circle = corals, star = ice cores, cross = instrumental).	62

Fig. 3.2 Location of proxies used in the ENSO reconstruction of W10, Tropical Teleconnection ENSO reconstruction. Shapes indicate different types of proxy records (circle = corals, star = ice cores).	63
Fig. 3.3 Location of proxies used in ENSO reconstruction studies by EG13A (35 of these), EG13B (36) and EG13C (37). Shapes indicate different types of proxy records (triangle = Tree-rings, circle = corals, star = ice cores, square = sediment). 64	64
Fig. 3.4 The seven regions containing proxies used in the ENSO reconstruction by LI13. Number in parentheses indicates number of Tree rings in each region. Also shown is the weighting of the data from each region.	65
Table 3.2 Proxy based ENSO reconstructions used in this thesis.	66
Table 3.3 Documentary-based ENSO chronologies used in this thesis.	68
Fig. 3.5 Time periods over which comparisons are made between various ENSO reconstructions and records.	70
Fig. 3.6 El Niño (orange) and La Niña (blue) events from the instrumental DJF SOI and the full PCR and CPS fitting reconstructions, 1979-2013.	73
Table 3.4 Hit rates for El Niño and La Niña events from PCR and CPS fitting full and mean of five sub-sampled iterations compared to the instrumental DJF SOI.	73
Fig. 3.7 DJF zonal wind composite anomalies (ms^{-1}) during El Niño events (1982/83, 1986/87, 1991/92, 1997/98, 2004/05 and 2009/10) and La Niña events (1988/89, 1998/99, 2007/08 2008/09, 2010/11, 2011/12), relative to the 1979-2013 climatology. Zonal wind data are taken from the ERA-Interim reanalysis. Predictor grid boxes used in the reconstruction are outlined in black, numbered 1-9 from west to east.	75
Fig. 3.8 DJF zonal wind composite anomalies (ms^{-1}) during El Niño events defined using the one standard deviation threshold of the instrumental SOI (1982/83, 1986/87, 1991/92, 1997/98, 2004/05 and 2009/10, relative to the 1979-2013 climatology). Zonal wind data taken from ERA-Interim reanalysis. Predictor grid boxes used in the reconstruction are outlined in black.	76
Fig. 3.9 DJF zonal wind anomalies (ms^{-1}) during La Niña events defined using the one standard deviation threshold of the instrumental SOI (1988/89, 1998/99, 2007/08, 2008/09, 2010/11 and 2011/12 relative to the 1979-2013 climatology). Predictor grid boxes used in the reconstruction are outlined in black.	77
Fig. 3.10 Eight (ST98, MANN00, MCG10, W10, EG13A, EG13B, EG13C and LI13) reconstructed ENSO indices and the instrumental DJF SOI (black), 1876-1977. Non-SOI indices are inverted.	79
Table 3.5 Pearson's correlation coefficients between the Instrumental SOI (DJF) and the Multi-Proxy ENSO reconstructions (Year ₁ comparisons). All correlations are significant at the 99% confidence level.	79
Fig. 3.11 El Niño (negative) and La Niña (positive) events, based on classification of ± 1.0 standard deviation from the mean (1876-1977) from all the multi-proxy reconstructions and the instrumental DJF SOI. Shown for Year ₁ (year that contains Jan).	80

Table 3.6 Hit Rates (Year ₁) between the instrumental DJF SOI and multi-proxy ENSO events with classification thresholds of ± 1.0 standard deviation 1876-1977 mean.	81
Table 3.7 Year ₀ and Year ₁ (in parentheses) correspondence between instrumental DJF SOI ENSO and QUINN87 and GF09, 1876-1977.	82
Table 3.8 Correlation between the normalised ENSO proxies 1876-1977. Bold highlights those which are significant at the 95% confidence level and non-SOI reconstructions are inverted to match the SOI.	83
Table 3.9 Correlation between the normalised ENSO proxies 1816-1854. Bold highlights those which are significant at the 95% confidence level and non-SOI reconstructions are inverted to match the SOI.	84
Fig. 3.12 Historical normalised ENSO indices from eight multi-proxy reconstructions (colour) and the four logbook-based SOI reconstructions (black) 1815/16 to 1853/54. The missing years in the logbook-based reconstruction are 1829/30 and 1844/45.	85
Table 3.10 Correlation coefficients between the logbook-based reconstructions and multi-proxy reconstructions 1815/16 – 1853/54. Those significant at the 95% confidence level are shown in bold. All indices were normalised over the period 1815-1853.	85
Fig. 3.13 Correspondence (Year ₁) between logbook-based ENSO events and multi-proxy ENSO events with classification thresholds of ± 1.0 standard deviation 1815-1854 mean. Year ₁ (that which contains Jan) shown.	86
Table 3.11 Correspondence rates (Year ₁) for El Niño and La Niña events 1815-1854 between logbook and multi-proxy reconstructions. Bold highlights those which meet or exceed the 40% threshold.	87
Fig. 3.14 El Niño chronologies from documentary sources, Year ₀ (black) and logbook-based reconstructions (grey) 1815/16 – 1853/54. Bars indicate El Niño event.	88
Table 3.12 Correspondence rates Year ₀ (Year ₁ in parenthesis) for El Niño and La Niña events 1815-1854 between logbook and GF09 ENSO reconstruction.	89
Fig. 3.15 Zonal wind anomalies (ms^{-1}) in 1818/19 and 1819/20, relative to the 1815/16 - 1833/34 DJF climatology.	92
Table 4.1 Hit rates for EP and CP El Niño events from PCR, CPS and the instrumental DJF SOI, compared to the NINO method, 1979-2013. PCR and CPS full fitting and the mean of five sub-sampled iterations are presented.	105
Fig. 4.1 DJF zonal wind composite anomalies (ms^{-1}) during EP El Niño events (1982/83, 1991/92 and 1997/98) and CP El Niño events (1994/95, 2002/03 and 2009/10), relative to the 1979-2013 climatology. Zonal wind data are taken from the ERA-Interim reanalysis. Predictor grid boxes used in the reconstruction are outlined in black, numbered 1-9 from west to east. Brackets indicate the number of events used to calculate each composite.	107
Fig. 4.2 DJF zonal wind anomalies (ms^{-1}) during EP El Niño events 1982/83, 1991/92 and 1997/98 relative to the 1979-2013 climatology. Predictor grid boxes used in the reconstruction are outlined in black.	108

Fig. 4.3 DJF zonal wind anomalies (ms^{-1}) during CP El Niño events 1994/95, 2002/03 and 2009/10, relative to the 1979-2013 climatology. Predictor grid boxes used in the reconstruction are outlined in black.....	109
Fig. 4.4 DJF Zonal wind anomalies (ms^{-1}) from logbook observations for CPS B El Niño years with 1824/25 and 1830/31 anomalies relative to the mean of 1815-1833 and 1833/34, 1838/39 and 1851/52 relative to the mean of 1834-1854.	111
Table 4.2 Classification of CPS B El Niño events into EP and CP El Niño.....	112
Fig. 4.5 DJF Zonal wind anomalies from logbook observations for PCR B El Niño years, 1835/36, 1842/43 and 1847/48 relative to the 1834-1854 mean. See Fig. 4.3 for the additional three PCR B events, which are also found in CPS B (1824/25, 1830/31, 1851/52).....	113
Table 4.3 Classification of PCR B El Niño events into EP and CP El Niño, Bold indicates events unique to PCR B, thus not already presented in Table 4.2 for CPS B.	114
Table 4.4 EP and CP skill of documentary records compared to target record of events from the NINO method, 1876-1977.	115
Fig. 4.6 Jan-Dec precipitation Composite anomaly plots for EP El Niño (1982/83, 1991/92 and 1997/98) and CP El Niño (1994/95, 2002/03 and 2009/2010) events during Year ₀ and Year ₁ of the events, using data from GPCP and made using NOAA composite plotting tool. Precipitation (mm/day) composite anomaly 1981-2010 climatology (NOAA 2017g).	116
Table 4.5 EP and CP skill of multi-proxy reconstructions compared to target record of events from the NINO method, 1876-1977.....	117
Table 4.6 EP and CP skill of multi-proxy reconstructions compared to target record of events from Schollaen et al. (2015), 1900-1977.....	118
Fig. 4.7 Oct-March Precipitation anomalies from GPCP from NOAA composite plotting tool for EP events defined from NINO method (1982/82, 1991/92 and 1997/98) and CP events (1994/95, 2002/03 and 2009/10), and locations of tree ring proxies used from STH98=yellow, MANN00=red, MCG10=red and pink, EG13B=green. The bottom plot shows the EP El Niño anomalies minus the CP El Niño.....	122
Fig. 4.8 DJF surface temperature anomalies from ERA-Interim for EP events defined from NINO method (1982/82, 1991/92 and 1997/98) and CP events (1994/95, 2002/03 and 2009/10), and locations of tree ring proxies used from STH98=yellow, MANN00=red, MCG10=red and pink, EG13B=green. The bottom plot shows the EP El Niño anomalies minus the CP El Niño.....	123
Fig. 4.9 Composite April-March SSTA (HaDISSTv4) for EP (above) and CP (below) El Niño years defined using NINO method. Dots indicate the location of coral records used by multi-proxy ENSO reconstructions (MCG10=Red; W10=purple and EGB13= green). The bottom plot shows the EP El Niño anomalies minus the CP El Niño anomalies.	126
Fig. 4.10 April-March GCPC precipitation (mm/day) composite anomaly (1981-2010) climatology for the three HadISST NINO method defined EP El Niño events (1982/83, 1991/92, 1997/98) and the three CP El Niño events (1994/5, 2002/3,	

2009/10). The bottom plot shows the EP El Niño anomalies minus the CP El Niño anomalies. Dots indicate the location of coral records used by multi-proxy ENSO reconstructions (MCG10=Red; W10=purple and EGB13= green)..... 127

Fig. 5.1 Pacific Walker Circulation in HadSLP2 (The Hadley Centre’s Sea Level Pressure reanalysis), 20CR, NNR (NCEP-NCAR), ERA40, ERA-Interim and SOCOL (a chemistry-climate GCM) ensemble mean (grey shading is ensemble SOCOL max and min) (Adapted from Compo et al. 2011)..... 137

Fig. 5.2 Comparison of the Niño 3.4 index (above) and the SOI (below) from a number of reanalyses JRA-55 (Japanese 55-year reanalysis) ERA-Interim, 20CR, ERA-20CM (ensemble run with no observations assimilated) and ERA-20C (Adapted from Poli et al. 2016)..... 138

Table 5.1 Reanalysis and SST analyses used in this thesis: original resolution, time span and reference. 140

Fig. 5.3 Mean Sea Level Pressure over Tahiti (upper) and Darwin (lower) boxes, December, January and February, 1900-2004, from ERA-20C reanalysis..... 141

Fig. 5.4 Mean sea level pressure (hPa) DJF climatology 1979/80 – 2004/05 from reanalysis datasets (all 1.5° x 1.5° latitude-longitude)..... 143

Fig. 5.5 DJF MSLP from reanalyses, 1979-2004. Black rectangles indicate the western and eastern regions used to define SOI from the reanalysis and model simulations. 144

Table 5.2 DJF MSLP from Darwin and Tahiti grid box (hPa), and MSLP gradient, 1979-2004 and 1900-1999. 144

Fig. 5.6 DJF MSLP from reanalyses and instrumental records Darwin (upper) and Tahiti (lower), 1900-2004. 145

Fig. 5.7 The DJF SOI from instrumental record and reanalyses, 1900/01 -2004/5, with 1979/2004 reference period for all records. 146

Table 5.3 Pearson’s correlation coefficient DJF SOI from the instrumental and reanalysis data over different analysis periods, all significant at the 99% confidence level. 147

Table 5.4 Number of El Niño and La Niña events from instrumental and reanalysis DJF SOI, 1979-2004 and 1900-1999. 148

Fig. 5.8 DJF surface temperature from SST analysis and reanalysis, 1979-2004. Black rectangle indicates the Niño 3.4 region. 151

Fig. 5.9 The DJF Niño 3.4 from the three reanalyses (ERA-20C, 20CR and ERA-Interim) and ERSSTv4, 1900-2004 (Anomaly in Appendix, Fig. B1) 152

Table 5.5 The Pearson’s correlation coefficient between the DJF Niño 3.4 indices from the ERSST record and reanalysis data over different analysis periods, all significant at the 99% confidence level. 152

Table 5.6 The DJF Niño 3.4 region mean surface temperature and standard deviation (std) over 1979-2004, 1900-1949, 1950-1999 and 1900-1999. 152

Table 5.7 Number of El Niño and La Niña events defined using the 1.0 standard deviation threshold from the Niño 3.4 anomaly 1979-2004. 154

Table 5.8 The Pearson’s correlation coefficient between the DJF SOI and the Niño 3.4 index within the instrumental record and the three reanalyses over various periods covering the twentieth century.	155
Fig. 5.10 DJF MSLP anomaly composites for El Niño events (left) and La Niña events (right) from reanalysis using the SOI one standard deviation threshold, 1979-2004. Numbers in brackets indicate number of events used to calculate each composite plot.....	156
Fig. 5.11 DJF MSLP anomaly composites for El Niño (left) and La Niña events (right), from reanalysis using the Niño 3.4 one standard deviation threshold, 1979-2004.....	157
Fig. 5.12 Surface temperature anomaly composites for El Niño events (left) and La Niña events (right) from reanalysis using SOI one standard deviation, 1979-2004.	157
Fig. 5.13 Mean DJF surface temperature anomaly from reanalysis during years with Niño 3.4 one standard deviation defined El Niño (left) and La Niña events (right), 1979-2004. Black rectangle indicates the Niño 3.4 region.....	158
Fig. 5.14 Mean zonal wind anomaly (ms^{-1}) composites for El Niño events (left) and La Niña events (right) from reanalysis using the one standard deviation of the SOI, 1979-2004.	158
Fig. 5.15 Mean DJF zonal wind anomaly (ms^{-1}) composite during years with Niño 3.4 one standard deviation defined El Niño (left) and La Niña events (right), 1979-2004.....	159
Fig. 6.1 Observed annual mean SST (Extended Reconstructed Sea Surface Temperature, ERSSTv3b), black boxes indicate regions of common SST bias within CMIP5 model simulations (Michael et al. 2013).....	164
Fig. 6.2 Multi-model mean simulation of 24 CMIP5 CGCMs minus observations (Hadley Centre Sea Ice and SST dataset - HadISST) of annual mean SST over the tropical Pacific (Li et al. 2016).	164
Table 6.1 The five historical climate model simulations used in this thesis: Model name, institution, atmospheric resolution, time span and reference.	169
Table 6.2 Forcings used as boundary conditions for Last Millennium simulations (adapted from Table 2 in Hope et al. 2016).	170
Fig. 6.3 DJF MSLP climatology (hPa) 1979-2004 from historical climate model simulations (left) and models’ MSLP climatology minus ERA-Interim reanalysis (right). Black rectangles indicate the western and eastern regions used to define SOI from the reanalysis and model simulations.....	173
Fig. 6.4 DJF MSLP (hPa) in the Darwin (top) and Tahiti (bottom) regions, 1900-2004 from instrumental record and historical climate model simulations.....	174
Table 6.3 MSLP from Darwin and Tahiti box regions, and MSLP gradient between the two regions, 1979-2004 and 1900-1999 DJF seasonal average from instrumental record, ERA-reanalysis and historical climate model simulations.	175
Table 6.4 Results from t-test comparing DJF MSLP from ERA-Interim of the datasets listed below, for Darwin and Tahiti, 1979-2004. Bold indicates where the null hypothesis can be rejected at the 95% confidence level.....	176

Fig. 6.5 The DJF SOI from historical climate model simulations and the DJF instrumental SOI 1900-2004 (calculated from a 1.5° x 1.5° latitude-longitude grid), normalised over 1979-2004.....	177
Table 6.5 Pearson’s correlation coefficient DJF SOI from the instrumental and climate model simulations over different analysis periods.	178
Table 6.6 Standard deviation of DJF SOI from the instrumental record, ERA-reanalysis and the historical climate model simulations 1979-2004 and 1900-1999.	179
Table 6.7 Number of El Niño and La Niña events defined by the instrumental, reanalysis and historical climate models using the threshold of one standard deviation from the DJF SOI to define events, 1900-1999.	180
Fig. 6.6 Mean DJF MSLP composite anomaly from ERA-Interim reanalysis and climate model simulations during years with SOI one standard deviation defined El Niño (left) and La Niña events (right), 1979-2004. Anomaly calculated from 1979-2004 DJF mean. Numbers in brackets indicate number of events used to calculate each composite plot. Black rectangles indicate the western (Darwin) and eastern (Tahiti) regions used to define the SOI from the reanalysis and model simulations.	182
Fig. 6.7 Mean DJF surface temperature composite anomaly from ERA-Interim reanalysis and climate model simulations during years with SOI one standard deviation defined El Niño (left) and La Niña events (right), 1979-2004. Anomaly calculated from 1979-2004 DJF mean. Numbers in brackets indicate number of events used to calculate each composite plot. Black rectangle indicates Niño 3.4 region.....	183
Table 6.8 Ranking of historical model simulations for SOI-based metrics compared to the DJF SOI instrumental record (BoM), (1900-1999).	185
Fig. 6.8 DJF surface temperature climatology, 1979-2004 from historical climate model simulations (left) and Model climatology minus ERA-Interim reanalysis (right). Black rectangle indicates the Niño 3.4 region.	186
Fig. 6.9 The DJF Niño 3.4 from the historical climate model simulations and ERSST, 1900-2004 (Anomalies shown in Appendix Fig. C5).....	188
Table 6.9 Pearson’s correlation coefficient DJF Niño 3.4 from the ERSST and climate model simulations over different analysis periods, bold indicates significance the 95% confidence level.	188
Table 6.10 Reanalysis and modelled DJF Niño 3.4 mean surface temperature and standard deviation, DJF 1979-2004 and 1900-1999.	189
Table 6.11 Number of El Niño and La Niña events defined by the instrumental, reanalysis and historical climate models using the threshold of one standard deviation from the Niño 3.4 anomaly (1979-2004 base period) to define events, 1900-1999.	190
Fig. 6.10 Mean DJF MSLP composite anomaly from ERA-Interim reanalysis and climate model simulations during El Niño (left) and La Niña events (right) years defined by one standard deviation of the Niño 3.4 index, 1979-2004. Anomaly calculated from 1979-2004 DJF mean. Numbers in brackets indicate number of	

events used to calculate each composite plot. Black rectangles indicate the western (Darwin) and eastern (Tahiti) regions used to define the SOI from the reanalysis and model simulations. 191

Fig. 6.11 Mean DJF surface temperature anomaly from ERA-Interim reanalysis and climate model simulations during El Niño (left) and La Niña events (right) years defined by one standard deviation of the Niño 3.4 index, 1979-2004. Anomaly calculated from 1979-2004 DJF mean. Numbers in brackets indicate number of events used to calculate each composite plot. Black rectangle indicates Niño 3.4 region. 192

Table 6.12 Ranking of historical model simulations for Niño 3.4 metrics compared to ERRST (1900-1999). 193

Fig. 6.12 The DJF Niño 3 (top) and Niño 4 (bottom) from HadISST and ERA-20C reanalysis, 1900-1999 (1.5° x 1.5° grid), Pearson’s correlation (r) between the indices from the two datasets is shown. 195

Fig. 6.13 The DJF Niño 3 index (top) and Niño 4 index (bottom) from HadISST and climate model historical simulations, 1900 -1999 (1.5° x 1.5° grid). 196

Table 6.13 Number of EP and CP El Niño events from reanalysis and historical climate models, 1900-1999. 198

Fig. 6.14 Mean DJF MSLP anomaly from ERA-20C reanalysis and climate model simulations during EP (left) and CP El Niño events (right), 1900-1999. Anomalies based on 1900-1999 DJF mean. Numbers in brackets indicate number of events used to calculate each composite plot. 198

Fig. 6.15 Mean DJF surface temperature anomaly from ERA-Interim and climate model simulations during EP (left) and CP El Niño events (right), 1900-1999. Anomalies based on 1900-1999 DJF mean. Black rectangle indicates Niño 3 (left) and Niño 4 (right) regions. Numbers in brackets indicate number of events used to calculate each composite plot. 200

Table 6.14 Ranking of historical models compared to Niño 3 and Niño 4 metrics from HadISST (1900-1999). 201

Fig. 7.1 Surface wind vector observations for selected years (1° latitude x 1° longitude grid) from ICOADS 2.5.1 used in ERA-20C. Fig. from Poli et al. (2013). 208

Fig. 7.2 Global wind zones at 1000mb in January, with the zero line representing the boundary between westerly and easterly zonal winds. Based on NCEP/NCAR Reanalysis 1970-1999, Fig. from Barry and Chorley (2010). 209

Fig. 7.3 Spatial correlation coefficients between the Niño 3.4 index and 20CR precipitation (left) and CMIP5 7 models MMM (right), model years 1976-2005 (SONDJF). Rectangular boxes indicate East, Central and West Pacific regions (Lewis and LeGrande 2015). 212

Fig. 7.4 Correlation between Sept-Feb rainfall and the Niño 3.4 index for six CIMP5 simulations, 1900-2000, (a-f), and observations, 1979-2008 (g). All linearly detrended before correlations were calculated. r = spatial pattern correlation coefficient between each model and the observed correlation patterns in the domain (Brown et al. 2016). 213

Fig. 7.5 Mean DJF zonal wind (ms^{-1}) 1979-2004 from ERA-Interim, ERA-20C and 20CR reanalysis at higher ($1.5^\circ \times 1.5^\circ$) resolution. Black boxes indicate grid boxes used in logbook-based reconstruction, with grid box 1 in the Atlantic Ocean, and grid boxes 2-4 from west to east around South Africa and west to east 5-9 in the Indian Ocean.	216
Fig. 7.6 DJF zonal wind (ms^{-1}) from logbook grid boxes, ERA-Interim, ERA-20C and 20CR, 1979-2004 ($7.5^\circ \times 8^\circ$ resolution).	217
Fig. 7.7 DJF zonal wind (ms^{-1}) for the nine logbook grid boxes from ERA-Interim (1979-2004), ERA-20C and 20CR reanalysis, 1900-2004.	218
Table 7.1 Pearson' correlation coefficient between mean DJF zonal wind in re-gridded logbook grid boxes, from ERA-20C and 20CR when correlated with ERA-Interim over 1979-2004.	219
Table 7.2 Pearson's correlation coefficient DJF zonal wind from ERA-20C and 20CR.	220
Fig. 7.8 DJF zonal wind (ms^{-1}) climatology, 1979/80-2004/05 from historical climate model simulations (left) and model zonal wind bias (Model climatology minus ERA-Interim reanalysis) (right).	221
Fig. 7.9 DJF Mean zonal wind (ms^{-1}) per grid box (1-9), 1900-1999 from ERA-20C and climate model simulations at $7.5^\circ \times 8^\circ$ latitude-longitude.	223
Fig. 7.10 The DJF mean zonal wind (ms^{-1}) in logbook predictor grid boxes, 1-9 ($7.5^\circ \times 8^\circ$) from ERA-20C and climate model simulations, 1900-1999.	224
Fig. 7.11 Pearson's correlation coefficients between mean DJF zonal wind and DJF SOI from ERA-Interim (top) and ERA-20C (below), 1979-2004. Data detrended prior to calculations of correlations. Black rectangles indicate logbook grid boxes. All values shown.	227
Fig. 7.12 Pearson's correlation coefficient between DJF Zonal wind per grid box (1-9) and DJF SOI from reanalysis. Dashed lines indicates 90% confidence level. Time series detrended prior to calculations of correlations.	227
Table 7.3 Metrics used to rank the models against each other when compared to ERA-20C.	229
Table 7.4 SOI-wind correlations ranks compared to ERA-20C, 1979-2004.	230
Fig. 7.13 Pearson's correlation coefficients between mean DJF zonal wind and DJF SOI from ERA-20C 1900-1999. Data detrended prior to calculations of correlations. Black rectangles indicate logbook regions. All values shown.	231
Fig. 7.14 Pearson's correlation coefficient between DJF zonal wind and DJF SOI from climate model simulations, 1900-1999. All time series detrended prior to calculation of correlations. Black rectangles indicate logbook grid boxes.	232
Fig. 7.15 Pearson's correlation coefficient between DJF Zonal wind per grid box (1-9) and DJF SOI from reanalysis and climate model simulations. Dashed lines indicates the 95% confidence value. Time series detrended prior to calculations of correlations.	233
Table 7.5 Number of logbook grid boxes with significant correlation (95% confidence level) of correct sign, and the number of grid boxes with the same sign correlation as ERA-20C, 1900-1999, between DJF zonal wind and the SOI.	234

Table 7.6 SOI-wind correlations ranks compared to ERA-20C, 1900-1999.	234
Fig. 7.16 30-year sliding/running correlation of DJF zonal wind and DJF SOI 1900-1999, ERA-20C. All time series were detrended before calculating running correlations. The year shown is the 15 th year of the window, dashed line indicates significance at the 90% confidence level.	236
Fig. 7.17 30-year sliding/running correlation of DJF zonal wind and DJF SOI 1900-1999, GISS-E2-R (orange) and IPSL-CM5A-LR (green) climate model simulations and ERA-20C (black). All time series were detrended before calculating running correlations. The year shown is the 15 th year of the window, dashed lines indicate significance at the 90% confidence levels.	238
Fig. 8.1 ENSO indices derived from a number of Last Millennium climate model simulations (850-1850 AD), based on the EOF of the DJF Niño 3.4 index. The grey bars indicate major volcanic events of the LM (1258, 1452, 1815). Fig. from Le (2017).	244
Fig. 8.2 Power spectrum density of GISS-E2-R (upper) and IPSL-CM5A-LR (lower) for the Niño 3.4 index from LM simulations (850-1850). Solid red line indicates model and the faded red shading is the 95% confidence levels. Fig. from Hope et al. (2016).	245
Fig. 8.3 30-year running correlations of Sept-Feb rainfall and the Niño 3.4 index (blue) and the 30-year running variance of the Niño 3.4 index (red). Correlations between the time series shown in the top-right corner of each panel. Fig. from Brown et al. (2016).	247
Fig. 8.4 The DJF SOI from GISS-E2-R (upper) and IPSL-CM5A-LR (lower) LM simulations relative to the 1750/51-1848/49 mean. Bold line indicates 30-year moving mean. Vertical lines indicate major volcanic eruptions 1258, 1452, 1815 as in Brown et al. (2016).	251
Fig. 8.5 Number of SOI-defined ENSO events per 100-year period of LM climate model simulations from GISS-E2-R (upper) and IPSL-CM5A-LR (lower), plus the number of events from the instrumental SOI over the twentieth century, indicated by data outlined in black. The 100-year period for the MWP is outlined in red, while the 200-years covering the LIA is outlined in light blue.	253
Fig. 8.6 El Niño and La Niña surface temperature (K) anomaly composites defined by one standard deviation of the DJF SOI from historical simulations, 1900-1999 from GISS-E2-R and IPSL-CM5A-LR. Anomalies based on 1900-1999 DJF mean. Numbers in brackets indicate number of events used to calculate each composite plot.	255
Fig. 8.7 El Niño and La Niña surface temperature (K) anomaly composites defined by one standard deviation of the DJF SOI from LM simulations, 1750-1848 from GISS-E2-R and IPSL-CM5A-LR. Numbers in brackets indicate number of events used to calculate each composite plot.	255
Fig. 8.8 El Niño and La Niña MSLPA (hPa) composites defined by one standard deviation of the DJF SOI from historical simulations, 1900-1999 from GISS-E2-R and IPSL-CM5A-LR. Anomalies based on 1900-1999 DJF mean. Numbers in brackets indicate number of events used to calculate each composite plot.	256

Fig. 8.9 El Niño and La Niña MSLPA (hPa) composites defined by one standard deviation of the DJF SOI from LM simulations, 1750-1848 from GISS-E2-R and IPSL-CM5A-LR. Numbers in brackets indicate number of events used to calculate each composite plot.....	257
Fig. 8.10 El Niño and La Niña zonal wind (ms^{-1}) anomaly composites defined by one standard deviation of the DJF SOI from historical simulations, 1900-1999 from GISS-E2-R and IPSL-CM5A-LR. Anomalies based on 1900-1999 DJF mean. Numbers in brackets indicate number of events used to calculate each composite plot.	258
Fig. 8.11 El Niño and La Niña zonal wind (ms^{-1}) anomaly composites defined by one standard deviation of the DJF SOI from LM simulations, 1750-1848 from GISS-E2-R and IPSL-CM5A-LR. Numbers in brackets indicate number of events used to calculate each composite plot.	258
Fig. 8.12 Pearson’s correlation coefficient between the mean DJF zonal wind and the DJF SOI from GISS-E2-R and IPSL-CM5A-LR for the years 1750-1848 of the LM simulations. Data are detrended prior to calculation of correlations and only those correlations significant at 90% confidence level are shown.	260
Fig. 8.13 As in Fig. 8.12 but for the years 1050-1149 from the two LM simulations.	261
Fig 8.14 As in Fig. 8.12 but for the years 1900-1999 from the two historical simulations.	261
Table 8.1 Pearson’s correlation coefficient between the detrended DJF SOI and the detrended DJF zonal wind in the predictor grid boxes ($1-9; 7.5^\circ \times 8^\circ$ resolution), from LM simulations, 1750-1848 and 1050-1149, and the historical simulations, 1900-1999. Bold indicates significance at the 90% confidence level.....	262
Fig. 8.15 30-year running Pearson’s correlation coefficients between mean DJF zonal wind and DJF SOI from GISS-E2-R LM climate model simulations 1750-1848. Data detrended prior to calculations of correlations. The year shown is the 15 th year of the window. Dashed lines indicate significance at the 90% confidence level.	264
Fig. 8.16 30-year running Pearson’s correlation coefficients between mean DJF zonal wind and DJF SOI IPSL-CM5A-LR climate model simulations 1750-1848. Data detrended prior to calculations of correlations. The year shown is the 15 th year of the window, dashed lines indicate significance at the 90% confidence level.....	265
Table 8.2 Standard deviation of 30-year running correlation over the LM period in each predictor grid box.....	266
Fig. 8.17 The DJF Niño 3 and DJF Niño 4 surface temperature from GISS-E2-R and IPSL-CM5A-LR LM simulation, 850-1848. Bold lines indicate 30-year running mean.	268
Fig. 8.18 Number of EP and CP El Niño events per 100-year period of LM simulation of GISS-E2-R (top) and IPSL-CM5A-LR (bottom), and the number of events from ERA-20C in the twentieth century, as outlined in black.....	269

Fig. 8.19 The CP-EP Ratio for the ten 100-year periods of the LM simulations of GISS-E2-R and IPSL-CM5A-LR, and that from ERA-20C reanalysis over the last century shown in black.	270
Fig. 8.20 EP and CP El Niño surface temperature (K) anomaly composites defined by the NINO method from LM simulations, 1750-1848 from GISS-E2-R and IPSL-CM5A-LR. Numbers in brackets indicate number of events used to calculate each composite plot.....	270
Fig. A1 Pearson’s correlation coefficient between de-trended mean local noon DJF zonal wind and the DJF Niño 3.4 index, 1979/80-2013/14. Only those significant at the 90% confidence level are shown. Grid boxes used in the reconstructions are outlined and are numbered 1 to 9 from West to East.....	311
Fig. A2 Numbered labels for the nine predictor grid boxes used in the logbook-based reconstruction.....	311
Table A1 Location and labels of the nine predictor grid boxes used in the logbook-based reconstructions.	312
Fig. A3 Number of days with both wind force and direction observations from logbooks per DJF season in the nine predictor grid boxes, 1750/51-1853/54.....	312
Table A2 The DJF SOI reconstructions for 1788/89, 1791/92, 1802/03 and 1811/12, using PCR and CPS. Climatology taken from 1834-1853.....	313
Table A3 Linear regression equations for the four historical reconstructions and their R^2 values, statistically significant ($p < 0.05$) trends are shown in bold.	313
Table A4 Mean normalised SOI and two-sampled t-test results for the two periods, 1815-1833 (Time period I) and 1834-1853 (Time period II) for the four SOI reconstructions.	313
Table A5 Pearson’s correlation coefficient between logbook DJF SOI reconstructions and the two early instrumental Jakarta based SOI records. Those significant at the 95% confidence level are indicated in bold.	314
Table A6 The logbook-based DJF SOI reconstructions, 1815/16 – 1853/54. NaN indicated years with missing data.	315
Table B1 Two sampled t-test comparing the DJF SOI from the instrumental record to that from the three reanalysis products, the DJF SOI is relative to the 1979-2004 mean. Comparisons made over four periods; 1979-2004, 1900-1949, 1950-1999 and 1900-1999.	316
Table B2 The years of El Niño and La Niña events defined from by one standard deviation of the DJF SOI from the instrumental record and the two historical reanalysis, 1900-1999. Bold indicates events found in the instrumental record and the selected reanalysis.....	316
Fig. B1 The DJF Niño 3.4 anomaly from the three reanalyses (ERA-20C, 20CR and ERA-Interim) and ERSSTv4, 1900-2004. Anomaly relative to the 1979-2004 base period.	317
Table B3 Two sampled t-test comparing the DJF Niño 3.4 index from the instrumental record to that from the three reanalysis products, the DJF Niño 3.4	

index is relative to the 1979-2004 mean. Comparisons made over four periods; 1979-2004, 1900-1949, 1950-1999 and 1900-1999.	317
Table B4 The years of El Niño and La Niña events defined from by one standard deviation of the DJF Niño 3.4 index from the instrumental record (ERSST) and the two historical reanalysis, 1900-1999. Bold indicates events found in the instrumental record and the selected reanalysis.	318
Fig. B2 DJF MSLP (hPa) composite anomaly for El Niño events (left) and La Niña events (right) from the two historical reanalysis using the SOI one standard deviation threshold, 1900-1999. Numbers in brackets indicate the number of events used to calculate each composite plot.	318
Fig. B3 DJF MSLP (hPa) composites anomaly for El Niño events (left) and La Niña events (right) from the two historical reanalysis using the Niño 3.4 index one standard deviation threshold, 1900-1999. Numbers in brackets indicate the number of events used to calculate each composite plot.	319
Fig. B4 Surface temperature (K) composite anomaly for El Niño events (left) and La Niña events (right) from the two historical reanalysis using SOI one standard deviation, 1900-1999. Numbers in brackets indicate the number of events used to calculate each composite plot.	319
Fig. B5 Surface temperature (K) composite anomaly for El Niño events (left) and La Niña events (right) from the two historical reanalysis using the Niño 3.4 index one standard deviation, 1900-1999. Numbers in brackets indicate the number of events used to calculate each composite plot.	319
Fig. B6 Zonal wind (ms^{-1}) composite anomaly for El Niño events (left) and La Niña events (right) from the two historical reanalysis using the SOI one standard deviation threshold, 1900-1999. Numbers in brackets indicate the number of events used to calculate each composite plot.	320
Fig. B7 Zonal wind (ms^{-1}) composite anomaly for El Niño events (left) and La Niña events (right) from the two historical reanalysis using the Niño 3.4 index one standard deviation threshold, 1900-1999. Numbers in brackets indicate the number of events used to calculate each composite plot.	320
Fig. C1 DJF MSLP (hPa) composite anomaly from ERA-20C reanalysis and climate model simulations during El Niño (left) and La Niña events (right) defined using the SOI one standard deviation, 1900-1999. Anomalies based on 1900-1999 DJF mean. Numbers in brackets indicate the number of events used to calculate each composite plot.	321
Fig. C2 Mean DJF surface temperature (K) composite anomaly from ERA-20C reanalysis and climate model simulations during El Niño (left) and La Niña events (right) defined using the SOI one standard deviation, 1900-1999. Anomalies based on 1900-1999 DJF mean. Numbers in brackets indicate the number of events used to calculate each composite plot.	322
Fig. C3 Mean DJF MSLP (hPa) composite anomaly from ERA-20C reanalysis and climate model simulations during El Niño (left) and La Niña (right) years defined using the Niño 3.4 one standard deviation, 1900-1999. Anomalies based on 1900-	

1999 DJF mean. Numbers in brackets indicate the number of events used to calculate each composite plot.....	323
Fig. C4 Mean DJF surface temperature (K) anomaly from ERA-20C reanalysis and climate model simulations during El Niño (left) and La Niña events (right) defined using the SOI one standard deviation, 1900-1999. Anomalies based on 1900-1999 DJF mean. Numbers in brackets indicate the number of events used to calculate each composite plot.....	324
Table C1 A two-sampled t-test comparing the DJF SOI from the instrumental record (BoM) to that from the climate model simulations, 1979-2004 and 1900-1999. Bold indicates that the null hypothesis cannot be rejected at the 95% confidence level..	324
Table C2 Ranking of historical model simulations for SOI-based metrics compared to the DJF SOI instrumental record (BoM) (1979-2004).	325
Table C3 Ranking of historical model simulations for SOI-based metrics compared to the DJF SOI from ERA-20C (1900-1999).....	325
Fig. C5 The DJF Niño 3.4 Anomaly (K) (1979-2004 base period) from historical climate model simulations and ERSST, 1900-2004.	326
Table C4 Results from t-test comparing the Niño 3.4 anomaly indices from ERSST to the climate model simulations, 1979-2004 and 1900-1999. Bold indicates that the null hypothesis cannot be rejected at the 95% confidence level.....	326
Table C5 Ranking of historical model simulations for Niño 3.4 metrics compared to ERSST (1979-2004).	327
Table D1 Two sample t-test to compare DJF zonal mean climatologies from ERA-Interim and ERA-20C/20CR for each grid box (1-9), 1979-2004. An H value of 0 indicates that the null hypothesis cannot be rejected at the 95% level.	328
Fig. D1 DJF Mean zonal wind (ms^{-1}) per grid box (1-9), 1979-2004 from reanalysis and historical climate model simulations at $7.5^\circ \times 8^\circ$ resolution.	328
Table D2 Two sample t-test to analyse difference between model simulation DJF zonal mean climatologies compared against ERA-Interim for each grid box (1-9), 1979-2004. Bold numbers indicate the number of logbook grid boxes where the null hypothesis cannot be rejected at the 95% level.	329
Table D3 Two sample t-test to analyse difference between model simulation DJF zonal mean climatologies compared against ERA-20C for each grid box (1-9), 1900-1999. Bold numbers indicate the number of logbook grid boxes where the null hypothesis cannot be rejected at the 95% level.	329
Fig. D2 Pearson's correlation coefficient between DJF zonal wind and the DJF SOI from climate model simulations, 1979-2004. All time series detrended prior to calculation of correlations. Black rectangles indicate logbook predictor grid boxes.	330
Fig. D3 Pearson's correlation coefficient between DJF zonal wind per grid box (1-9) and the DJF SOI from reanalysis and climate model simulations. Dashed lines indicates 90% significance value. Time series detrended prior to calculations of correlations.....	331
Table D4 Number of logbook grid boxes with significant correlation (90% confidence level) of correct sign, and the number of grid boxes with the same sign	

correlation as ERA-Interim, 1979-2004, between the DJF zonal wind and the DJF SOI.	331
Fig. D4 Pearson’s correlation coefficients between mean DJF zonal wind and the DJF Niño 3.4 index from ERA-Interim (top) and ERA-20C (below), 1979-2004. Data detrended prior to calculations of correlations. Black rectangles indicate logbook regions. All values shown.	332
Fig. D5 Pearson’s correlation coefficient between DJF zonal wind per grid box (1-9) and the DJF Niño 3.4 index from reanalysis. Dashed lines indicates significance at the 90% confidence level. Time series detrended prior to calculations of correlations.	333
Fig. D6 Pearson’s correlation coefficient between DJF zonal wind and the DJF Niño 3.4 index from climate model simulations, 1979-2004. All time series were detrended prior to calculation of correlations. Black rectangles indicate logbook grid boxes.	334
Fig. D7 Pearson’s correlation coefficient between DJF zonal wind per grid box (1-9) and the DJF NINO3.4 index from reanalysis and climate model simulations. Dashed lines indicates significance at the 90% confidence level. Time series detrended prior to calculations of correlations.	335
Table D5 Number of logbook grid boxes with significant correlation (90% confidence level) of correct sign, and the number of grid boxes with the same sign correlation as ERA-Interim, 1979-2004, between DJF zonal wind and the DJF Niño 3.4 index.	335
Fig. D8 Pearson’s correlation coefficients between mean DJF zonal wind and the DJF Niño 3.4 index from ERA-20C 1900-1999. Data detrended prior to calculations of correlations. Black rectangles indicate logbook regions. All values shown.....	336
Fig. D9 Pearson’s correlation coefficient between DJF zonal wind and the DJF Niño 3.4 index from climate model simulations, 1900-1999. All time series detrended prior to calculation of correlations. Black rectangles indicate logbook predictor grid boxes.	337
Fig. D10 Pearson’s correlation coefficient between DJF zonal wind per grid box (1-9) and the DJF Niño 3.4 index from reanalysis and climate model. Dashed lines indicates significance at the 90% confidence level. Time series detrended prior to calculations of correlations.	338
Table D6 Number of logbook grid boxes with significant correlation (90% confidence level) of correct sign, and the number of grid boxes with the same sign correlation as ERA-20, 1900-1999, between DJF zonal wind and the DJF Niño 3.4 index.....	338
Table D7 Ranking of historical climate models for zonal wind and DJF Niño 3.4 index metrics compared to the ERA-20C, 1900-1999.	339
Fig. D11 30-year running correlation of DJF zonal wind and DJF SOI 1900-1999, 20CR. All time series were detrended before calculating running correlations. The year shown is the 15 th year of the window, dashed line indicates significance at the 90% confidence level.	339

Fig. D12 30-year running correlation of DJF zonal wind and the DJF SOI 1900-1999, climate model simulations and ERA-20C. All time series were detrended before calculating running correlations. The year shown is the 15 th year of the window, dashed line indicates significance at the 90% confidence level.....	340
Fig. E1 The DJF SOI from GISS-E2-R LM simulation in the 20 years surrounding the three major volcanic events of the LM, 1258, 1452 and 1815 AD. Vertical lines indicate years of eruption.....	341
Fig. E2 The DJF SOI from IPSL-CM5A-LR LM simulation in the 20 years surrounding the three major volcanic events of the LM, 1258, 1452 and 1815 AD. Vertical lines indicate years of eruption.	342
Table E1 Number of El Niño and La Niña events during 100-year periods from Last Millennium simulations. (based on one standard deviation of the SOI from the mean of the 100-year period; SOI normalised over 1750-1848), and ENSO frequency (years).	343
Fig. E3 El Niño and La Niña surface temperature (K) anomaly composites defined by one standard deviation of the DJF SOI from LM simulations, 1050-1149 from GISS-E2-R and IPSL-CM5A-LR. Numbers in brackets indicate the number of events used to calculate each composite plot.....	343
Fig. E4 El Niño and La Niña MSLPA (hPa) composites defined by one standard deviation of the DJF SOI from LM simulations, 1050-1149 from GISS-E2-R and IPSL-CM5A-LR. Numbers in brackets indicate the number of events used to calculate each composite plot.	344
Fig. E5 El Niño and La Niña zonal wind (ms ⁻¹) anomaly composites defined by one standard deviation of the DJF SOI from LM simulations, 1050-1149 from GISS-E2-R and IPSL-CM5A-LR. Numbers in brackets indicate the number of events used to calculate each composite plot.	344
Fig. E6 Pearson's correlation coefficient between the mean DJF zonal wind and the DJF SOI from GISS-E2-R and IPSL-CM5A-LR 1750-1848, covering the logbook-based reconstruction period. All correlations shown. Calculated at 1.5° x 1.5° degree resolution.....	345
Fig. E7 Pearson's correlation coefficient between the mean DJF zonal wind and the DJF SOI from GISS-E2-R and IPSL-CM5A-LR 1750-1848, covering the logbook-based reconstruction period. All correlations shown. Calculated at 7.5° x 8° degree resolution.....	346
Fig. E8 Pearson's correlation coefficient between the mean DJF zonal wind and the DJF SOI from GISS-E2-R and IPSL-CM5A-LR for the 1050-1149 of the LM simulations covering the logbook-based reconstruction period. All correlations shown. 1.5° x 1.5° degree resolution.....	347
Fig. E9 Pearson's correlation coefficient between the mean DJF zonal wind and the DJF SOI from GISS-E2-R and IPSL-CM5A-LR for the 1050-1149 of the LM simulations covering the logbook-based reconstruction period. All correlations shown. 7.5° x 8° degree resolution.....	348

Table E2 Number of EP and CP El Niño events during 100-year periods from the Last Millennium simulations (based on one standard deviation from mean of 100-year period).	349
Fig. E10 EP and CP El Niño surface temperature (K) anomaly composites defined by one standard deviation of NINO method from the LM simulations, 1050-1149 from GISS-E2-R and IPSL-CM5A-LR. Numbers in brackets indicate the number of events used to calculate each composite plot.	349

List of Abbreviations

20CR – The Twentieth Century Reanalysis Project

ACRE – Atmospheric Circulation Reconstructions over the Earth

BF3 – Beaufort Force 3

BoM – Bureau of Meteorology

CLIWOC – Climatological Database for the World's Oceans

CMIP– Climate Model Intercomparison Project

CP – Central Pacific

CPS – Composite Plus Scale

ECMWF – European Centre for Medium-Range Weather Forecasts

EEIC – English East India Company

ENSO – El Niño Southern Oscillation

EP – Eastern Pacific

ERSST – Extended Reconstructed Sea Surface Temperature

GISS- Goddard Institute for Space Studies

HadC – Hadley Centre

HadISST – Hadley Centre Sea Ice and Sea Surface Temperature dataset

ICOADS – International Comprehensive Ocean-Atmosphere Data Set

IPCC – Intergovernmental Panel on Climate Change

IPSL- Institut Pierre Simon Laplace

ISMR – Indian Summer Monsoon Rainfall

ISPD – International Surface Pressure Databank

LIA – Little Ice Age

LM – Last Millennium

MADA – Monsoon Asia Drought Atlas

MIROC- Model for Interdisciplinary Research on Climate

MMM – Multi-Model Mean

MPI- Max Plank Institute

MSLP(A) - Mean Sea Level Pressure (Anomaly)

MWP – Medieval Warm Period

NADA – North America Drought Atlas

NOAA – National Oceanic and Atmospheric Administration

PCR – Principal Component Regression

PMIP – Paleoclimate Modelling Intercomparison Project

PWC – Pacific Walker Circulation

RE – Reduction of Error

RECLAIM – Recovery of Logbooks and International Marine data

SLP – Sea Level Pressure

SODAsi – Simple Ocean Data Assimilation – sparse input

SOI – Southern Oscillation Index

SST(A) – Sea Surface Temperature (Anomaly)

EG13 – Emile-Geay et al. (2013)

GF09 – Gergis and Fowler (2009)

GH08 – Garcia-Herrera et al. (2008)

LI13 – Li et al. (2013)

MANN00 – Mann et al. (2000)

MCG10 – McGregor et al. (2010)

ORT00 – Ortleib (2000)

QUINN87 – Quinn et al. (1987)

ST98 – Stahle et al. (1998)

W10 – Wilson et al. (2010)

1. Introduction

1.1 Introduction

Understanding the climate of the past is vitally important for placing current and future climate changes within a historical context (Gergis and Fowler 2009). El Niño Southern Oscillation (ENSO) is a major mode of climate variability which impacts global climate on interannual timescales. Recent strong ENSO events, such as the 2015/16 El Niño, have highlighted the large socio-economic and environmental impacts of this climatic phenomenon, and thus the importance of understanding its natural variability (Schiermeier 2015). In addition, projected changes in the climate system due to anthropogenic climate change raise questions surrounding the future behaviour of natural modes of climate variability such as ENSO (Cai et al. 2015). This thesis explores ENSO during the pre-instrumental, historical period, focusing on the early-to-mid nineteenth century using a range of non-instrumental data sources. This enables a better understanding of ENSO behaviour on time scales longer than those available from the modern instrumental record.

This chapter provides background information on the current understanding of the El Niño Southern Oscillation, its importance, its impacts and its instrumental record. This thesis uses an underexploited source of historical weather observations to provide new explorations into the history of ENSO. This chapter introduces the observations from ships' logbooks as a source of historical data for climate studies. In addition, the data sources which are commonly used to investigate ENSO in the past are then introduced. These include documentary records, multi-proxy reconstructions and climate model simulations. The key aims of this thesis are then discussed and placed within the context of existing research, and the anticipated outcomes of this research are identified. Finally, the structure of this thesis is presented.

1.2 Mean Pacific climate and El Niño Southern Oscillation

ENSO is the main driver of interannual climate variability within the tropics and its impacts affect weather patterns around the globe (Diaz and Markgraf 1992). It is a coupled atmosphere-ocean phenomenon associated with a see-saw pattern of

atmospheric pressure and sea surface temperatures (SSTs) across the equatorial Pacific with periodicities of two to seven years. It is important to the global climate due to the near-global extent of its teleconnections and the widespread impacts of its two extreme phases, El Niño and La Niña (Barry and Chorley 2010). The El Niño events of 1982/83 and 1997/98 were two of the most extreme events on record, and an unusually prolonged El Niño persisted between 1990 and 1995. More recently, in 2015/16 an El Niño event which was one of the strongest on record caused wide-spread disruption around the globe (Schiermeier 2015). There is currently little agreement as to whether changes in the strength and characteristics of ENSO events in recent decades are due to natural variability or to anthropogenic forcings, and there is still uncertainty regarding how ENSO may change in the future (Christensen et al. 2013).

In order to understand ENSO, an understanding of the mean tropical Pacific climate is needed. The main oceanic and atmospheric features of the tropical Pacific are characterised simply in Fig. 1.1 which, relative to ENSO conditions, are often termed ‘normal conditions’. Fig. 1.1 shows a key component of this mean Pacific climate, the zonal gradient in SST across the equatorial Pacific. In the western Pacific, there is a build-up of relatively warm SST in the Indo-Pacific Warm Pool. The eastern Pacific, especially just off the coast of equatorial South America, is relatively cooler with coastal upwelling bringing cool, deep ocean waters to the surface. This ‘cold tongue’ extends further to the west along the equator. The upwelling of water in the east results in a shallower thermocline in this region and thus a tilted thermocline from east to west. The SST gradient across the equatorial Pacific is a key driver of atmospheric circulation, namely the Walker Circulation. The Walker Circulation is the large scale zonal atmospheric circulation which drives surface winds and promotes regions of suppressed and enhanced convection. Relatively low Mean Sea Level Pressure (MSLP) and relatively high SSTs in the western Pacific and over the Maritime Continent result in the upward movement of moist air, leading to a large amount of convective activity and precipitation in the west. This results in the movement of relatively drier air to the east aloft, and then sinking air over the eastern Pacific associated with high MSLP and suppressed convection (Laing and Evans 2015). To complete this zonal circulation cell, air is transported from the eastern to western Pacific at the surface via the equatorial easterly trade winds. The easterly surface winds, also contribute to the build-up of

warmer waters in the western Pacific and upwelling in the eastern Pacific off the coast of South America. Therefore, the atmosphere and ocean are strongly coupled across this region.

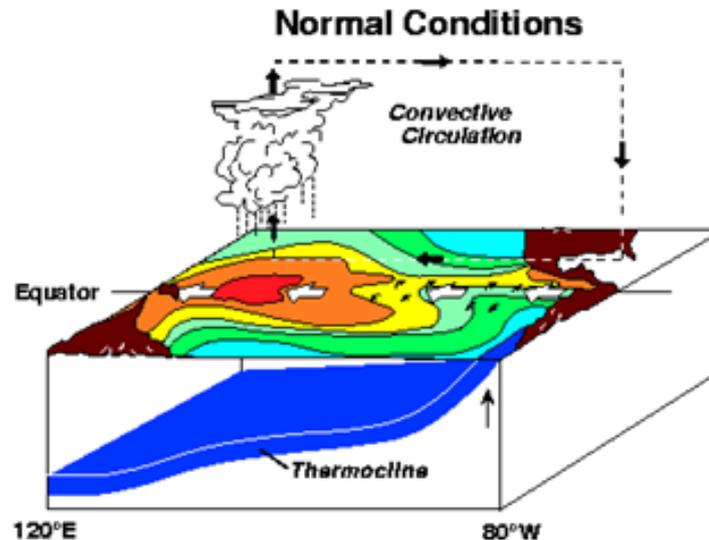


Fig. 1.1 Typical Normal conditions over the equatorial Pacific (National Oceanic and Atmospheric Administration, NOAA, 2017a).

ENSO events occur when these normal atmospheric and oceanic circulation patterns are disturbed. The term El Niño is used to describe an unusually warm current off the coast of Peru, known locally to Peruvian fishermen for centuries. An 1891 article in the Lima Geographical Society bulletin described an unusual counter current observed by sailors which appeared just after Christmas time, the '*corriente del Niño*' (Philander 1990; Wang and Fiedler 2006). A warm southward current resulted in a switch from cold to warm ocean conditions annually in northern Peru. However, in some years the warm current was exceptionally warm and was thought to be linked to unusual climatic and oceanic phenomena (Wang and Fiedler 2006). The occurrence of this unusually warm current on interannual time scales was henceforth described as El Niño and the annual occurrence largely overlooked as focus turned to these anomalously warm events (Wang and Fiedler 2006).

The Southern Oscillation is the atmospheric component of ENSO, first described by Sir Gilbert Walker in 1920s, and is associated with changes in the Walker Circulation (Walker 1932). Walker noted that the zonal circulation cell oscillated, with certain

years seeing a weakening and/or a reversal of the MSLP gradient and surface winds across the equatorial Pacific. He termed these changes from normal conditions the Southern Oscillation. The term El Niño Southern Oscillation with its modern day meaning as a coupled atmosphere-ocean phenomenon was not introduced until 1960s with the work of Professor Jacob Bjerkness, who linked the known atmospheric and oceanic features together and proposed an integrated, coupled phenomenon (Bjerkness 1969). Using air and sea temperature observations from Canton Island in the central Pacific, the links between warmer SST and low MSLP leading to enhanced precipitation were noted and an explanation of the link between observations and the circulation patterns associated with Sir Gilbert Walker's Southern Oscillation was presented (Bjerkness 1969). It was not until the exceptionally strong El Niño event of 1982/83 that ENSO became more extensively studied by the climate community. This El Niño event spurred scientists to launch efforts to improve their understanding of this phenomenon and increase observation and monitoring in hope of increasing prediction skill (Wang and Fiedler 2006). The strong ENSO events of 1997/98 and 2015/16 have been key in raising public awareness of this climatic phenomenon.

ENSO events are associated with changes in SST, MSLP, surface winds and outgoing longwave radiation associated with convective activity. El Niño events occur when 'normal', ENSO-neutral, conditions are disturbed and there is a redistribution of the areas of high and low MSLP anomalies, changes in the SST gradient, and in the strength and the direction of zonal winds, as seen in Fig. 1.2a. During El Niño events, the Indo-Pacific Warm Pool is shifted eastwards and the western Pacific cools by around 1°C at the surface (Gagan et al. 2004). The thermocline slope reduces as the gradient in SST is reduced over the equatorial Pacific. The Walker Circulation is shifted and the equatorial surface trade winds weaken, and in some cases are reversed. The centre of deep convection and high precipitation is shifted eastward, in a region of convergence of surface winds. As a result, high pressure anomalies are found in the western Pacific and over the Maritime Continent. El Niño conditions typically peak in boreal winter and then terminate in boreal spring (Bellenger et al. 2014). El Niño events generally rapidly decay and evolve into opposite phase, La Niña, by the following winter (Zhou et al. 2014). However, not all El Niño events are followed by

a La Niña event, as the tropical Pacific can often revert to normal conditions after an El Niño event.

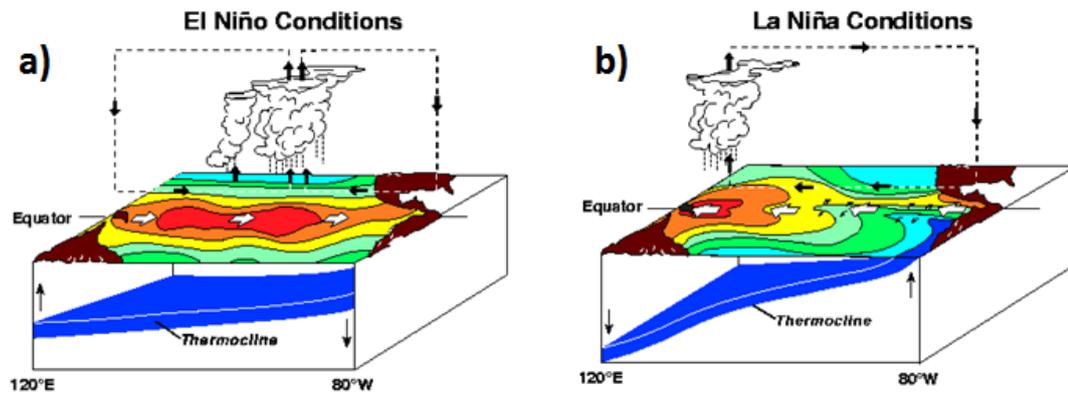


Fig. 1.2 Typical a) El Niño and b) La Niña conditions over the equatorial Pacific (NOAA 2017a).

During La Niña events the normal conditions in the equatorial Pacific are intensified, as seen in Fig. 1.2b. The easterly trade winds are stronger, resulting in enhanced upwelling along the coast of South America and a build-up of warmer waters in the western Pacific, leading to an intensified Indo-Pacific Warm Pool. Due to this increase in SST over the Maritime Continent there is enhanced convection, increased precipitation and low SLP anomalies over the far western Pacific. Due to enhanced upwelling in the east, the thermocline is shallow in this region and the slope of the thermocline increases. La Niña events also generally peak in boreal winter, however after this peak La Niña events can often persist longer than El Niño events, decaying at a slower rate and persisting until the following winter (Zhou et al. 2014).

1.3 ENSO teleconnections and impacts

As well as changes in the atmosphere and the ocean over the equatorial Pacific, ENSO events are associated with changes in weather patterns around the globe, termed ENSO teleconnections. ENSO events are known to have a range of impacts making them an important climatic feature to understand in order to reduce the impacts. Typical ENSO teleconnections are shown in Fig. 1.3 for El Niño and Fig. 1.4 La Niña events, showing great variety in the typically conditions in different areas of the globe during ENSO events. However, it must be noted that all ENSO events are different and have varied

teleconnections (Capotondi et al. 2015). ENSO teleconnections are generally more wide-spread and robust in regions surrounding the Pacific, although it can be seen from Fig. 1.3 ad Fig. 1.4 that more remote regions can still have noticeable teleconnections. There is even a strong ENSO influence in remote regions such as the Antarctic Peninsula and West Antarctica (Lachlan-Cope and Connolley 2006), not shown on Fig. 1.3 or Fig. 1.4.

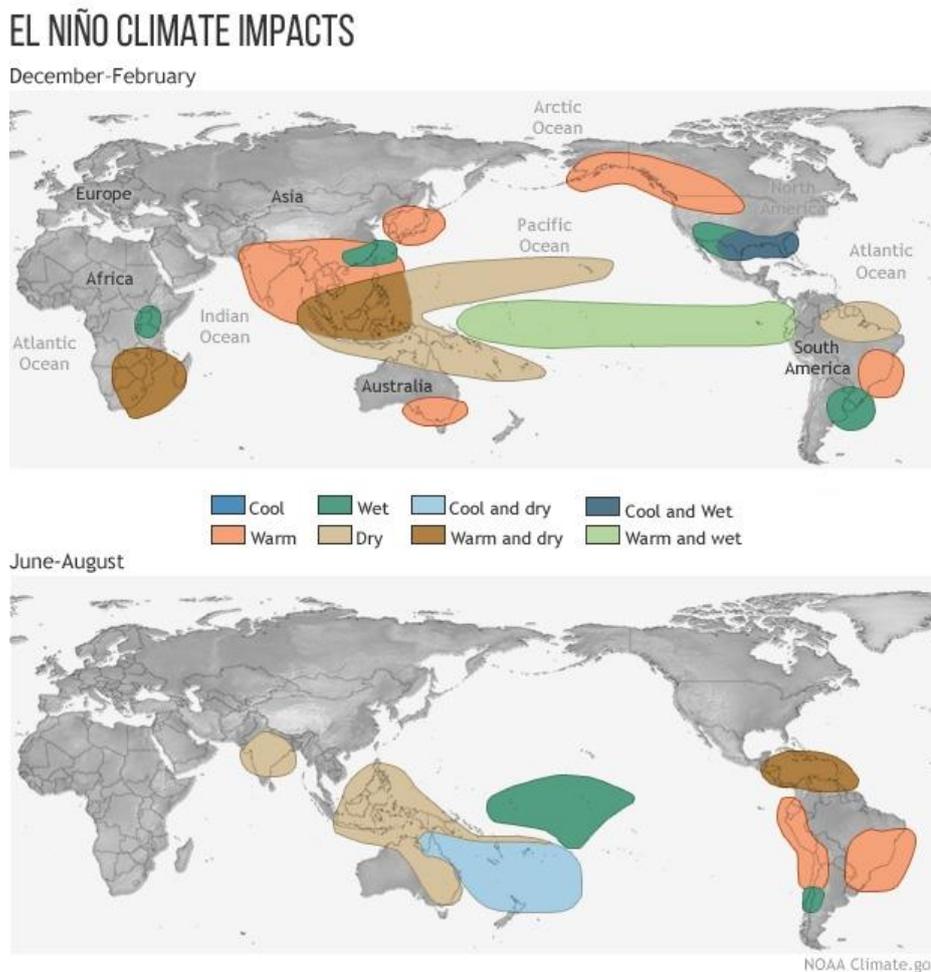


Fig. 1.3 Typical El Niño climate impacts during DJF and JJA seasons (NOAA 2017b).

During El Niño events, it is typically warmer and drier over large regions spanning the western Pacific and the Maritime Continent. This is a result of suppressed convection in this region which can often lead to a range of impacts such as the spread of wildfires and destruction of productive agricultural land. Drought conditions in this region can also impact agricultural productivity. For example, the 2015-16 El Niño was one of the strongest on record and led to drought and extensive wildfires in

Indonesia (Schiermeier 2015). The wild fires can have significant impacts on health, in Malaysia and parts of Indonesia, exacerbated forest fires lead to extreme pollution events, with an unusually high air pollution index within cities, during events such as the 1997/98 El Niño (Khandekar et al. 2000). As seen in Fig. 1.3, during the Indian Monsoon season (June-August) of El Niño years, it is often drier over India as a result of a delayed and/or weakened monsoon; this is a common impact reported during El Niño events (Krishna Kumar et al. 1999). The reduced monsoon has dramatic impacts upon agriculture and water resources, upon which over a billion people rely. In the eastern Pacific, and in coastal South America El Niño events typically lead to warm, wet conditions (Fig. 1.3). The 1997/98 El Niño also resulted in severe flooding and landslides along the west coast of South America, and excessive precipitation in Peru due to the 2015 El Niño resulted in widespread flooding which led to a state of emergency being declared (Schiermeier 2015).

El Niño events, and the extreme weather it is associated with, can have significant economic impacts, such as reduced agriculture output, food shortages and food-price inflation (Cashin et al. 2014). Key commodities such as coffee and cocoa are produced in regions affected by drought or excess precipitation during ENSO events, and thus strong events lead to significant changes in the price of these global commodities (Ubilava 2011). This impact filters around the globe, affecting regions not directly impacted by the extreme weather associated with ENSO. The reduction in coastal upwelling off the coast of Peru during El Niño events can devastate the Peruvian fishing communities as the nutrient-rich waters which normally support thriving marine ecosystems don't upwell to the surface and fish stocks are drastically reduced (Broad et al. 2002). In the western Pacific, the changes in SST also cause huge impacts on the marine ecosystem with strong El Niño events leading to extensive coral bleaching in regions such as the Great Barrier Reef (Ampou et al. 2017).

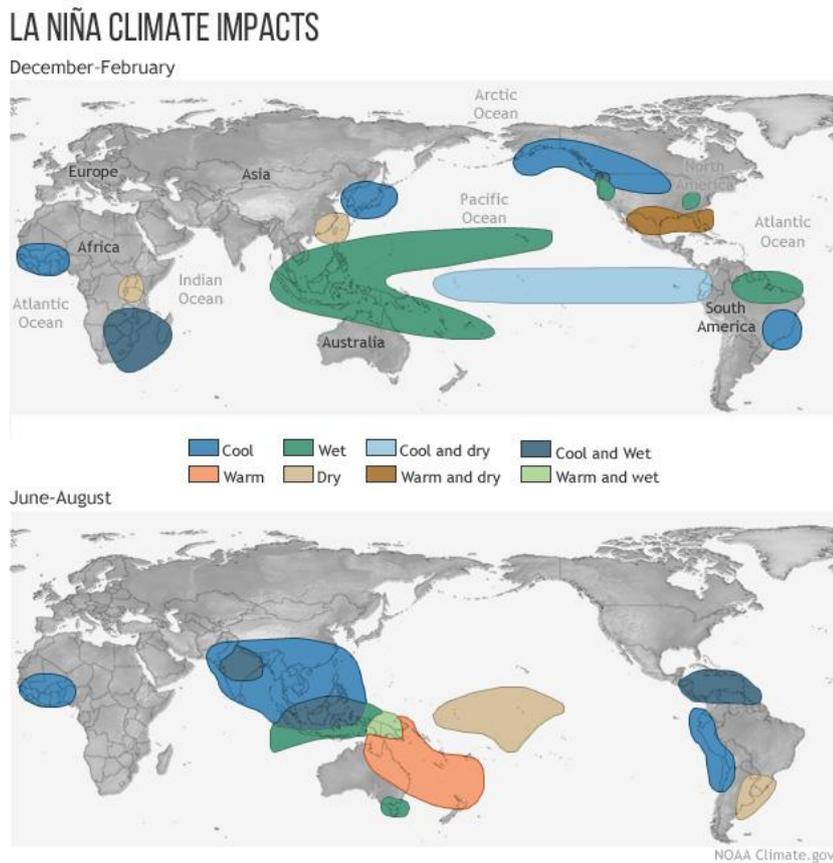


Fig. 1.4 Typical La Niña climate impacts during DJF and JJA seasons (NOAA 2017c).

The teleconnections associated with La Niña events seen in Fig. 1.4 are equally as wide-spread, and generally opposing to El Niño events, having a similar range of impacts including flooding, droughts, landslides and wild-fires. Some regions experience opposite impacts during El Niño and La Niña events, whereas more remote regions can be more strongly impacted by one type of ENSO event than the other. For example, Fig. 1.4 shows cool anomalies in western Africa during La Niña events, but no significant teleconnection during El Niño events. More focus is often given to the impacts of El Niño events because of the recent higher number of strong El Niño events, and because the teleconnections associated with El Niño events tend to be stronger than those for an equally as strong La Niña event due to the asymmetry of ENSO events (Hoerling et al. 1997; Zhang et al. 2014).

The terms El Niño and La Niña cover a wide range of conditions due to the differences in intensity and characteristics of different events (Philander 1990). There is large event-to-event diversity between El Niño events with a range of amplitudes, patterns

and evolutions possible (Capotondi et al. 2015). Recently, there has been increasing research into the different types of El Niño events often termed ‘ENSO flavours’ or ‘ENSO diversity’, and two main types of El Niño event have emerged. The El Niño conditions described above and shown in Fig. 1.2a follow the classical definition of El Niño events, which are termed Eastern Pacific, canonical or classical events (Ashok et al. 2007). However, the 2004/05 El Niño event raised questions on the differences between events and brought rise to the identification and appreciation of a new type of El Niño event which is given multiple terms within the literature; the Warm Pool (Kug et al. 2009), Central Pacific (Kao and Yu 2009), Date-line El Niño (Larkin and Harrison 2005) or ‘El Niño Modoki’ (Ashok et al. 2007). All these terms essentially refer to El Niño events which are characterised by a zonal tripole pattern in SST anomalies, with the highest positive SST anomalies focused around the central Pacific around the international date line (Kug et al. 2009). At either side of this warm centre are cooler SST regions, thus creating the tripole pattern in both the ocean and, in turn, the atmosphere (Ashok et al. 2007). Central Pacific El Niño events can have unconventional teleconnections (Kug et al. 2009), and thus have a difference signature on weather patterns and related impacts.

It is important to improve understanding of the range of different El Niño events and their varying characteristics in order to better understand their varied teleconnections. This is especially true due to the increased frequency and persistence of El Niño Modoki events in recent decades (Ashok et al. 2007). This is addressed throughout this thesis. Improving understanding of ENSO events and their impacts helps with prediction of future impacts and identification of possible mitigation strategies which can aim to reduce the expected impacts (Khandekar et al. 2000). For example, the use of an ENSO forecast can provide planning and reduce uncertainty in key industries such as the Peruvian fisheries and rice production (Broad et al. 2002). The important impacts of ENSO over recent decades provide a key rationale for this research into historical ENSO.

1.4 The instrumental record of ENSO

The atmospheric, instrumental-based ENSO index, the Southern Oscillation Index (SOI), extends back to 1876, while the oceanic index, the SST Niño 3.4 index can be

extended back to 1856 (Ropelewski and Jones 1987; Kaplan et al. 1998). These indices provide the most reliable record of recent ENSO behaviour. The SOI can be extended further back into the mid nineteenth century with the use of historical pressure observations from Indonesia, discussed further in Chapter 2 (Können et al. 1998). However, the instrumental record remains too short to fully capture natural decadal-to centennial-scale ENSO variability and to investigate if recent ENSO characteristics are within the range of this natural variability. Fig. 1.5 shows the SOI, which is a measure of the pressure difference from Tahiti, in the central Pacific and Darwin, in Northern Australia (Bureau of Meteorology 2017; National Institute of Water and Atmospheric Research 2017). There is clear interannual variability in this index, with a strong negative SOI indicative of an El Niño event and a strong positive SOI indicating a La Niña event. The frequency of ENSO events is not consistent over time, with variations in periodicity seen within the instrumental record (Allan 2000). These vary from periodicities of 3-4 years in the period 1870-1910, 5-7 years in 1910-1920 and again in 1930-1960, and periodicities of 4-5 years from 1970-1990 (Allan 2000). The positive ENSO skew in amplitude is also visible in SOI time series with El Niño events typically of larger magnitude than La Niña events. This is seen recently, with three strong El Niño events (1982/83, 1997/98, 2015/15) dominating the recent ENSO record.

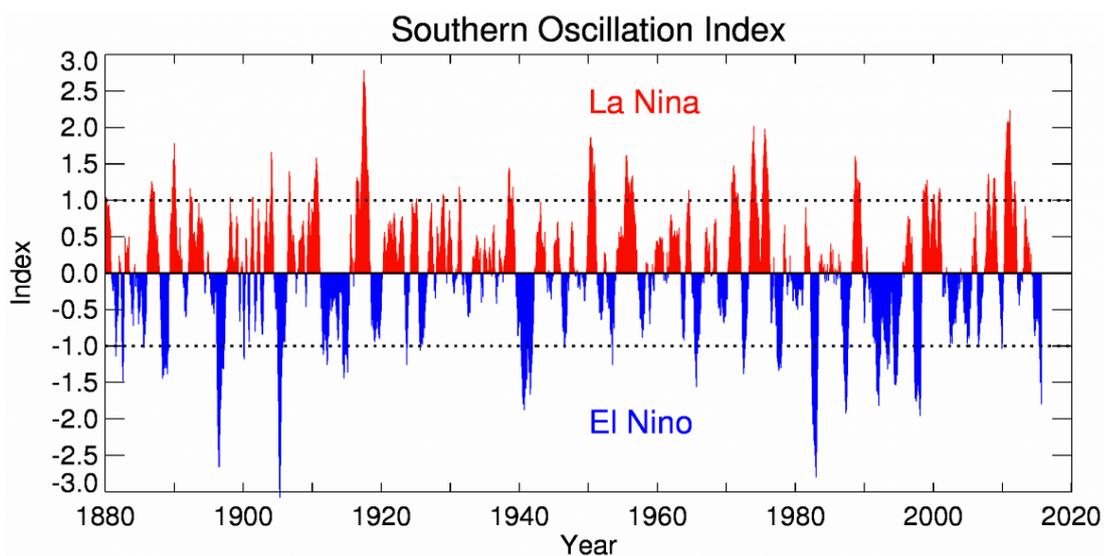


Fig. 1.5 Southern Oscillation Index (SOI) since 1880 (National Institute of Water and Atmospheric Research 2017).

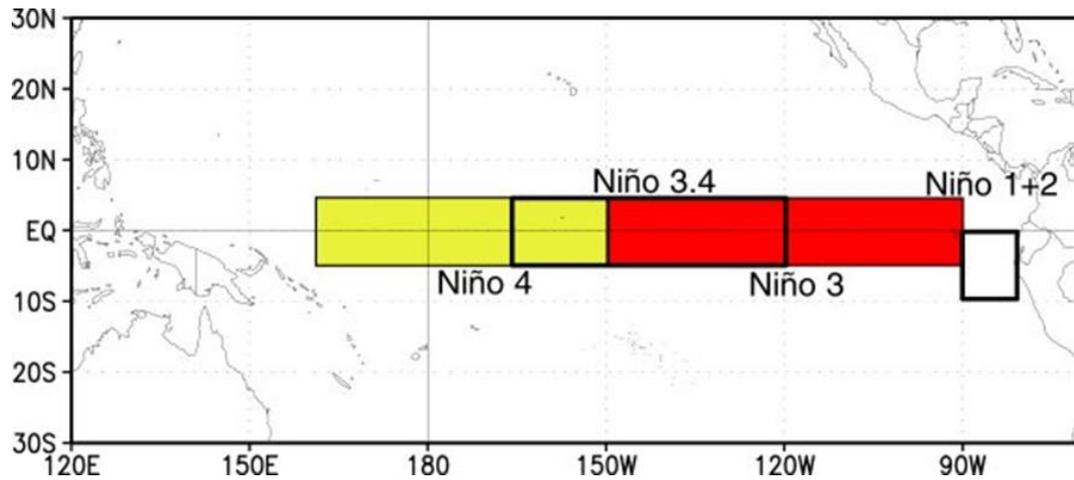


Fig. 1.6 The SST regions of the Niño indices. Niño 1+2 are the ‘classical’ El Niño regions. Niño 3.4 SST region commonly used in calibrating reconstructions of ENSO, while Niño 4 covers the central Pacific region. (NOAA 2017d).

The regions from which SST indices are extracted are shown in Fig. 1.6. The Niño 3.4 region is the most commonly used Sea Surface Temperature Anomaly (SSTA) index as it captures warming in both the central and eastern Pacific. The Niño 4 region captures central Pacific warming, while Niño 1+2 focus on warming in the far eastern Pacific. The SSTA indices, especially Niño 3.4, and the SOI are well correlated over their period of overlap, suggesting strong coupling between the ocean and the atmosphere. SST indices identify El Niño with positive anomalies and La Niña events as negative anomalies in the index, the opposite to that in the SOI. These instrumental records provide valuable information of the frequency and duration of ENSO events over the past 160 years. However, in order to extend our knowledge of ENSO further past these instrumental records a range of alternative sources of information must be used. This is the key focus of this thesis, to explore ENSO using a range of sources available in the historical period prior to the instrumental ENSO record. These sources of information will now be briefly introduced; however, more detail is provided within the relevant chapters which use the various data sources.

1.5 Ship logbook observations as a source of information for historical ENSO

Firstly, the observations from historical ships’ logbooks are used within this thesis to reconstruct ENSO during this pre-instrumental period. This is a unique data source

which has previously been used to reconstruct historical weather conditions and climatological phenomenon, but rarely used to investigate historical ENSO (Jones and Salmon 2005). The logbooks provide direct, daily observations on the state of the weather at sea which can be compiled into monthly and seasonal averages in order to reconstruct a time series of the large scale atmospheric circulation and pressure patterns (Wheeler et al. 2010). The wind observations, wind force and wind direction, are thought to be a more direct measure of pressure and circulation than other climatic variables such as temperature and precipitation making the wind observations from logbooks an important source of data regarding historical climatology (Jones and Salmon 2005). Therefore, the teleconnection between ENSO and surface winds are assessed within this thesis, driven by the use of observations from historical ships' logbooks, in which wind observations are available in abundance and recorded to a fairly common and consistent standard (Garcia-Herrera et al. 2005; Können and Koek 2005). Additional sources of historical observations of wind, air and sea temperature, and pressure are continuously being recovered, and following digitisation are poised to provide an additional valuable source of weather observations for this pre-instrumental period (Allan et al. 2011).

Large digitisation efforts over the last decade or so, such as the Climatological Database for the World's Oceans, CLIWOC (García-Herrera et al. 2005), have resulted in an expansion in the use of ships' logbooks as a source of historical climate data. Prior to this, logbook studies focused on very specific features for example synoptic conditions during key naval battles such as The Battle of Trafalgar in 1805 (Wheeler 2001), using journey duration as a proxy for the strength of trade winds in the South Atlantic (Farrington et al. 1998) and across the well-known shipping route of the 'Manila galleons' (Garcia et al. 2001), and use of the Hudson Bay company archives to reconstruct ice free seasons and the formation and break up of ice in the 18th and 19th century (Catchpole 1992). As a result of recent digitisation projects, reconstructions of large-scale pressure patterns and of regional precipitation (Gallego et al. 2005; Küttel et al. 2010; Hannaford et al. 2015) are now possible. The recently released International Comprehensive Ocean-Atmosphere Data Set (ICOADS) version 3.0 provides a recent and comprehensive collection of historical wind observations and has been used to create continuous historical records of climatic

features such as West African Monsoon (Gallego et al. 2015) and the Indian Monsoon (Ordóñez-Pérez 2016). Although this dataset is not fully utilised here due to the timing of its release, there is an overlap in the observations used within Chapter 2 and that from ICOADS v.3. Jones and Salmon (2005) presented the only previous attempt at ENSO reconstruction using the data from ships' logbooks and the ENSO reconstruction presented in Chapter 2 builds upon this previous research in a number of ways, fully exploiting this resource and developing the methodologies to do so.

The use of ships' logbook observations to reconstruct ENSO during the pre-instrumental period is chosen as it provides a record of direct 'semi-instrumental' weather observations from which a climatic reconstruction can be drawn (García-Herrera et al. 2005; Woodruff et al. 2005). This provides a bridge between indirect proxy observations and the advent of routine terrestrial-based instrumental ENSO records.

The observations from the ships' logbooks which are initially investigated within this thesis covers the period 1750 to 1854 AD, extending over 100 years before the range of the instrumental ENSO record. Daily weather observations were recorded with navigational and safety purposes in mind (García-Herrera et al. 2005). It is only afterwards that the contribution these records could have for climatological research was realized. The value of the observations for meteorology and climatology purposes was first commented on by Francis Beaufort who in 1809 questioned 'What better data could a patient meteorological philosopher desire?' in reference to the daily and sub-daily wind and weather data of Navy logbooks (Wheeler 2001). More recently this sentiment was expressed by Hubert Lamb who describes this data as 'a vast treasure trove waiting to be used' (Lamb 1995 pp. 89).

1.6 Data sources for historical- and paleo-ENSO

In addition to the observations from ships' logbooks, a number of data sources are used in this investigation of historical ENSO. The observations from ships' logbooks are the focus of Chapter 2, which explains its use in producing a new reconstruction of ENSO. Chapter 3 introduces documentary and proxy records commonly used to reconstruct ENSO. Documentary records have been used as a key indicator of ENSO events in the past with multiple El Niño chronologies existing using documents

primarily in South America (Quinn et al. 1987; Ortleib 2000; García-Herrera et al. 2008). They provide dated records of weather events in specific locations, however there is some uncertainty in the interpretation of the records and the reliability of the direct link to ENSO events (Ortleib 2000). A range of proxy records can provide information of paleo-ENSO including corals (Cobb et al. 2003), ice cores (Thompson et al. 1993), tree rings (Stahle et al. 1998) and sedimentary records (Wolff et al. 2011). Multiple existing multi-proxy ENSO reconstructions are compared within this thesis (e.g McGregor et al. 2010; Wilson et al. 2010; Emile-Geay et al. 2013a). Differences in ENSO chronologies and ENSO diversity from these various sources over the historical period are assessed in Chapter 3 and Chapter 4 respectively.

Reanalysis data are commonly used in climatic studies as they provide a useful global dataset of a number of variables which provide a coherent representation of global atmospheric circulation (Dee et al. 2011). The number of observations and the type of data assimilated into the reanalyses decreases back in time, therefore a modern reanalysis, ERA-Interim (Dee et al. 2011), is used in Chapter 2 for validation. Although it covers a shorter period, it covers the modern satellite era in which a few million observations can be assimilated to create the reanalysis (Poli et al. 2016). Chapter 5 focuses on the longer reanalyses, ERA-20C (Poli et al. 2016) and the twentieth century reanalysis, 20CRv2 (Compo et al. 2011). These span the entire twentieth century and therefore are used to assess their performance throughout a period with a variable number of observations, and to gain a longer term perspective on recent ENSO behaviour.

Chapter 6 introduces the final data source used in this investigation of historical ENSO; the output from climate model simulations. Climate models are a mathematical representation of reality which provide an understanding of the climate of the past, the present and future projected changes. Historical and paleo-climate simulations from models which made up the Coupled Model Intercomparison Project Phase 5 (CMIP5) are used here to investigate the representation of ENSO in current and past climates (Taylor et al. 2012; Schmidt et al. 2014). The use of this multitude of data sources allows for a thorough investigation of the behaviour of ENSO in the past, as analysis of both reconstructions and climate model simulations provides the most comprehensive summary of our understanding of the past (Jiang et al. 2015).

1.7 Aims of the thesis

The aim of historical- and paleo-climatology is to investigate past climates in order to improve understanding of natural climate variability and to inform future climate projections (Barry and Chorley 2010; Nash and Adamson 2014). The defining focus throughout this thesis is to explore ENSO during the pre-instrumental period in order to gain a better understanding of ENSO behaviour on longer time scales than available from the instrumental record. To achieve this, methods, concepts and data from historical- and paleo-climatology are used. There are four main aims which direct this research, which will now be discussed within the context of their key objectives, their rationale and the anticipated implications.

AIM 1: To reconstruct El Niño Southern Oscillation using the observations from historical ship logbooks (Chapter 2).

Objective 1A. The wind observations from ships' logbooks will be used in statistically-based reconstructions of ENSO during the historical period covering 1815-1854.

Objective 1B. Reconstruction methodologies will be tested in order to assess which method produces an ENSO reconstruction that corresponds best with the instrumental record during the modern period, using reanalysis data as an alternative to the logbook data. Methods of assessing the uncertainty of using a data source with temporally and spatially varying number of observations available will also be developed.

This builds on previous work by Jones and Salmon (2005), who carried out a preliminary reconstruction of ENSO using this data source and including this historical period, but who were limited by data availability and used a less developed methodology. This research will aim to raise awareness of the utilisation of logbook data for climatic research, building methodologies and highlighting regions in which future digitisation efforts could be focused in order for their contribution to historical ENSO studies to be even more valuable. Producing a reconstruction using the observations from ships' logbooks will build upon previous proxy-based research by presenting analysis based upon direct observations of the weather, bridging a gap

between proxy and modern instrumental records. A new record will therefore be produced, which is a key output from this research. This leads onto the next aim of this thesis. A large range of proxy and historical based ENSO reconstructions are available and provide a platform against which the logbook-based ENSO reconstruction can be evaluated.

AIM 2: To compare the logbook-based ENSO reconstructions to previous documentary and multi-proxy reconstructions of ENSO (Chapters 3 and 4).

Objective 2A. The logbook-based ENSO reconstruction methodologies will also be assessed over the modern period (1979-2013) to see how well ENSO variability events are captured compared to the instrumental record.

Objective 2B. Documentary and multi-proxy ENSO reconstructions will be compared to the instrumental SOI during a period of overlap (1877-1977) to assess how well they capture ENSO variability, and El Niño and La Niña events.

Objective 2C. The logbook-based reconstructions, documentary and multi-proxy reconstructions will be compared during the historical period, 1815-1854, to assess the agreement between them.

Objective 2D. In addition to the assessment of conventional El Niño and La Niña events, the representation of ENSO diversity within the various reconstructions will also be explored.

The main anticipated output of this aim is to improve the understanding of how well multiple ENSO reconstructions capture ENSO variability, events and ENSO diversity in order to inform the interpretation of the various records and assess their possible limitations in representing the full diversity of ENSO events. It provides an updated comparison of multi-proxy reconstructions building on previous review work (Gergis et al. 2006; Wilson et al. 2010) by using the most up-to-date reconstructions available, and with the inclusion of the new logbook-based reconstruction. In terms of ENSO diversity, this has only been explicitly addressed within proxy reconstructions a

handful of times (Schollaen et al. 2015; Karamperidou et al. 2015; Capotondi et al. 2015), and not within logbook-based reconstructions or within the documentary ENSO chronologies. It is extremely important to assess any bias within the reconstructions towards Eastern Pacific or Central Pacific events as this will be influencing the event chronologies taken from the reconstructions and will impact the understanding of ENSO diversity in the pre-instrumental period.

AIM 3: To investigate ENSO within reanalyses spanning the twentieth century and within historical climate model simulations (Chapters 5, 6 and 7).

Objective 3A. A number of reanalysis datasets will be assessed in terms of their climatology and representation of ENSO indices and spatial-teleconnection patterns, 1900-1999.

Objective 3B. The climatology and ENSO representation within historical climate model simulations (1900-2004) of five CMIP5/PMIP3 models will be compared to the reanalysis data and the instrumental record in order to rank the models based on selected ENSO metrics.

Objective 3C. The ENSO-wind teleconnection which is used within the logbook-based reconstruction will then be assessed over the twentieth century to investigate its stationarity.

ERA-20C is a newly released twentieth century reanalysis, and therefore here one of the first extensive analyses of its representation of ENSO is provided. Poli et al. (2016) presented a comparison of ENSO indices within the reanalysis, however here the spatial composites of key variables during ENSO events will be assessed and compared to the instrumental record of ENSO. Ranking methods will be used for climate model analysis. Historical simulations are often used to assess the ability of models to represent climatic features, but often do not reduce the number of models assessed over the LM. Here, the historical climate models are ranked and only the best performing models are then used to investigate the climate over the LM, and thus using the historical simulations for one of their key original purposes. The ENSO-wind

teleconnection is not commonly assessed within reanalysis or climate models as previous ENSO-teleconnection studies often focus on temperature or precipitation (Lewis and LeGrande 2015; Brown et al. 2016), which is more directly linked to commonly used proxy records. Therefore, focusing on wind as a key variable provides a unique aspect to this thesis.

AIM 4: To assess the representation of ENSO within CMIP5/PMIP3 climate model simulations spanning the Last Millennium (Chapter 8).

Objective 4A. The frequency of ENSO events of the Last Millennium, 850-1850 AD, will be assessed within climate model simulations.

Objective 4B. The stationarity of ENSO-wind teleconnections over the Last Millennium will also be assessed using the same climate model simulations.

Objective 4C. ENSO diversity over the Last Millennium will be investigated within the climate model simulations.

Using climate model simulations which span the Last Millennium provides a unique opportunity to assess the behaviour of ENSO over an extended time period. Although climate model simulations are not without their biases, they provide an important source of information in the pre-instrumental era, especially the simulation of spatial patterns associated with ENSO. The ENSO-wind teleconnection is again assessed over this period, providing for the first time an assessment of the stationarity of this teleconnection over the regions where wind was used as predictors for the logbook-based reconstructions. ENSO diversity is explored over the Last Millennium, again providing a longer-term perspective of the types of El Niño events. This builds upon previous research which uses LM climate simulations to assess the stability of ENSO teleconnections, namely temperature and precipitation, in different regions (Lewis and LeGrande 2015; Brown et al. 2016).

There is a wide range of previous research on ENSO as it is a vitally important component of the climate system. Wang et al. (2012) provide an extensive review of current ENSO understanding and the theories surrounding ENSO mechanisms,

expanding on the background information provided in Section 1.2 and 1.3, while Capotondi et al. (2015) provide a current assessment of the understanding of ENSO diversity. Gergis et al. (2006) and Wilson et al. (2010) both provide a comprehensive discussion of multi-proxy ENSO reconstruction methods, while Bellenger et al. (2014) discuss the representation of ENSO within climate models. This thesis builds on the work of this previous ENSO literature by assessing ENSO reconstruction methods and presenting new, complimentary methods. It also expands on previous research on ENSO diversity by assessing the ENSO diversity skill of a range of ENSO reconstructions and in climate model simulations at a level of detail not provided from previous research. Finally, it provides a focus on ENSO reconstruction by using winds as the key ENSO-teleconnection variable, this is alternative to the more commonly used hydroclimate variables of temperature and precipitation.

1.8 Thesis structure

The remainder of this thesis is organised and presented as follows:

Chapter 2 introduces improved methodologies for reconstructing ENSO using data from ship's logbooks. It presents the logbook-based ENSO reconstructions and investigates the methods used in detail. It also presents a new methodology to improve the understanding of using a data source with limited data availability and assesses the implications this has on the reconstructions produced. *Chapter 2 includes work from the published manuscript: Barrett HG, Jones JM, Bigg GR (2017) Reconstructing El Niño Southern Oscillation using data from ships' logbooks. Part I: methodology and evaluation. Clim Dyn DOI:10.1007/s00382-017-3644-7.*

Chapter 3 presents a review of previous ENSO reconstructions from documentary and multi-proxy databases. It compares a range of ENSO reconstructions to the logbook-based reconstructions presented in Chapter 2. *Chapter 4* investigates ENSO diversity within the various reconstructions and the implications of different ENSO flavours within the reconstructions. *Chapters 3 and 4 include work from the published manuscript: Barrett HG, Jones JM, Bigg GR (2017) Reconstructing El Niño Southern Oscillation using data from ships' logbooks. Part II: Comparisons with existing ENSO reconstructions and implications for reconstructing ENSO diversity. Clim Dyn DOI 10.1007/s00382-017-3797-4.*

Chapter 5 assesses the representation of ENSO within the historical reanalysis, comparing reanalyses spanning the twentieth century with the instrumental record in order to assess which twentieth century reanalysis best captures ENSO characteristics.

Chapter 6 uses the best reanalysis from Chapter 5 to compare to historical climate model simulations in over the twentieth century from selected CMIP5 climate models. Focus is given to the comparison of ENSO indices, Mean Sea Level Pressure and surface temperature climatology, and ENSO diversity.

Chapter 7 then assesses the zonal wind climatology and the ENSO-wind teleconnection within the reanalysis and historical climate models over the twentieth century. The stability of this teleconnection is investigated to assess changes in the strength of the relationship over time. *Chapter 8* then uses the models that best represent ENSO and ENSO-wind teleconnections in order to investigate ENSO variability, amplitude, frequency and teleconnections over the Last Millennium.

Chapter 9 summarises the main findings from the previous chapters linking together the findings of the logbook-based reconstruction, multi-proxy comparisons, and climate model simulations in order to provide the most holistic understanding of the ENSO climate of the past. The conclusion chapter also discusses the implications of the key findings of this thesis, assesses the limitations and provides suggestions for further research within this field.

2. Reconstructing El Niño Southern Oscillation using wind observations from ships' logbooks, 1815-1854

This chapter is based upon the publication 'Barrett HG, Jones JM, Bigg GR (2017) Reconstructing El Niño Southern Oscillation using data from ships' logbooks. Part I: methodology and evaluation. Climate Dynamics doi: 10.1007/s00382-017-3644-7' and includes text and figures from this publication.

2.1 Introduction

This chapter presents a number of reconstructions of El Niño Southern Oscillation (ENSO) using observations from ships' logbooks, 1815/16- 1853/54. Firstly, it introduces this historical data source, then explains and expands upon previous methodologies used for using reconstructing past climatic phenomenon. These methods are then applied to the historical observations in order to produce new reconstructions of ENSO in the early-to-mid nineteenth century. This addresses the first main aim of this thesis as outlined in Chapter 1, and provides logbook-based ENSO reconstructions which can then be compared to previous ENSO reconstructions from documentary and multi-proxy records in the following chapters.

The logbooks of ships which historically travelled the world's oceans contain a vast amount of meteorological information useful for studies of historical climate. These logbooks present a unique and exciting record of daily weather observations including wind force and direction, air temperature, pressure, and the state of the weather and the sea from the pre-instrumental era (Garcia-Herrera et al. 2005). Jones and Salmon (2005) were the first to attempt a reconstruction of El Niño Southern Oscillation using logbook observations from across the Indian Ocean (16°N to 40°S). However, gaps in the spatial and temporal coverage of the data were found to be a major restriction to their reconstruction (Jones and Salmon 2005). Koek and Können (2005) highlighted the challenge of reconstructing ENSO from this data source due to the lack of observations from the Pacific Ocean. This chapter builds on the initial attempt by Jones and Salmon (2005), by using an increased amount of data available since their work, and by focusing only on regions with strong ENSO-teleconnections. The main

output will be a reconstruction of ENSO during the period 1815-1854, which lies prior to the instrumental record of ENSO.

As introduced in Chapter 1, El Niño Southern Oscillation is the dominant mode of interannual climate variability in the tropics and is an important part of the natural climate system (Diaz and Markgraf 1992). The most reliable record of ENSO behaviour is from instrumental data, which are obtained from *in-situ* measurements of atmospheric and oceanic variables. The Southern Oscillation Index (SOI), an atmospheric index of the pressure difference between Tahiti and Darwin, extends back to 1876 (Ropelewski and Jones 1987). This was extended further using early pressure data from Jakarta and Tahiti to 1841, and from rain-day counts from Jakarta back to 1829 (Können et al. 1998). The record of Jakarta rain-day counts was used to create an SOI covering June-November, 1829-1850, and provides the most direct record to compare to the logbook-based ENSO reconstructions presented here. The use of instrumental records from Indonesia to construct the SOI provides an alternative method to the commonly used Tahiti-Darwin SOI, but still exploits the differences in pressure between the central Pacific and the regions to the west of the Pacific. The Niño 3.4 index, based on SST, has been extended back to 1877 (Bunge and Clarke 2009). Current understanding of ENSO behaviour during the early to mid-nineteenth century comes from a range of data sources. Proxy records, such as tree rings, corals and ice cores, are widely used as indicators of past climate and environmental change but have a range of uncertainties which are attached to them. They are less direct measures of the climate than the data from ships' logbooks. Recent multi-proxy studies have attempted to bring together a range of reconstructions, however a lack of consensus on ENSO behaviour on an interannual time scale remains (Emile-Geay et al. 2013a). Climate models provide an alternative source of information on past ENSO variability, but their ability to produce realistic ENSO simulations is still limited, and conflicts between models are apparent in multi-model comparison projects (Bellenger et al. 2014). In this chapter, the logbook-based reconstructions are compared to the early SOI records from the Jakarta observations, as both records are from direct, well-dated observations (Können et al. 1998). Comparison between the logbook and multi-proxy reconstructions will be discussed further in a Chapter 3 and Chapter 4.

Created as part of an international project completed in 2004, the Climatological Database for the World's Ocean (CLIWOC) provides access to a digital database of over 280,000 daily weather observations from mostly Dutch, British, Spanish and French national archives (Können and Koek 2005). The CLIWOC database spans 1662 to 1855, however, over 99% of the observations occur in the period 1750 to 1855 (Küttel et al. 2010). They provide a significant extension to modern instrumental records. However, the process of digitization is expensive and time-consuming therefore following the CLIWOC project, there were still over 90% of European, mostly British, logbook collections yet to be digitised (Garcia-Herrera et al. 2005). An additional 891 logbooks have since been digitised in the UK from English East India Company (EEIC) records, providing a further 273,000 observations from 1789 to 1834 (Brohan et al. 2012). This chapter considers the combined databases in order to utilise the highest number of observations possible. Despite these two large digitisation projects there is still a long way to go in order to fully exploit this resource (Wheeler 2014).

The presence of El Niño was first noted by a knowledgeable ship captain near the end of the nineteenth century (Wang and Fiedler 2006). It is fascinating that we can now use data from ships over 100 years before the discovery of this phenomenon, to reconstruct an El Niño index. The use of data from ships' logbooks was never intended for climatic study, it was vital for the navigation and safety of the ship. The contributions of both the ships providing the data, and the ships that first discovered El Niño, are much greater than could ever have been envisioned from the sailors at the time.

This chapter is organised as follows: *Section 2.2* describes the data sources used to carry out the historical ENSO reconstructions. Focus on spatial and temporal availability of the logbook data is provided. *Section 2.3* explains the data selection procedure, the reconstruction and validation methodologies, and introduces a method of assessing the impact of changing the number of observations per season on reconstruction skill. *Section 2.4* presents the results of the logbook data analysis with a focus on the wind force terms used. It also presents the results of an evaluation of the reconstruction methods and presents the historical ENSO reconstructions, along with comparison to the early Jakarta rain-day SOI record. The reconstructions and key

findings are then discussed in *Section 2.5*, while *Section 2.6* summarises the key findings of this chapter. Chapters 3 and 4 compare the results of this chapter to previous ENSO reconstructions, documentary and proxy-based, and discusses the implications of these results. This chapter addresses the first aim of this thesis and Objectives 1A and 1B, as outlined in Chapter 1.

2.2. Data

Wind direction and wind strength derived from ships' logbooks were used to reconstruct the Southern Oscillation Index. Due to a lack of overlap between the logbook data and the instrumental SOI, reanalysis data from the modern period (1979-2014) were used to establish statistical relationships between zonal wind and the SOI. These relationships were then applied to the logbook wind data to reconstruct the historical SOI.

2.2.1 Reanalysis data

Sub-daily 10m zonal wind values were obtained from ERA-Interim Reanalysis data (Dee et al. 2011). These data cover the period from 1979 to present, with 2013/2014 being the last DJF season used here. Reanalysis wind data were re-gridded to $7.5^{\circ} \times 8^{\circ}$ latitude-longitude, following previous studies using logbook data which found the use of an $8^{\circ} \times 8^{\circ}$ latitude-longitude grid provided a good resolution to enable suitable logbook data density, without compromising the climatology (Gallego et al. 2005; Jones and Salmon 2005; Küttel et al. 2010; Hannaford et al. 2015). The re-gridding of the reanalysis data was carried out using Climate Data Operators (CDO, <https://code.zmaw.de/projects/cdo>). ERA-Interim data are available at four synoptic time steps 00UTC, 06UTC, 12UTC and 18UTC. In order to best match the logbook data, which is taken at local noon, the daily value in each grid box was taken as the one closest to its local noon. These daily 'noon' values were then converted into monthly and seasonal values, based on the four standard meteorological seasons (DJF, MAM, JJA and SON). Only DJF values are used for the reconstructions as this is the season in which ENSO events typically reach their peak intensity (Bellenger et al. 2014), but also in which a suitable amount of logbook data availability was found.

2.2.2 Instrumental ENSO data

A monthly SOI was obtained from the Australian Bureau of Meteorology (BoM) extending from 1876 to present (Bureau of Meteorology 2015). The Troup SOI method was used, which calculates the standardised anomaly of the Mean Sea Level Pressure difference between observations from Tahiti and Darwin. The seasonal SOI was calculated from these monthly values. Other indices of ENSO are available and preliminary investigations included a range of indices, but the SOI was selected as it is an atmospheric measure of ENSO and is therefore closely linked to winds and circulation patterns, so had the strongest correlations with zonal winds (not shown).

2.2.3 Ships' logbook data

The focus of this reconstruction is on the wind data obtained from ships' logbooks, combining the data from the Climatological Database for the World's Oceans and digitised EEIC records. Wind direction was observed on-board ships using a 32-point compass and reported with respect to the direction from which wind was blowing (Wheeler 2005). For wind force, a range of descriptive terminology was used to make observations, as quantitative measures of wind speed had not yet been introduced. The Beaufort Force scale was not introduced into the Royal Navy until 1838, however prior to this there is evidence to suggest a fairly standardized vocabulary for recording wind force was already in place, with terms passed down through generations of sailors (Wheeler 2005). The various wind force terms found within the logbooks of the different countries were recorded in CLIWOC's 'Multilingual Meteorological Dictionary' (Garcia-Herrera et al. 2003). This included British, Dutch, Spanish and French terms, with translations carried out on a country by country basis. This reference guide was also used for transforming the wind force terms found within the additional EEIC logbooks (Brohan et al. 2012).

The digitization and processing of the raw data involves extensive efforts. Various corrections and adjustments have to be made to make to data suitable for use in scientific analysis. These include adjustments of dates, location, wind direction, wind force and notes of the general state of the weather. A wealth of metadata such as information about the ship and its notable encounters is also available in addition to climatologically important information. Following the corrections and adjustments,

which are carried out under the standard set by each digitization project (eg. CLIWOC), a process of data verification is carried out in order to assess the reliability of the data. The main method used in previous studies to test the reliability of data in records is to undertake consistency checks, comparing data from vessels in convoy (Können and Koek 2005).

Observations of wind direction were the most common of all the meteorological variables in logbooks from pre-instrumental era (Wheeler 2005). In accordance with modern day wind direction observations, winds were reported with respect to the direction from which they were blowing from and taken using a 32-point compass (Wheeler 2005). Adjustments to wind direction observations were carried out in order to take into account magnetic inclination, which describes the angle between magnetic north and true north. There can be up to 30° difference between the magnetic and true north with variations in time and location (Können and Koek 2005). Although sailors knew about this variation, there is a lack of historical evidence to suggest whether they accounted for this and adjusted their observations of wind direction to be with respect to true north rather than magnetic north (Können and Koek 2005). The CLIWOC project assumed this correction was not made, and thus made adjustments to all wind direction data to convert them with respect to true north (Können and Koek 2005). This was carried out with the use of pre-existing databases of magnetic variation (Wheeler 2014). Different digitisation projects use similar techniques, but it is important to be aware of the adjustments which have been made in order to understand the dataset which are being used for further climatic or meteorological analysis.

Daily observations of wind force were also extracted from the ships logbooks. Descriptive terminology was used to describe the force of the wind and recorded in the logbooks. However, it has been suggested that by 1750 nautical terms had developed into a fairly standardized vocabulary, with common terms being passed down through the ranks (Barriopedro et al. 2014). Studies have found that there is a larger degree of homogeneity in recording of wind force than initially anticipated (Oliver and Kington 1970). Therefore, many wind force terms had become fairly standardized before the formal adoption of the modern day Beaufort scale by the Royal Navy in 1838.

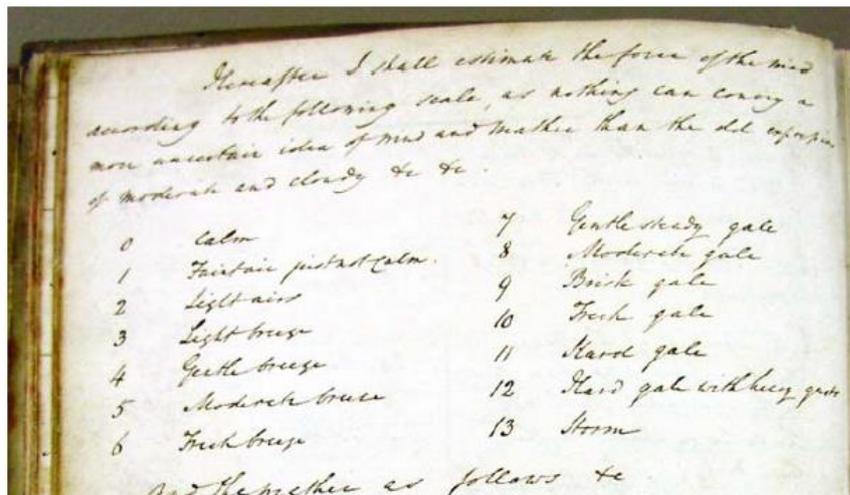


Fig. 2.1 Francis Beaufort's wind scale as originally proposed in the logbook of HMS *Woolwich* facing the page dated 13th Jan 1806 (Courtesy of UK Met Office, within Wheeler and Wilkinson 2005).

The scale devised by Beaufort written within a logbook over 20 years before it was officially adopted (Fig. 2.1), suggesting that these phrases were common even at the start of the 1800's (Wheeler and Wilkinson 2005). This helped in the production of the 'CLIWOC Multilingual Meteorological Dictionary' (Garcia-Herrera et al. 2003). Over 99% of the wind force terms found within the logbooks were able to be matched to a Beaufort scale equivalent (Garcia-Herrera et al. 2006). A final adjustment carried out concerned transforming from Beaufort terms to the modern day SI units of wind speed (ms^{-1}) in accordance with International Marine Meteorological Archive standards (Können and Koek 2005). The value of deck-estimated wind data is still appreciated today, with non-instrumental wind force data collected from modern day ships used by the UK Met Office to help prepare their synoptic forecasts (Wheeler 2005).

Other data adjustments were carried out in order to correct for differences in spatial and temporal differences associated with historical observations. Dates and location are needed to correctly position the meteorological data within the database. Before the shift to the use of modern day Gregorian calendar in 1753 the Julian calendar was in common usage (Barriopedro et al. 2014). Therefore, data from pre-1753 had to be adjusted to account for this. Another adjustment was needed to account for the

difference between the start of the nautical day and the civil day as the nautical day starts at noon (12pm midday), 12 hours in advance of the civil day (Wheeler et al. 2010).

It is important to be able to interpret and make corrections to the recorded latitude and longitude in order to best describe the spatial patterns present. It was not until 1884 that Greenwich was adopted as the universal meridian, therefore data pre-1884 use a range of zero meridians (Garcia-Herrera et al. 2006). It is often the case that the zero meridian used in voyages was the last point of land seen (eg. Lizard Point when leaving England), when reaching new land on route, this would then be used and the zero meridian replaced with this location. In compiling the CLIWOC database, over 600 locations were found to be used as zero meridians throughout the voyages (Garcia-Herrera et al. 2006). All of the ship routes were plotted and adjusted to use modern day zero meridian. Even though this resulted in largely realistic routes for most journeys, some routes were showing up to be going over land. This is shown in Fig. 2.2 which shows the initial outward route of HMS Surprise to be over land after the adjustment to modern day meridian (Können and Koek 2005). The main reason for this discrepancy is thought to be due to the accumulation of errors involved with the process of dead reckoning (Können and Koek 2005). Dead reckoning was often used in order to estimate the position of the ship. The previous days recorded location was used along with the estimated distance travelled and speed of the ship during that day. Due to the implausibility of the suggested route, data adjustment procedures were carried out by CLIWOC team who applied incremental adjustments to obtain a most likely route as shown in Fig. 2.2 (Können and Koek 2005).

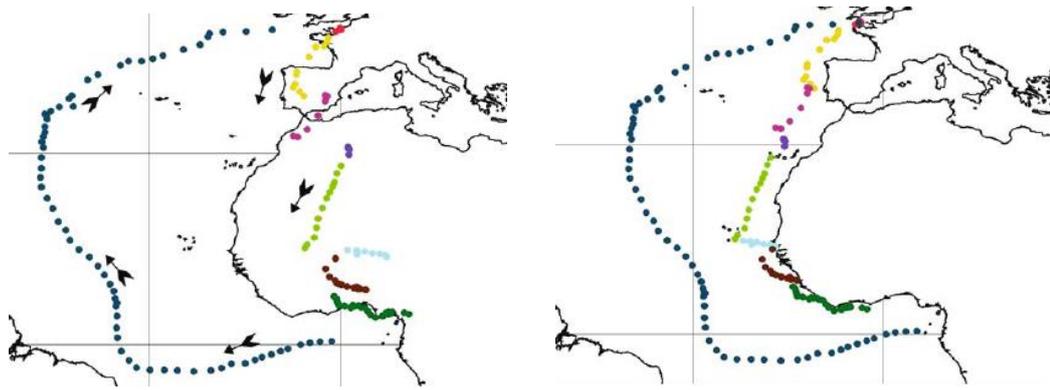


Fig. 2.2 The un-adjusted (left) and adjusted (right) route of HMS Surprise (1750-1751). The route is adjusted to take into account the modern day zero meridian and with incremental adjustments made to get route over the ocean (Können and Koek 2005).

When combined, the CLIWOC and EEIC digitised databases provide over half a million records globally over the period 1750 to 1854. However, 43% of these do not include observations of both wind force and direction. Both wind force and wind direction are required to calculate zonal wind, therefore only those logbook entries with both variables can be used here. The spatial and temporal distributions of these remaining 283,943 wind observations are shown in Figs. 2.3 and 2.4. The logbook observations are concentrated along the tracks of key historical shipping routes, which has resulted in their uneven global distribution (Fig. 2.3). The highest concentrations of data points are found in the Atlantic and Indian Oceans, and around the tip of South Africa. Very few observations are found over the Pacific Ocean as there were no regular shipping routes in this area during this period (Gallego et al. 2005).

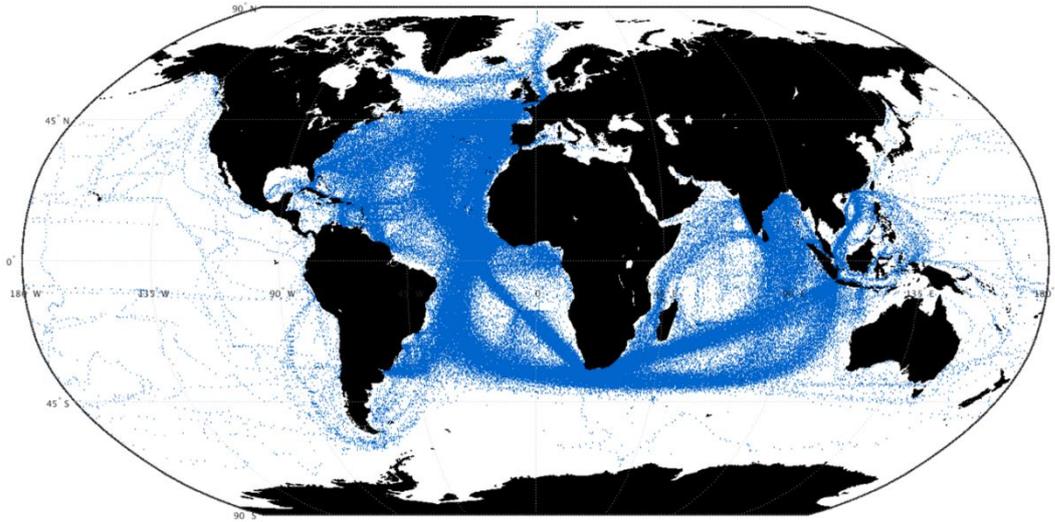


Fig. 2.3 Location of data entries from ships' logbooks with both wind force and wind direction observations from the digitised archives of CLIWOC and EEIC, 1750-1854.

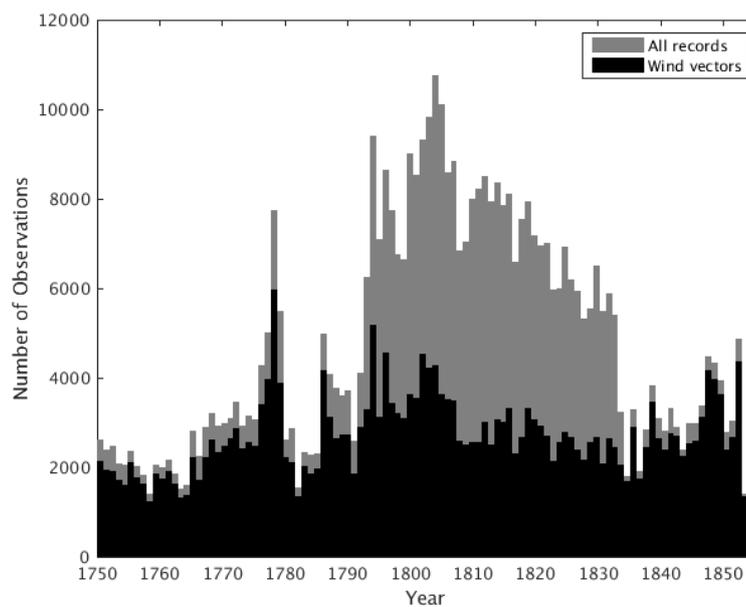


Fig. 2.4 Yearly distribution of all logbook data entries (grey) and those with wind vector observations (black), 1750-1854.

There is interannual variability in the number of wind observations globally from 1750 to 1854 (Fig. 2.4). Although the number of records increases in the period of EEIC data coverage, 1789 to 1834, the number of records with wind observations does not increase to a similar extent. This is a result of CLIWOC focusing on obtaining wind

observations whereas, digitisation of EEIC logbooks prioritised temperature and pressure measurements (Brohan et al. 2012).

Although the coverage of observations appears highly concentrated in Fig. 2.3, the data density on a given day is extremely sparse and is a key limitation to studies using this data source at present. To overcome this data sparsity, daily observations were averaged and aggregated into seasonal values, and seasonal reconstructions were carried out. Wind observations were averaged and aggregated into a 7.5° by 8° latitude-longitude grid, matching the resolution of the re-gridded ERA-Interim zonal winds. As ships often travelled in convoy, observations taken on the same day from within a given grid box were averaged into a daily value and counted as one observation for that day. Given that the logbook observations were taken at local noon, the effect of averaging observations from the same day is not influenced by the effects of the diurnal cycle.

2.3 Method

2.3.1 Selection of predictor grid boxes

Ideally, logbook data from the equatorial Pacific would be used as this is the centre of action for ENSO. However, the lack of logbook data in this core region results in dependence on regions with strong ENSO teleconnections. To identify these regions, Pearson's correlation coefficients were calculated between the re-gridded, seasonal 'noon' zonal winds from ERA-Interim and the seasonal Southern Oscillation Index, 1979/80 to 2013/14. Both datasets were detrended prior to calculation of correlations. Grid boxes whose correlations were significant at the 90% confidence level were considered as potential grid boxes suitable for use in the reconstructions. The correlation coefficients during the DJF season are shown in Fig. 2.5. Strong positive correlations are found during this season across much of the central and Eastern Indian Ocean, and the Maritime Continent, a region with strong ENSO teleconnections (Aldrian and Susanto 2003). During El Niño events high pressure anomalies are typical over the western Pacific and negative zonal wind anomalies are typical over the Maritime Continent (Wang and Fiedler 2006). The opposite is the case for La Niña events, with positive zonal wind anomalies typically in this region. These patterns lead to the large area of positive correlation between the SOI and zonal wind over the

Maritime Continent and into the Indian Ocean. Logbook data from this region of strong teleconnections is used for the reconstruction. A significant ENSO signal is also found within the zonal winds in other regions of the globe, which means that logbook data from more remote regions can also be exploited. In the Appendix, Fig. A1 shows the correlations between the DJF Niño 3.4 index and the DJF zonal wind over the same period and at the same spatial resolution as in Fig. 2.5 and finds similar but opposite correlations to that when using the SOI.

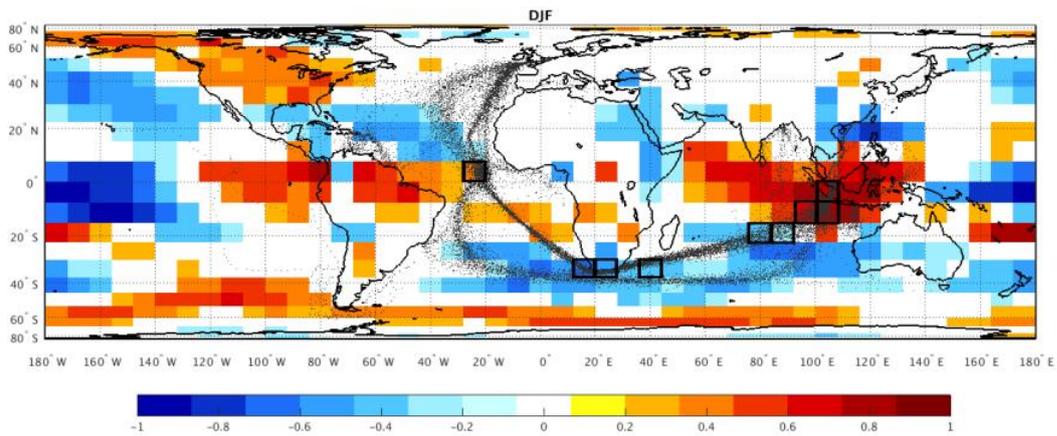


Fig. 2.5 Pearson's correlation coefficient between de-trended mean local noon DJF zonal wind and DJF Southern Oscillation Index, 1979/80-2013/14. Only those significant at the 90% confidence level are shown. Grid boxes used in the reconstructions are outlined and are numbered 1 to 9 from West to East. Black dots indicate location of DJF logbook wind observations 1815/16-1853/54.

After analysis of regions with significant correlations, the main factor limiting use of individual grid boxes for reconstruction is the availability of wind data from logbooks, also indicated in Fig. 2.5. Following extensive analysis of the availability of DJF observations per grid box per season, the nine grid boxes indicated in Fig. 2.5 were selected for use in the reconstruction. These span the eastern and central Indian Ocean (grid boxes 5-9), around the coast of South Africa (grid boxes 2-4) and one from the mid-Atlantic (grid box 1), these are labelled in Appendix Fig. A2 and the exact locations explained in Table A1. There are significant correlations between zonal winds and the SOI during DJF seasons in these grid boxes, with coefficients ranging from +0.86 to -0.60. The use of observations from different regions follows recent

ENSO reconstructions which have used proxy data from multiple teleconnection regions to provide a more robust ENSO signal (Braganza et al. 2009).

The logbook data availability from these nine grid boxes during DJF seasons 1815/16 to 1853/54 are shown in Fig. 2.6 and back to 1750 in Appendix Fig. A3. Between 1815 and 1854 there is an average of 13 observations per grid box per DJF season, however there is large interannual variability in the availability of data. In a given year, the amount of data across the nine grid boxes is also non-uniform. Prior to 1815 data availability in these grid boxes is reduced and missing data hinders reconstruction. There are only four DJF seasons between 1750 and 1815 where data are available in all nine grid boxes (1788, 1791, 1802 and 1811), therefore the years after 1815 were the focus of the ENSO reconstructions, with SOI reconstructions for these four additional years available in the Appendix (Table A2).

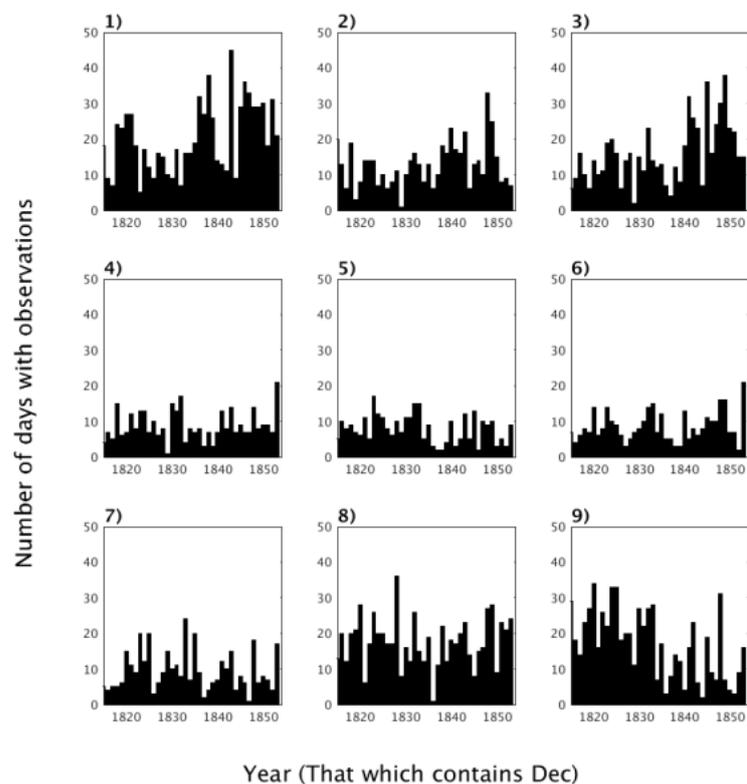


Fig. 2.6 Number of days with observations per DJF season in the nine predictor grid boxes, 1815/16-1853/54.

In order to provide as continuous a historical reconstruction as possible, DJF seasons in which individual grid boxes had less than three observations per season were infilled

via longitudinal interpolation. Following methods of previous studies (Küttel et al. 2010; Hannaford et al. 2015), where available, wind observations from adjacent grid boxes which exceeded the selected number threshold were used. Interpolation was carried out longitudinally as the boxes to the east/west had better data availability than those to the north/south due to the historic shipping routes in the regions used. For example, for a given grid box if there were less than the minimum threshold of observations for the DJF season in a given year, then the DJF value from that year in the adjacent grid box which exceeds the threshold would be used, and the difference between the ERA-Interim reanalysis climatological means of the two grid boxes in the reanalysis applied, to gain a value for the grid box with low observations. If not enough observations were available, then the year was deemed missing.

2.3.2 Reconstruction and validation methods

Two common approaches to past climate reconstruction are Composite-Plus-Scale (CPS) and Climate Field Reconstruction (CFR) (Mann et al. 2005). Principal Component Regression (PCR) is a well-established method of CFR and is used in previous studies containing logbook data (Jones and Salmon 2005; Küttel et al. 2010; Hannaford et al. 2015). CPS and CFR are complementary methods, with CFR providing reconstruction of spatial patterns based on assumptions of stationarity between the predictor and large scale patterns, and CPS involving a simpler procedure based on local statistical relationships (Mann et al. 2008). Recent ENSO reconstructions have used both Composite-Plus-Scale and Climate Field Reconstruction methods (Wilson et al. 2010, Emile-Geay et al. 2013a). Here, both the PCR and the CPS reconstruction methodologies are used.

Principal Component Regression reduces the dimensionality of climate data and helps to separate climate signals from noise. The PCR method used created a covariance matrix from the zonal wind data, which was then decomposed into eigenvalues and eigenvectors (Taylor et al. 2013). No areal weighting scheme was applied. The PCs are then correlated with the detrended SOI and those with the highest correlation coefficients (PC 1, 2 and 6) retained for reconstruction (Jones and Widmann 2003).

Fitting regression was carried out using the selected PCs and the resultant reconstruction was compared to the instrumental SOI to assess the performance of the

model. Fitting and validation statistics were used in order to inform selection of the optimum number of PCs to retain, until the highest skill scores were obtained. The relationships established during the fitting period were then applied to the historic period, with zonal winds as predictors and SOI as predictand.

Various methods of CPS have been used in past work; here the approach outlined by the PAGES 2K Consortium has been employed (PAGES 2K Consortium 2013). This method essentially used the mean DJF zonal winds of each grid box as a ‘proxy’ time series for the SOI. It used a weighting based on the correlation between detrended zonal winds and detrended SOI (Neukom et al. 2013; PAGES 2K Consortium 2013). Pearson’s moment correlation coefficients between these variables were calculated for each grid box over the calibration period, 1979-2013. These weightings were then applied to the logbook zonal wind anomaly time series from each grid box. An average was then calculated from these nine time series and the resulting time series normalised to mean zero and unit variance. Therefore, the CPS provides a reconstruction of normalised SOI, whereas PCR method has a larger range of variability, more similar to the instrumental SOI.

For both methods, seasonal zonal wind anomalies were calculated using detrended, re-gridded ERA-Interim zonal wind, over the period 1979/80-2013/14. ERA-Interim zonal winds and DJF Southern Oscillation Index are detrended prior to model fitting in order to avoid an influence from modern trends in the reconstruction method. However, after fitting and validation, the non-detrended data are used to produce the reconstruction. Uncertainties were calculated using ± 1.96 standard deviations of the residuals from the fitting as the 95% confidence intervals (PAGES 2K Consortium 2013).

2.3.3 Assessing reconstruction skill

A number of statistics were used to assess the quality of the reconstruction models following the example of previous ENSO-reconstructions studies (Emile-Geay et al. 2013a). Pearson’s correlation coefficients and coefficient of determination were calculated between the fitting reconstruction and the instrumental SOI. Cross validation was also carried out using the same period as the model fitting (1979-2013) to test the regression relationships (Wilks 2005). This method has been used in

previous PCR-based reconstructions where the period over which regressions are derived is too short for other validation methods (Jones and Widmann 2003; Hannaford et al. 2015). Reduction of Error (RE) skill scores were also calculated as an objective measure to assessing the performance of the statistical reconstruction models (Cook et al. 1994). RE scores range from $-\infty$ to +1, where 1 suggests complete agreement with the predictand, 0 indicates that the reconstruction is of same value as the climatology, and negative scores suggest no meaningful information can be gained from the reconstruction.

2.3.4 Threshold tests

Due to the limited availability of logbook data, seasonal averages must be calculated using only a restricted number of days of observations. Previous logbook studies have used different thresholds for defining the minimum number of observations that should be used in the calculation of seasonal averages. Küttel et al. (2010) and Neukom et al. (2013) only used seasons with observations for at least three days in each $8^\circ \times 8^\circ$ grid box, whereas Hannaford et al. (2015) used a threshold of at least one-tenth of the number of days in the season. The key aim is to find a balance between high spatial resolution and a good signal-to-noise ratio from the observations (Küttel et al. 2010).

A minimum observation threshold test was carried out to assess the impact of using only a few observations per season from modern reanalysis data to calculate the seasonal means. These thresholds ranged from one day of observations per season per grid box up to 10 days of observations. This range was chosen as it covers both the threshold used by Küttel et al. (2010) and Neukom et al. (2013), and that used by Hannaford et al. (2015). At each threshold, the PCR and CPS methods were carried out 100 times over the fitting period using seasonal datasets composed of randomly selected daily mean values from ERA-Interim reanalysis data. The eigenvectors for PCR, and the correlation weights for CPS, were those calculated with the full seasonal dataset, as these are what are used in the historical reconstructions. The seasonal mean was calculated using the dataset obtained from the selected threshold number of observations in each grid box. At each threshold (1 to 10), the mean correlation

coefficients between the reconstructed SOI and the instrumental SOI, and the RE skill scores, were calculated.

The minimum data thresholds tests are based on a ‘worst-case scenario’, where each grid box contains only the selected minimum threshold number of observations. In reality the number of logbook observations per grid box is non-uniform (Fig. 2.6), with thresholds being exceeded in most of the grid boxes. Therefore, an additional test was carried out to directly replicate the data availability in each of the years of the logbook-based reconstruction. For each year of the logbook-based reconstruction, fitting and validation was carried out using the data availability for that given year. For example, the data availability in 1815/16 DJF season in the predictor grid boxes was replicated for all seasons throughout the fitting period (1979-2013), and the RE score calculated. Therefore, each year of the historical reconstruction has its own RE score assigned to it based on a fitting which replicates its data availability throughout a 35-year period. The RE scores obtained from this analysis were then used to assign a reconstruction skill to each of the years of the reconstruction, flagging up the years that need to be treated with more caution than others.

The initial threshold tests, which investigated the skill of the reconstruction methods using observations ranging from one per grid box per season up to 10 observations found that CPS performs better than PCR when a limited number of observations are available. Table 2.1 compares the correlation between the instrumental SOI and the resulting fitting reconstructions as well as presenting the RE for both PCR and CPS methods. For both the PCR and CPS reconstructions, the correlation coefficient (r) and the RE skill score increased with the number of observations used to calculate the seasonal mean, but are lower than those obtained when all days within a season are used (ie. full reconstructions). For CPS, all but the lowest threshold have mean RE skill scores which are positive, indicating a reconstruction skill higher than climatology when more than one observation per grid box is used. The PCR method performs with less skill when using such few observations, with negative skill scores obtained when six or less observations are used per grid box, suggesting additional caution is needed when using the PCR method with few observations.

Table 2.1 Pearson’s correlation coefficient (r) and RE Skill score from fitting reconstructions (1979/80-2013/14) carried out using 1 to 10 days of observations per grid box and full season of observations, from both PCR and CPS methods. The r and RE values shown are the mean from 100 tests at each threshold.

Number of days of observations per DJF season	Pearson’s correlation coefficient between reconstructed and instrumental SOI (mean of 100)		RE Skill Score (mean of 100)	
	PCR	CPS	PCR	CPS
1	0.30	0.48	-4.16	-0.04
2	0.43	0.59	-1.66	0.17
3	0.47	0.66	-0.87	0.31
4	0.53	0.69	-0.46	0.38
5	0.59	0.71	-0.18	0.41
6	0.59	0.73	-0.07	0.46
7	0.64	0.75	0.11	0.49
8	0.66	0.76	0.20	0.51
9	0.68	0.78	0.27	0.56
10	0.69	0.78	0.30	0.56
Full	0.89	0.87	0.80	0.75

However, the minimum threshold results are based on ‘worst case scenarios’. In reality, the number of observations in each grid box is not the same in a given year and varies throughout the time period. Therefore, when replicating the actual data availability from the logbook years a more directly applicable RE scores for the historical reconstructions is obtained. The RE value for the individual year relates to the skill of a reconstruction that could be carried out assuming the data availability during that given year is replicated over the entire reconstruction period, and the results are shown in Fig. 2.7.

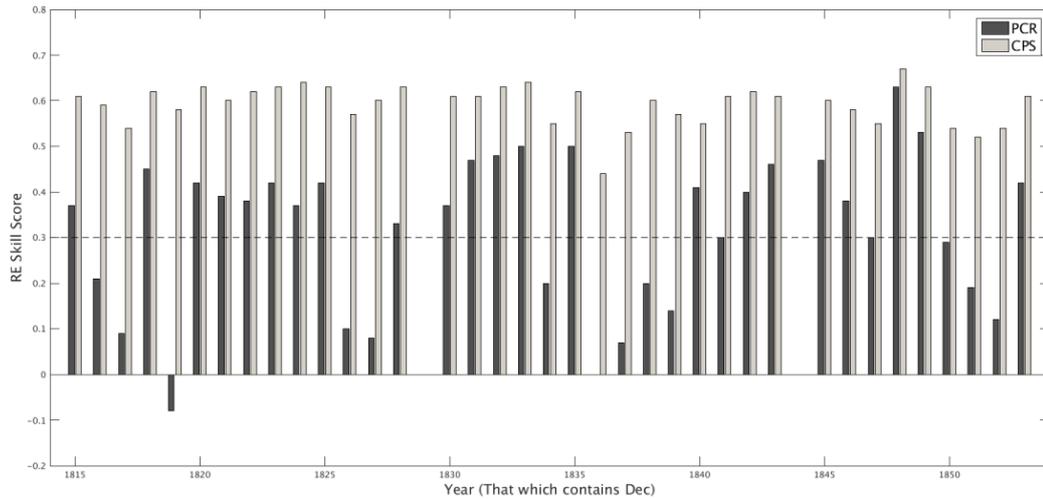


Fig. 2.7 RE skill scores obtained from PCR (dark grey) and CPS (light grey) reconstructions carried out over the fitting period, replicating the logbook data availability from each of the DJF seasons, 1815-1854. 1829/30 and 1844/45 are missing data years. Dashed line indicated RE=0.30.

When fully replicating the data availability of the period 1815/16-1853/54 during the fitting period the mean RE score obtained is 0.32 for the PCR method and 0.60 for CPS. The PCR RE scores are lower than the CPS RE scores for the same data availability, but all are positive except for in 1819/1820. Any positive RE score is considered successful, making the difference between success and failure very small in some cases (Cook et al. 1994). Positive values closer to zero have less skill than those closer to one, however all are classed as ‘skilful’. In order to account for the reduced skill of positive values in close proximity to zero, a threshold was selected, below which additional caution in the reconstruction should be taken. From the results in Fig. 2.7, an RE score of less than 0.30 was selected as the value at which a year within the historical reconstruction which should be treated with caution. All of the CPS scores are higher than this threshold, whereas 13 of the PCR scores are lower than 0.3 and therefore are highlighted as years with additional uncertainty in the historical PCR reconstructions.

2.3.5 Data origin and wind force terms

Prior to carrying out the reconstructions, the wind observations from the logbooks were analysed. Previous analysis of the wind force conversion terms highlighted that

translations were carried out on a country by country basis, and that improvements could still be made on the calibration of wind force terms between countries (Koek and Können 2005). As a result of this, the country of origin of the observations used here has been analysed. The temporal coverage of observations from different countries varies throughout the CLIWOC period. Dutch ships dominate the later portion of the CLIWOC period and contributions from UK logbooks to CLIWOC end in 1829. The additional digitised EEIC observations only span from 1787 to 1834. The observations from the grid boxes used in the reconstructions have been divided into the countries from which the ships originate. Fig. 2.8 shows the Dutch and EEIC contributions, which together make up 96% of the observations used during the period 1815-1854. An additional 3% of observations are taken from UK CLIWOC ships and <1% from Spanish ships, which are not included in this section of analysis. The period 1815-1833 is dominated by EEIC observations (68%), whereas after 1833 all of the observations are from Dutch ships. This distinct shift in origins of the data source could be important when interpreting the resulting seasonal zonal winds and the reconstructed SOI.

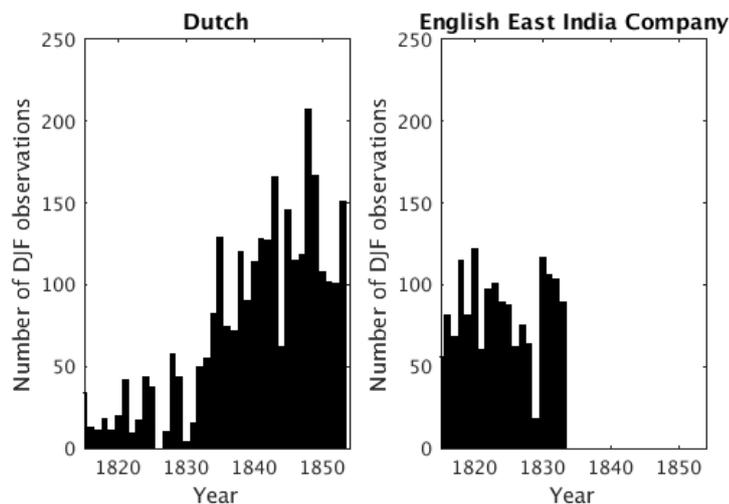


Fig. 2.8 The number of DJF wind observations from the regions used in the SOI reconstruction which come from Dutch and English East India Company during DJF seasons 1815/16 -1853/54.

To examine the differences in wind force terms used per country, the frequency of Beaufort Force equivalent wind terms used by the Dutch and EEIC ships during the period of data overlap, 1815-1833, is assessed in Fig. 2.9. The data from the nine grid

boxes were grouped together, based on similar climatology, into three sub-regions. These sub-regions consist of tropical grid boxes (1, 7, 8 and 9), those from the trade wind region (5 and 6) and grid boxes from around the southern tip of South Africa (2, 3 and 4). The Beaufort Scale ranges from 0 to 12, with higher numbers relating to stronger winds. Looking at both the Dutch and EEIC observations, grid boxes located in the tropics have a higher frequency of low Beaufort Force values, and some of the most extreme values are found in the grid boxes around southern Africa.

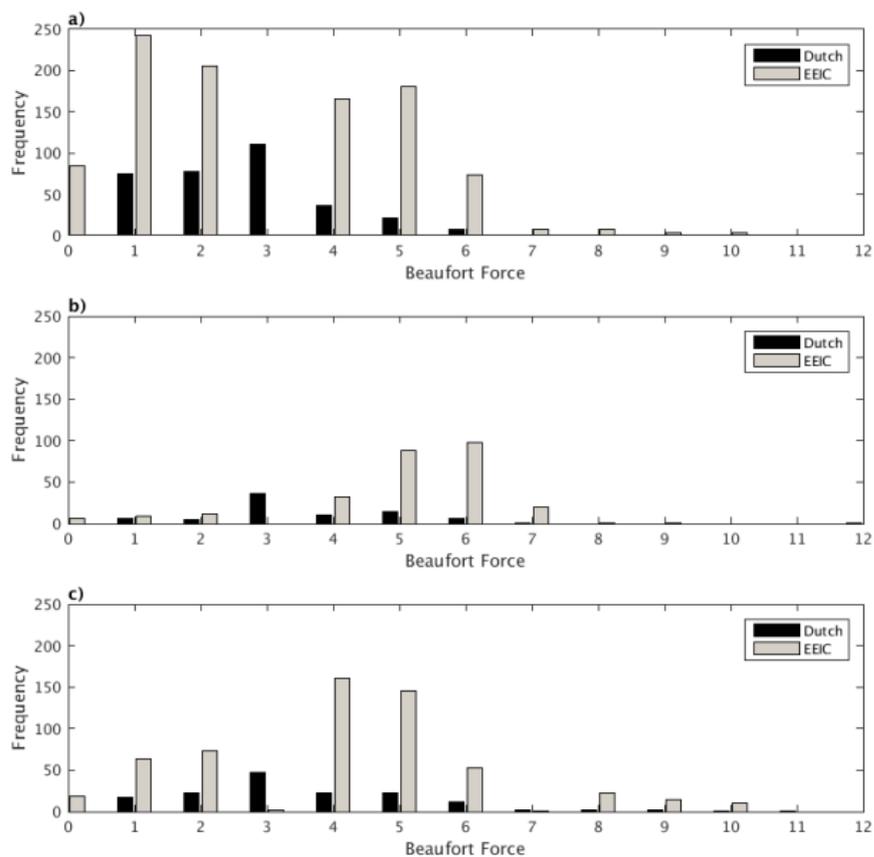


Fig. 2.9 Frequency of Beaufort-equivalent wind force terms used by the Dutch and EEIC in the a) Tropics (Grid boxes 1 and 7-9) b) Trade wind region (5-6) and c) southern African grid boxes (2-4).

Although the Dutch observations come from a smaller sample size, the general distributions from the two countries can be compared. A main feature of the distributions is the almost complete lack of Beaufort Force 3 recorded in the EEIC data. It was used less than 0.01% in the EEIC observations, whereas it is the most common Beaufort Force found in the Dutch data, being used in 32% of the records.

The EEIC observations are generally of higher wind forces than the Dutch data. A two sampled Kolmogorov-Smirnov test (Wilks 2011) was used in each of the three sub-regions to test the null hypothesis that the wind force data from the Dutch and EEIC ships come from the same distribution. In all three regions, the null hypothesis can be rejected at 99% confidence level, indicating a statistically significant difference in the distributions of the wind force terms used by the Dutch and EEIC from within the same regions.

These preliminary investigations highlighted the difference in wind force terms used by the different countries, and the dominance of Dutch data in the later part of the record. As a result, this was taken into account during the ENSO reconstructions. Firstly, reconstructions were carried out using a climatological mean calculated over the entire period, 1815-1854 (known as variant 'A' later). Following this an alternative method was employed which used two separate means, one for the period dominated by British records, 1815-1833, and one for the later period, dominated by Dutch records, 1834-1854 (variant 'B'). This second method aims to address the impact of using data from countries with differences in the terminology for wind force observations.

2.4 Results

2.4.1 Reconstruction skill

The reconstruction skill of both the PCR and CPS methods were assessed over the fitting period, 1979-2013, using the re-gridded ERA-Interim data and instrumental Southern Oscillation Index for the DJF season. Fig. 2.10 shows the SOI reconstructions along with the instrumental SOI for this period. Due to differences in the methodologies, the CPS method produces a normalised reconstruction, whereas the PCR method produces a reconstruction with variability similar to the non-normalised instrumental record.

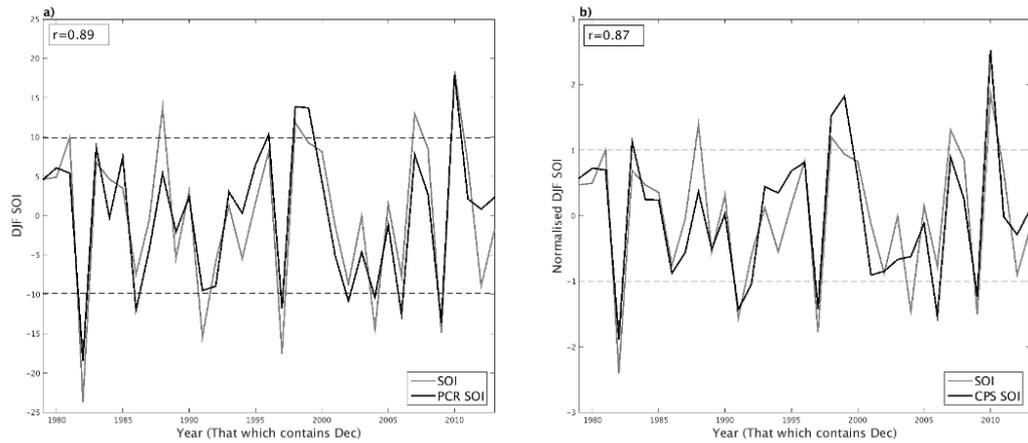


Fig. 2.10 Instrumental DJF Southern Oscillation Index and reconstructed DJF SOI using (a) Principal Component Regression and (b) Composite-Plus-Scale (normalised values), 1979/80-2013/14. Dashed lines represent ± 1 standard deviation, the threshold used to define the main ENSO events.

For the PCR method, Principal Components 1, 2 and 6 were retained due to their high correlation with the instrumental SOI. The coefficient of multiple determination during the fitting period is 0.89 (significant at the 99% confidence level) (Fig. 2.10), thus 80% of the variance of the SOI is explained by the PCR model. The correlation between the PCR-reconstructed SOI and the detrended instrumental SOI from cross validation is 0.86, with an RE score of 0.73. For the CPS reconstruction, a correlation of 0.87 (significant at the 99% confidence level) was found during the fitting period, thus 75% of the variance is explained by the CPS method (Fig. 2.10). For both methods, when comparing the non-detrended SOI with the reconstructed SOI, high RE values are obtained, with a RE score of 0.80 for the PCR method and 0.75 for the CPS method. Overall, the PCR method has slightly higher reconstruction skill than the CPS, however both perform well and are therefore both used to carry out historical reconstructions.

2.4.2 Modern wind patterns during strong El Niño and La Niña events

To test these modern reconstructions further and to see if the selection of predictor grid boxes is physically sensible, the wind patterns from some of the strongest ENSO events in this record were analysed. The DJF zonal wind anomalies in two recent strong El Niño events, 1982/83 and 1997/98, and two strong La Niña events, 1998/99 and 2010/11 are shown in Figs. 2.11 and 2.12, respectively. Although the 1982/83 and

1997/98 El Niño events were both strong events, the difference in the strength of the wind anomalies in predictor grid boxes is clear from Fig. 2.11. The negative anomalies in the Indian Ocean are stronger in 1997/98, whereas there is a much larger region of positive anomalies around the tip of Southern Africa in the 1982/83 event. Therefore, using grid boxes from both these key regions in the historical reconstructions increased the chance of identifying key anomalous years in the past. The same is true for La Niña events (Fig. 2.12), with much stronger positive anomalies in the eastern Indian Ocean in the 2010/11 event, compared to the 1998/99 event.

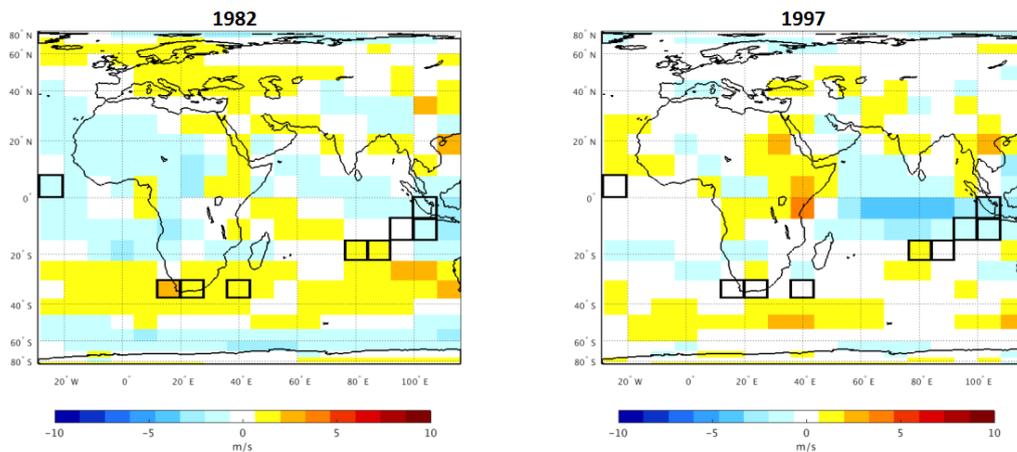


Fig. 2.11 DJF zonal wind anomalies (ms^{-1}) during El Niño events in 1982/83 (left) and 1997/98 (right), relative to the 1979-2013 climatology. Predictor grid boxes used in the reconstruction are outlined in black.

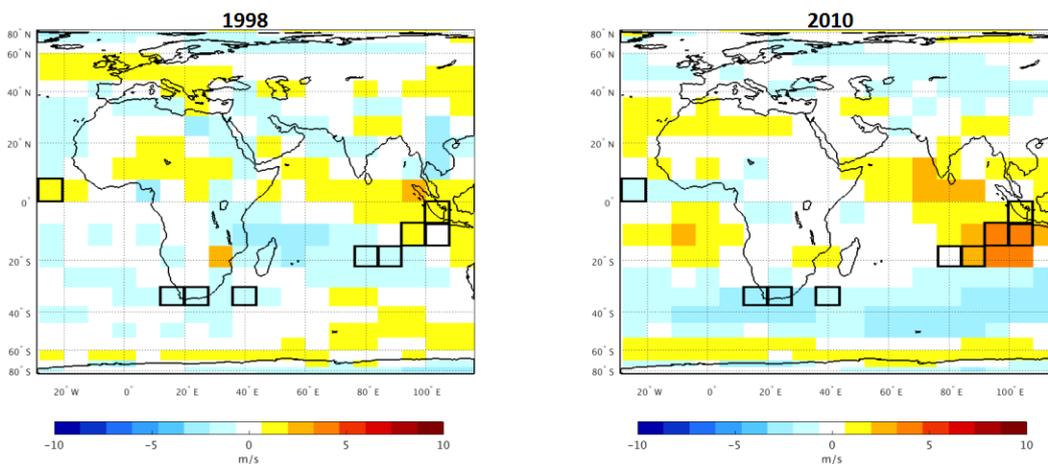


Fig. 2.12 DJF zonal wind anomalies (ms^{-1}) during La Niña events in 1998/99 (left) and 2010/11 (right), relative to the 1979-2013 climatology. Predictor grid boxes used in the reconstruction are outlined in black.

Even though the expected anomalies are not found in all of the grid boxes used in the four events, the reconstructed SOI for these years show values very close to the true SOI (Fig. 2.10). In most cases, the sign of the wind anomalies in the predictor grid box fits with what is expected from the correlations indicated in Fig. 2.5. Therefore, the use of anomalies from different regions is beneficial for the reconstruction approach and helps provide a more robust ENSO signal.

2.4.3 Historical ENSO reconstructions

Following, fitting and validation of both reconstruction methods, historical reconstructions were carried out, using the model fitted over the whole period. Fig. 2.13 shows the DJF SOI reconstructions using PCR and CPS over 1815-1854. Variant A reconstructions, (PCR A and CPS A) were calculated using a climatological mean from the entire period, 1815-1853, whereas, variant B (PCR B and CPS B) were calculated using two climatological values, one covering 1815 to 1833 and one for the later period 1834 to 1854. This takes into account the change in data source from mostly EEIC in the earlier period to all Dutch records after 1833. Years with RE scores less than 0.3 are indicated with the historical PCR reconstructions (light grey bars), taking into account the results of the minimum data availability threshold testing.

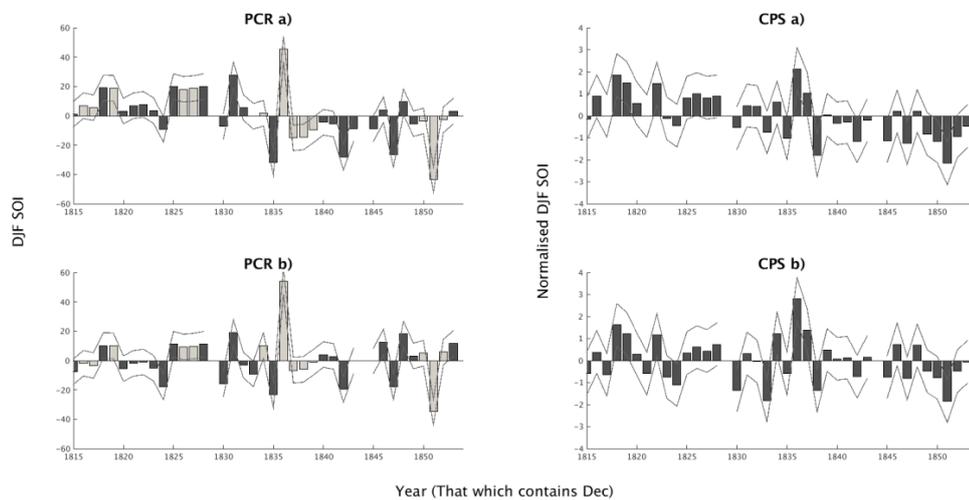


Fig. 2.13 Reconstructions of DJF SOI 1815/16-1853/54 using Principal Component Regression (PCR) and Composite-Plus-Scale (CPS) (normalised), with (a) the same mean throughout and (b) using 1815-1833 and 1834-1853 means. Thin lines show the 95% confidence intervals. Light grey bars indicate years with RE scores less than 0.3. No data for 1829/30 and 1844/45.

Table 2.2 Pearson’s correlation coefficients between historical reconstructions of DJF SOI 1815/16-1853/54 using Principal Component Regression (PCR) and Composite-Plus-Scale (CPS) (normalised), with (a) the same mean throughout and (b) using 1815-1833 and 1834-1853 means. All significant at 99% confidence level.

	PCR A	PCR B	CPS A	CPS B
PCR A	----			
PCR B	0.86	----		
CPS A	0.81	0.65	----	
CPS B	0.67	0.75	0.88	----

Table 2.2 shows the correlation coefficients between the four historical reconstructions. They are all significantly correlated and show a high degree of similarity to each other. The strongest correlations are found between the two CPS reconstructions, $r=0.88$, followed by the two PCR reconstructions, $r=0.86$. However, correlations between the reconstructions using the different methods remain high, suggesting good agreement between the two reconstruction methodologies. Within the PCR reconstructions, some of the more extreme events are those with additional uncertainty (i.e. $RE < 0.3$). However, these events are still classified as main ENSO events in the CPS reconstructions, where reconstruction skill is higher. Therefore, similar signals are being provided by the two methods.

In order to assess the differences between the reconstructions, additional statistical analysis was performed. Least squares fitting was carried out to identify any statistically significant trends in the four reconstructions. A negative trend in SOI over the reconstruction period is suggested from Fig. 2.13, with the SOI generally decreasing during the period 1815/16 to 1853/54. Results from linear regression for the four reconstructions found that a statistically significant ($p < 0.05$) negative trend is found when only one climatology is used throughout the entire period, for both the PCR and CPS methods (Table A3). When variant B is used, to take into account the change in dominance from EEIC to Dutch observations, this significant linear trend is no longer present in either of the statistical reconstructions.

Analysis of the mean and two-sampled t-tests was carried out to assess the difference in SOI before and after 1833/34 from each of the four reconstructions (Table A4). It was found that the reconstructions without the adjustments in the mean (Variant A) have significantly different SOI in the earlier period compared to the later period. In contrast, for those reconstructions where adjustments to the mean were made (Variant B), this statistically significant difference was not found. Therefore, the impact of using logbook data from two periods dominated by countries which use different wind force terminology is significant and should be addressed.

To identify the main ENSO events from these reconstructions, El Niño events were classified as seasons with a negative SOI value more than one standard deviation from the mean, and La Niña as a positive SOI more than one standard deviation from the mean (Stahle et al. 1998). Fig. 2.14 shows DJF seasons classified as El Niño and La Niña in the four reconstructions. Two events are found in all four of the reconstructions and have the most extreme SOI values, an El Niño event in 1851/52, and a La Niña event in 1836/37. These are named as high confidence events. Attention is then given to those ENSO events found in at least three of the four logbook-based reconstructions, which are seen as medium confidence events. Between 1815/16 and 1853/54, five additional events fall into this category; two La Niña events 1818/19 and 1819/20, and three El Niño events 1835/36, 1842/43 and 1847/48. Taking into account the significantly more positive SOI values found in the reconstructions of variant A prior to 1834, the criteria for medium confidence events is relaxed to only the two reconstructions using variant B needed to classify El Niño events prior to 1834. This results in two additional El Niño events, 1824/25 and 1830/31. Grouping together these high and medium confidence events, there are six main El Niño events and three main La Niña events during the logbook-based reconstruction period.

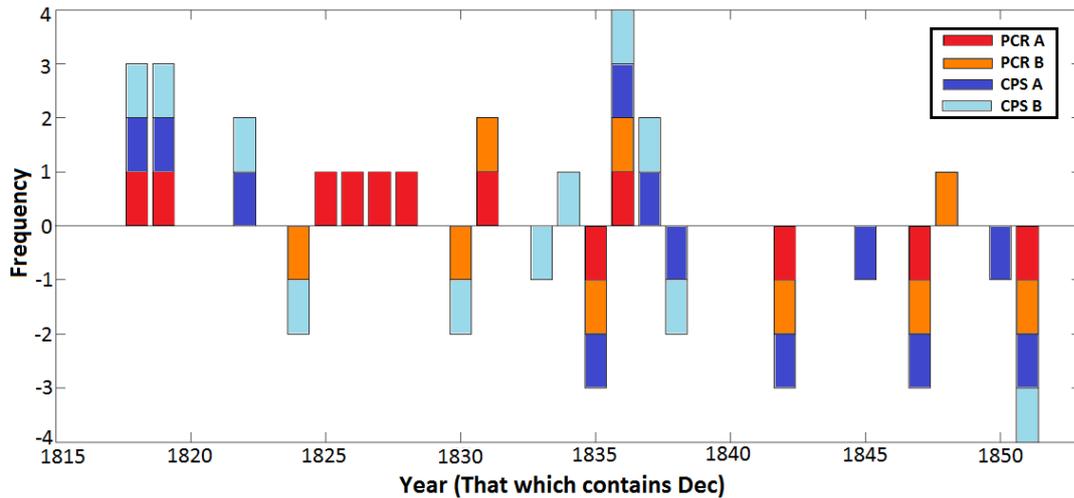


Fig. 2.14 Frequency of El Niño (negative SOI events) and La Niña (positive SOI events) in logbook-based reconstructions, 1815/16 – 1853/54.

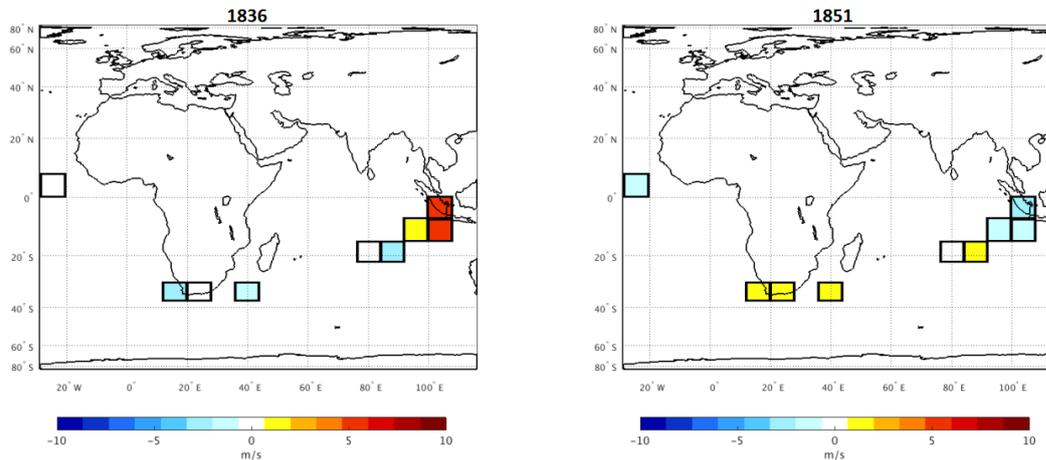


Fig. 2.15 La Niña, 1836/37, (left) and El Niño, 1851/52, (right) zonal wind anomalies (ms^{-1}), relative to the 1834/35 – 1853/54 DJF climatology.

The zonal wind anomalies in the predictor grid boxes in the two high confidence events are shown in Fig. 2.15. There is a clear difference in the signal between El Niño, 1851/52 and La Niña, 1836/37. In 1851/52, negative zonal wind anomalies are found in the three grid boxes closest to Indonesia. This negative wind anomaly was also apparent in the more recent El Niño 1997/98 and in a similar region in 1982/83 (Fig. 2.11). In contrast, strong positive zonal wind anomalies are found in this region during the historical La Niña 1836/37, as well as modern events in 2010/11 and 1998/99 (Fig. 2.12).

The El Niño signal from the grid boxes around the southern tip of Africa is of positive wind anomalies, which is seen more strongly in the 1982/83, modern event than the 1997/98 event, however such a signal fits with that suggested from the modern correlation coefficients between zonal wind and SOI in this region (Fig. 2.5). The signal from these grid boxes is less clear, in the 1836/37 La Niña event, as negative anomalies would be expected. However, this raises the issue of inter-event diversity of ENSO, which is returned to in Chapter 4. The spatial pattern of wind anomalies during different events are not always the same, therefore using data from various regions helps to identify different signals. However, overall similarities are found in the typical spatial wind patterns from both historical and modern events.

Additional caution is needed when analysing those years with PCR RE scores less than 0.3. The limited logbook data availability from these two years resulted in RE scores less than 0.3 when using the PCR method, and are therefore flagged as more uncertain years within the reconstruction. However, the CPS reconstructions in these years also indicate strong ENSO events, which supports the classification of these events within the historical record.

2.5 Discussion

In Chapters 3 and 4 the new logbook-based SOI reconstructions are compared to previous proxy and documentary ENSO reconstructions, and the implications of this comparison for the nature of ENSO reconstruction methods are discussed. However, here the strength and limitations of the methodology are discussed, and the logbook-based reconstruction is compared with a previous SOI reconstruction which is also based on direct meteorological measurements (Können et al. 1998).

2.5.1 Difference in wind force terminology

An important finding of this chapter is the statistically significant difference between the distributions of wind force terms used by the Dutch and EEIC ships. A shift from using observations from mostly EEIC to entirely Dutch records during the period of reconstruction is found, because of the use of both EEIC and CLIWOC data, but in part influenced by the spatial domain. The majority of previous logbook studies focus on the North Atlantic and only use CLIWOC data (Gallego et al. 2005; Küttel et al.

2010). Hannaford et al. (2015) used both CLIWOC and EEIC data from grid boxes around southern African region, but no clear shift was identified in their study.

The major difference in the distributions of wind force terms between the Dutch and EEIC observations is the very limited appearance of Beaufort Force 3 (BF3) in the British records but the common use of it in the Dutch records. Previously, Koek and Können (2005) highlighted a difference in the frequency of Beaufort force values used by different CLIWOC countries, especially in the range of Force 3 to 5. Beaufort Force 3 is described as a ‘Gentle Breeze’ under current British classification, and the CLIWOC multi-lingual dictionary found the following historic terms to be representative of equivalent wind force: ‘gentle breeze’, ‘gentle gale’, ‘gentle trade’, ‘feint gale’, ‘light gale’ and ‘easy gale’, with the latter two uncommon after 1700 (Garcia-Herrera et al. 2003). Of the 22,000 British wind force terms analysed by CLIWOC, less than 0.5% were BF3 equivalent terms (Wheeler and Wilkinson 2005). Of the observations used in the SOI reconstruction, BF3 was only recorded twice in the British wind force records. The reason for this limited use of BF3 by the British ships is unknown. As the Beaufort terms are conversions of descriptive terms by CLIWOC, it may either be because BF3 equivalents were under recorded in the original logbook observations, or that BF3 equivalents were under-converted by CLIWOC. Determination of which of these is true, is outside of the scope of this chapter. In contrast to the British records, Koek and Können (2005) found that Beaufort Force 3 is over-represented in the Dutch records, and this is supported in the results of this chapter. The Beaufort Force 3 equivalent terms in Dutch all include the term ‘bramzeilskoelte’ which refers to the topgallant sail, the upper-most sail that could be used at this wind speed (Koek and Können 2005).

It was also found that in general, the Beaufort Force terms used by the EEIC were higher than those used by the Dutch within the same geographical regions over the same time period. One reason for this is that the method of describing wind force used by the Dutch was distinctly different than that used by EEIC and other countries from within the CLIWOC database. The Dutch method of wind force observations relied heavily on the wind’s influence on the sails used, which provides a more explicit means of determining the order of these terms, making them more refined than is possible in the other languages (Garcia-Herrera et al. 2003). In contrast the British

relied more on the state of the sea to describe wind force (Wheeler and Wilkinson 2005). This different approach to wind determination is suggested as the likely reason for the differences in the distribution of wind force terms used.

Another reason for this difference could be that some of the Dutch CLIWOC data post 1826 are taken from 'Extract Logbooks', meteorological summaries of original logbook data. These Extract Logs were not made until the 1860s, after the adoption of the Beaufort scale in the Netherlands (Koek and Können 2005). The conversions from wind force terms to Beaufort scales were carried out at the time the Extract Logs were created, and these translations were found to be of better quality than those done by the CLIWOC team (Koek and Können 2005). Therefore, the wind force terms taken from these logbooks could be more reliable.

The implications of higher wind force terms from EEIC than Dutch records is a bias towards stronger winds in the period dominated by EEIC and weaker winds in the period of Dutch observations. A tendency for stronger winds in the first half of the record could promote an increased frequency of La Niña events in the first part of the reconstruction, as during La Niña events stronger westerly winds are typical in the predictor grid boxes closest to the Maritime Continent. This was found in all of the historical reconstructions carried out, but more so in those in which the difference in wind force observations was not taken into account. Seven of the eight La Niña events found in the PCR A reconstruction were prior to 1831, thus within the period dominated by stronger EEIC observations and all four of the El Niño events in this reconstruction were post 1835.

The impact of relying solely on Dutch logbooks post 1834 has not been investigated in previous papers which carried out climate reconstructions using logbooks over this period. Wheeler (2005) carried out an assessment of the coherency of wind observations between vessels sailing in convoy and found good agreement. However, a major limitation in this analysis was the tendency for ships in convoy to be from the same country rather than from a mix of nationalities and therefore a mix of recording approaches. The difference between wind observations from different countries has been highlighted here as a potential source of inconsistency and an area that requires further investigation. Less than 5% of British logbooks were exploited by CLIWOC

and none after 1828, therefore a vast amount of records are available for future digitisation, as outlined in a recent survey of logbooks in UK archives (Wilkinson 2009). Digitisation of these British records would result in the ability to carry out reconstructions using data from solely British sources. An analysis of wind force terms and reconstructions using only British observations would enable a fuller assessment of the impact of using a split of Dutch and British records.

2.5.2 Difference in reconstruction methods: PCR and CPS

Reconstructions using both PCR and CPS were carried out to assess the impact of using different methodologies on the resulting historical reconstructions. Recent multi-proxy ENSO studies employed both techniques and found no major differences between the ENSO reconstructions found using CPS and PCR (Wilson et al. 2010). The two methods can be seen as complimentary approaches (Mann et al. 2008). Both methods assume stationarity in the statistical relationships over time, a common limitation in studies of the past (Gergis et al. 2006). Results from this chapter suggest good agreement between the two approaches and both provided good reconstruction skill over the fitting period. Nevertheless, it is worth pointing out that the PCR method has slightly higher skill during the fitting period. Although correlations are strong between reconstructions, the ENSO event chronologies differ significantly between reconstructions. Only two ENSO events of high confidence were found: La Niña 1836/37 and El Niño 1851/52. An additional, seven medium confidence events were found: five El Niño (1824/25, 1830/31, 1835/36, 1842/43 and 1847/48) and two La Niña events (1818/19 and 1819/20). Chapter 3 will analyse the discrepancies in these events-based chronologies further.

2.5.3 Threshold tests on number of observations per season

Threshold tests were carried out to assess the reconstruction skill using a limited number of observations. These showed that, as would be expected, the higher the number of observations used per season, the better the reconstruction skill. Nevertheless, these tests demonstrated that even with very few observations reasonable reconstruction skill can be obtained, with the CPS method performing better than PCR with very limited data availability. Additional analysis which replicated the logbook data available in the predictor grid boxes used in the historical

reconstruction, enabled the reconstruction skill in each year to be evaluated, and identified years within the PCR reconstructions where reconstruction skill is lower than $RE=0.3$. It is noteworthy, however, that the CPR reconstruction for all years had an RE significantly greater than 0.3 (Fig. 2.7). Overall, the threshold results showed that valuable reconstructions can be obtained even when using limited data sources. However, the benefit of further digitisation to increase data density of logbook wind observations is clear, as reconstruction skill can be increased if the number of observations per season can be increased.

2.5.4 Comparison to Jakarta based records of the SOI

While Chapters 3 and 4, focus on a comprehensive comparison with previous ENSO reconstructions, here a comparison of one of the logbook-based SOI reconstructions (CPS B) with the Jakarta rain-day SOI and a single site pressure SOI are presented (Können et al. 1998), as both are from direct observations and can be used together to obtain a more robust ENSO signal from this period. The Jakarta-rain day SOI spans the Indonesian dry season, June-November and the SOI was calculated from a continuous road maintenance record of dry days in Jakarta, 1830-1850. CPS B is focused on for comparison, as the CPS methodology has the higher skill when using sparse data availability (Fig. 2.7) and variant B accounts for the difference in wind force terms between the earlier and later part of the reconstruction period. However, the Pearson's correlation coefficient when comparing the Können et al. (1998) records with the four logbook-based reconstructions can be found in Table A5. The logbook CPS B DJF reconstruction and the Jakarta rain-day SOI are significantly correlated at the 95% confidence level (Fig. 2.16). Over the period 1830 to 1850 the strongest El Niño event in CPS B is during 1833/34. The Jakarta rain-day SOI also indicates 1833 to be the most extreme negative SOI (El Niño) during this period. The strongest La Niña in the CPS B reconstruction during this period is in 1836/37 (and extends into 1837/38). In 1837, the strongest positive SOI (La Niña) is found in the Jakarta rain-day SOI. Therefore, there is a good match between the two most extreme ENSO events during this period. The strong agreement between these records suggests that there is a robust signal from these early instrumental based reconstructions of SOI, giving additional confidence to our logbook-based reconstructions.

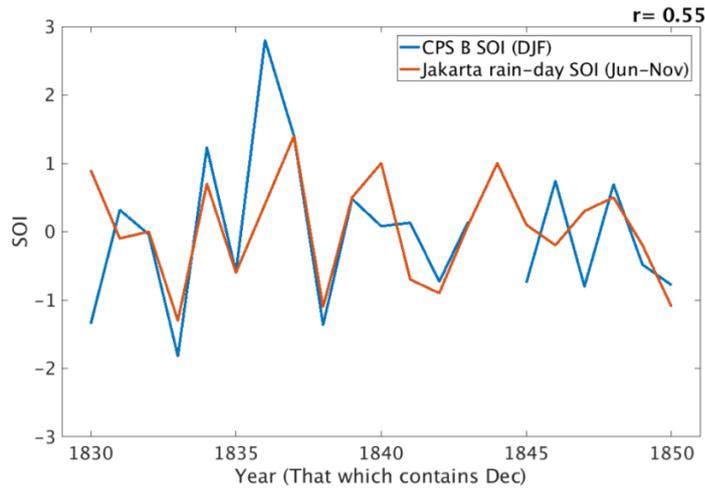


Fig. 2.16 CPS B logbook-based DJF SOI (1830/31 – 1850/51) and Jakarta rain-day June-Nov SOI (1830 – 1850).

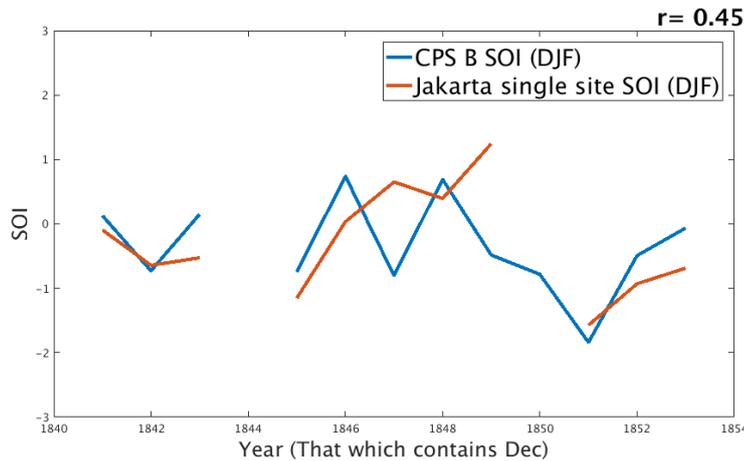


Fig. 2.17 CPS B logbook-based DJF SOI reconstruction and Können et al. (1998) Jakarta single-site DJF SOI (1841/42 – 1853/54). Blank years indicate years with missing data.

The CPS B reconstruction of the DJF SOI along with the Jakarta single site SOI is shown in Fig. 2.17. This single-site SOI record, also from Können et al. (1998), covers a shorter period than the rain-day SOI, spanning from 1841/42-1853/54, and also contains a number of years with missing data. This DJF SOI was calculated using pressure readings from the military doctor’s house based in a botanic garden in Buitenzorg, 50km south of Jakarta. It shows some agreement with CPS B logbook-based SOI but is limited due to its temporal coverage. Overall, there is an indication

of agreement between the records from Können et al. (1998), especially the rain-day SOI with the logbook-based SOI over the period of overlap.

2.6 Summary

The development of ENSO reconstructions using wind observations from ships' logbooks has been presented. Temperature and pressure data from ships' logbooks provide an additional source of data, but are also subject to data availability issues, and additional sources of measurement uncertainty. Therefore, they are not used in the reconstruction approach presented here. Thus, Jones and Salmon's (2005) preliminary logbook-based SOI reconstruction has been built upon by using a larger amount of data, a focus on regions with strong ENSO teleconnections and an analysis of multiple methods of reconstructions. Two different reconstruction methodologies were used, PCR and CPS and both were found to be suitable methods for reconstruction ENSO using data from ships' logbooks. These SOI reconstructions can be found in Table A6.

Methodologies to assess the impact of data availability on reconstruction quality have been developed. It was found that good reconstruction skill can be obtained with limited data availability. As expected, the reconstruction skill increases with the number of observations available. However, it was found that statistically significant skill can be obtained when using surprisingly few observations per season. These results are regionally specific, reflecting the ability of few observations to capture seasonal climate in the regions covering the predictor grid boxes. Therefore, it is recommended that this analysis is repeated by researchers developing reconstructions using logbook data from other geographical regions.

A statistically significant difference in the mean and distribution of the basic Dutch and EEIC wind force observations was found, which leads to a potentially spurious trend in the reconstructions. This is addressed by splitting the study period into two, an earlier period dominated by British records (1815/16-1833-34) and a later period from which only Dutch observations are available (1834/35-1853/54). For these two periods, different climatological values were calculated and applied to the reconstruction methods, thus producing two reconstructions for each statistical reconstruction method. The resulting historical SOI time series had strong correlations with each other, however some differences were found when converting this into an

events-based chronology. The most robust CPS B logbook-based reconstruction compares well to the Jakarta rain-day SOI, (Können et al. 1998) which spans much of the reconstruction period. The strong agreement between these two records indicate a common signal from direct observations, which together can be used to obtain a more robust understanding of ENSO during this period. In Chapter 3 both sets of reconstructions are compared against existing ENSO reconstructions, to determine which method produces a reconstruction with a higher agreement to existing reconstructions.

Digitisation of additional logbook-based records would enable a higher number of observations to be used. The SOI reconstruction skill largely depends on the data availability during a given year, as it varies between grid boxes and seasons. Therefore, the number of observations needed to obtain good reconstruction skill may be different in other seasons and/or geographical regions with different climatological variability, and should therefore be assessed on an individual basis. Fig. A3 shows, for example, that the current SOI reconstruction could be greatly extended in temporal length, in principle back to 1750, with targeted additional digitisation of existing logbooks. There thus remains a potentially rich source of marine records for studies of past climate variability.

3. Comparison of ENSO reconstructions from ships' logbooks with existing multi-proxy and documentary-based ENSO reconstructions

This chapter is based upon the publication 'Barrett HG, Jones JM, Bigg GR (2017) Reconstructing El Niño Southern Oscillation using data from ships' logbooks. Part II: Comparisons with existing ENSO reconstructions and implications for reconstructing ENSO diversity. Climate Dynamics DOI 10.1007/s00382-017-3797-4.' and includes text and figures from this publication.

3.1 Introduction

Understanding of pre-instrumental variability in El Niño Southern Oscillation is a key challenge for paleoclimatology. As discussed in Chapter 1, instrumental records are too short to fully capture the natural range of ENSO variability and to assess ENSO behaviour in the pre-industrial era (Braganza et al. 2009). Therefore, in order to understand the natural variability of ENSO, proxy records and historical documents that span beyond the instrumental period need to be used (Gergis et al. 2006). Chapter 2 presented new ENSO reconstructions which utilise the wind observations found within ships' logbooks during the period 1815/16 to 1853/54. The observations from ships' logbooks provides a record with low dating uncertainty and are based on direct, *in-situ* observations of the weather, thus providing a unique reconstruction from a period when routine observations of the weather were uncommon over land. Four historical reconstructions of the boreal winter, December, January, February (DJF), SOI, were presented in Chapter 2 using methods of Principal Component Regression (PCR) and Composite-Plus-Scale (CPS). The logbook-based reconstructions were found to have good agreement with a Jakarta-based SOI, which uses rain-day counts, from 1830-1850 (Können et al. 1998), suggesting a common signal from these early observation-based records. This chapter will compare the reconstructions from ships' logbooks to previous proxy and documentary-based reconstructions of ENSO, in order to add to the understanding of ENSO variability during the early to mid-nineteenth century.

Proxy data are natural archives of the climate and environment of the past. They are indirect measures of past climatic changes and have been used extensively in the reconstruction of ENSO (Gergis et al. 2006). Here, ENSO reconstructions that cover the historical period which coincides with the reconstruction from ships' logbooks are focused upon. Recent ENSO reconstructions take a multi-proxy approach, using a range of data sources including corals, tree rings, speleothems and ice-cores (McGregor et al. 2010; Emile-Geay et al. 2013a). The multi-proxy approach captures ENSO signals from different regions and from different sources. This helps to address the problem of the assumption of stationarity in the relationships between ENSO and the climate signal in the regions in which the proxies are located. This is useful as no two ENSO events, or their resulting teleconnections, are the same so combining the different signals from multiple regions provides the most robust ENSO signal (Braganza et al. 2009; Capotondi et al. 2015). The differences in the spatial distribution of proxy data used in teleconnection regions can result in differences in the ENSO reconstructions. Therefore, it is important to consider the locations of the proxies from which reconstructions are made, and the implications of these differences.

Additional indications of ENSO behaviour in the pre-instrumental era can be obtained from documentary records. Documentary records are the focus of studies within historical climatology and include sources such as weather journals, governmental and military records, and documents from religious institutions. These historical documents are 'human' climate proxies and provide data from sources that were often not originally intended for use in climatology (Nash and Adamson 2014). The work of Quinn et al. (1987) played a key role in promoting the use of historical documents for El Niño reconstruction. Quinn et al. (1987) created an El Niño events-based chronology using publications from the west coast of northern South America and the neighbouring parts of the Pacific Ocean, extending back to AD 1525. This record provides a benchmark in historical ENSO reconstruction studies (Gergis and Fowler 2009).

The observations from ships' logbooks are an alternative source of documentary data, with weather observations taken daily on board ships from across the World's Oceans. Although the use of data from ships' logbooks has increased in the field of historical

climatology recently, there has only been one prior use of this data source which directly focused on ENSO reconstruction before that presented in Chapter 2. Jones and Salmon (2005) developed a preliminary reconstruction of the SOI using ships' logbook data, but were hindered by poor data availability. They found weak correlations with previous ENSO reconstructions over the period 1750 to 1850. The logbook-based SOI reconstructions presented in Chapter 2 built upon Jones and Salmon's (2005) initial reconstruction by using additional data, concentrating on regions with strong ENSO signals and developing their methodologies further. Therefore, comparison of the logbook-based reconstructions with their work, and the proxy studies with which they compared their SOI index, investigates the improvement of these changes to the reconstruction methods.

This chapter is organised as follows: *Section 3.2* introduces the range of proxy and documentary reconstructions used for comparison to the logbook-based ENSO reconstruction. *Section 3.3* describes the methods of comparison used. *Section 3.4* presents the results of the ENSO reconstruction comparisons, with a focus on the continuous ENSO indices from proxy sources, followed by events-based chronologies from both proxy and documentary evidence. *Section 3.5* summarises the key findings and areas of uncertainty, while *Section 3.6* concludes this chapter and indicates areas of further research. This chapter, along with Chapter 4, addresses Aim 2 and Objectives 2A-2D, as outlined in Chapter 1 which focus on the comparison of the logbook-based ENSO reconstructions to previous documentary and multi-proxy reconstructions of ENSO.

3.2 Data

3.2.1 Logbook-based DJF SOI reconstructions

Four reconstructions of the DJF SOI using wind observations from ships' logbooks were presented in Chapter 2 (Fig. 2.13). These extend from 1815/16 to 1853/54, a period dictated by the availability of a sufficient number of digitised logbook observations. The wind force and direction observations were taken from the ships' logbooks and converted into seasonal zonal wind anomalies at a 7.5° by 8° latitude-longitude gridded resolution. Nine predictor grid boxes from a number of regions were selected due to the existence of statistically significant ($p < 0.1$) correlation between

ERA-Interim zonal wind and SOI during the modern period 1979 to 2013, and suitable logbook data availability. These grid boxes are located in the Indian Ocean, around the southern tip of Africa and in the Atlantic Ocean. Firstly, PCR was carried out, followed by a simpler method of CPS. Two unadjusted reconstructions were obtained and labelled PCR A and CPS A, with variant A referring to reconstructions which were based on a climatological average calculated over the entire reconstruction period, 1815-1854. These methods were then repeated, but with adjustments made to the climatological means to take into account the shift from dominance of English East India Company observations in the first part of the record, to Dutch observations in the later half. A statistically significant difference in the wind force observations between vessels from these two countries was found in Chapter 2 (Fig. 2.9), and the adjustment in climatological means attempts to address this bias. Therefore, a separate climatological mean was applied to the earlier period (1815-1833) and the later period (1834-1853), and the two time series combined. The resulting reconstructions are labelled variant B, PCR B and CPS B. Chapter 2 suggested that the CPS B method performs best, based on higher RE skill scores from the CPS method and the adjustment in climatology to reduce the bias. However, all four are investigated here, as although an artificial change in the mean is strongly suspected, it is not completely certain that this change is not real.

Over the modern era, 1979-2013, ERA-Interim data were used to test the two reconstruction methods. It was found, in Chapter 2 (Fig. 2.10), that using the full ERA-Interim dataset over this period, strong positive correlation coefficients with the instrumental SOI are obtained for both PCR and CPS. Additionally, in order to obtain a dataset which best represents the data availability of the logbook-based SOI reconstructions during the modern era, ERA-Interim is sub-sampled at the availability similar to that from the logbook era in the predictor grid boxes. This was carried out in Chapter 2 in order to assess the sensitivity of the reconstruction methods to limited data availability. It was found that even with a limited number of observations a reasonable level of reconstruction skill can be obtained using the PCR and CPS methods. Therefore, a sub-sampled PCR and CPS time series which replicates this data availability in the modern period, 1979-2013, is also used in comparison to the instrumental record of ENSO. The sampling replicates the average number of

observations per predictor grid box over the logbook period, 1815-1853 (Table 3.1) for one random iteration and the mean of five iterations as in Küttel et al. (2010). This provides the best possible replication and representation of the logbook-based reconstruction during the modern period and therefore allows the most realistic comparison between analysis of the modern period and that of the logbook period.

Table 3.1 Mean number of DJF logbook observations per grid box (1-9), 1815-1854 used in sub-sampled ERA-Interim data.

Grid box	1	2	3	4	5	6	7	8	9
Average number of logbook wind observations (1815-1853)	20	13	16	9	8	9	9	18	17

As in Chapter 2, a seasonal SOI was calculated from monthly values of the instrumental pressure difference between Tahiti and Darwin that is available from 1876 to present from the Australian Bureau of Meteorology (BoM 2015). The SOI for boreal winter (DJF) season is used as a key indicator of ENSO behaviour.

3.2.2 Multi-Proxy ENSO reconstructions

A range of multi-proxy ENSO reconstructions exist from previous studies. There is no consensus on which multi-proxy ENSO reconstruction provides the most realistic record of past ENSO behaviour. Proxy data sources which have been used to reconstruct ENSO include tree-rings, ice cores, corals and multi-proxy databases (Gergis and Fowler 2009; McGregor et al. 2010; Emile-Geay et al. 2013a). Methods of proxy-based reconstruction have a range of uncertainties and limitations associated with them, such as dating errors, local environmental influences and non-stationary teleconnections. Therefore, comparison and calibration in multi-proxy analysis is suggested as the best way to gain a comprehensive record of climatic changes from the range of sources (Gergis and Fowler 2009).

Stahle et al. (1998) used tree rings from southern USA and Mexico, a region with strongly established teleconnections to ENSO, along with one teak tree from Java to reconstruct a boreal winter (DJF) Southern Oscillation Index extending from 1706-

1997. This ENSO reconstruction is directly comparable to the logbook-based reconstructions as both produce a time series of DJF SOI, and PCR methods were used in both studies. This was one of the reconstructions used by Jones and Salmon (2005) for comparison with their SOI reconstruction. Mann et al. (2000a), also used by Jones and Salmon (2005) for ENSO comparison, present a reconstruction of Oct-March Niño 3 SST, 1408-1978, compiled from 112 annually resolved proxies. The use of such annually resolved proxies provides good comparability to the annual logbook-based SOI reconstruction. In both cases, use of proxies from tropical and extra-tropical regions ensures multiple ENSO teleconnections are used to inform the reconstruction.

McGregor et al. (2010) used 10 commonly used ENSO reconstructions, including Stahle et al. (1998) and Mann et al. (2000a). This combined ENSO reconstruction resulted from the use of 155 proxies, including tree rings, corals and ice cores from across the Pacific and from key teleconnection regions such as the United States. The inclusion of the data used by the reconstructions by Stahle et al. (1998) and Mann et al. (2000a) within McGregor et al.'s (2010) proxy index means that an additional degree of similarity will be expected when comparing these three reconstructions. The locations of these three proxy networks can be seen in Fig. 3.1.

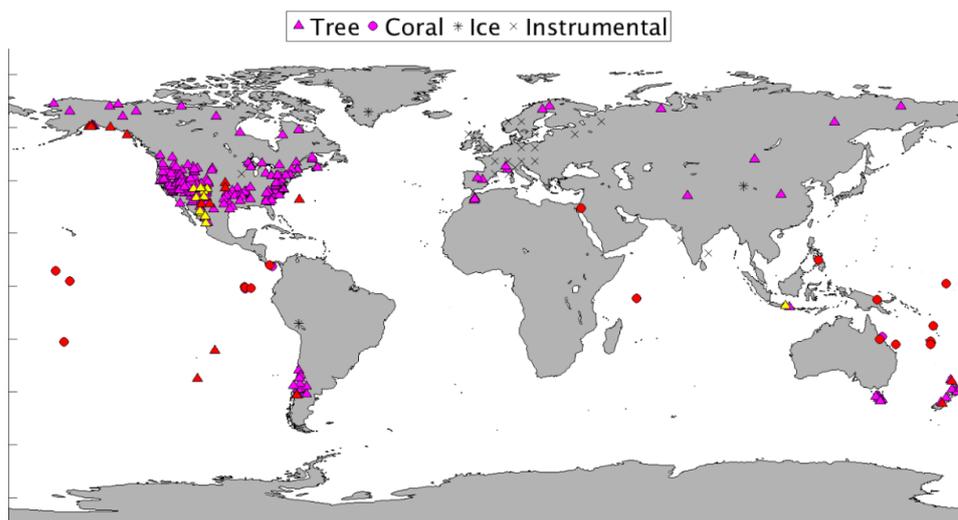


Fig. 3.1 Location of proxies used in ENSO reconstruction studies by ST98 (yellow triangles only), MANN00 (purple) and MCG10 (yellow, purple and red), with the exception of Cook et al. (2008) used in MCG10 as specific location data unavailable. Shapes indicate different types of proxy records (triangle = Tree-rings, circle = corals, star = ice cores, cross = instrumental).

Wilson et al. (2010) provide a normalised Niño 3.4 reconstruction which uses annually resolved proxy data from teleconnection regions within the tropics. Eleven coral proxies and one tropical ice core record were used for their teleconnection reconstruction which spans from 1540 to 1998 (Wilson et al. 2010). Therefore, a smaller number of proxies are used compared to McGregor et al. (2010), however the focus on tropical teleconnections only suggests that there may be less noise in the ENSO signals from the proxies used in this reconstruction. The proxies used are located mostly in the western Pacific and Maritime Continent, with a few in the central Pacific, Fig. 3.2.

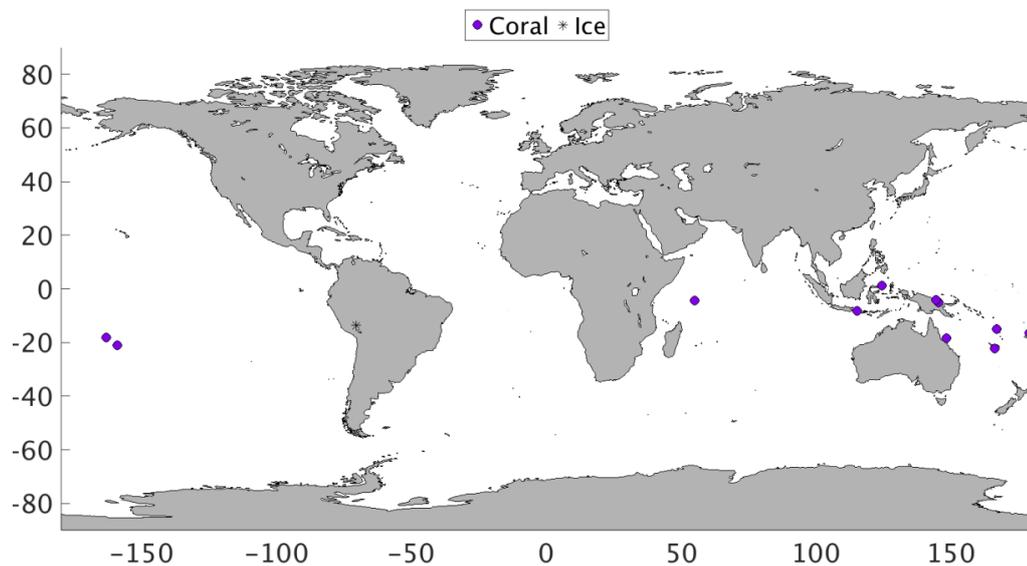


Fig. 3.2 Location of proxies used in the ENSO reconstruction of W10, Tropical Teleconnection ENSO reconstruction. Shapes indicate different types of proxy records (circle = corals, star = ice cores).

Emile-Geay et al. (2013a) analysed 57 proxies from the centre of action of ENSO and a number of teleconnection regions including the Maritime Continent, Southern Africa and Mexico/Southern USA. A difference in this reconstruction was its focus on low-frequency ENSO variability and the inclusion of proxies with a temporal resolution of 5 years or less and with dating errors of up to ± 5 years. In the annually resolved proxies, the dating errors would be expected to be less and may, therefore contain less uncertainty than this low-frequency reconstruction (McGregor et al. 2010). Another criterion for their proxy selection was for them to be equatorward of 35° , which differs from the other multi-proxy reconstructions, which use some records from higher

latitudes. They present three different Niño 3.4 SST (DJF) reconstructions, based on the first principal component of three historical SST analyses. Of the initial 57 proxies, reduced networks are used for each of the three reconstructions based on existence of significant correlation between the proxy record and the Niño 3.4 from each historical SST (Emile-Geay et al. 2013a). For the purpose of this thesis, these three reconstructions will be referred to as EG13A, EG13B and EG13C relating to that produced from ERSSTv3 (Smith et al. 2008), HadSST2i (Rayner et al. 2006) and Kaplan SST (Kaplan et al. 1997) respectively. The screening process resulted in 35 (EG13A), 36 (EG13B), and 37 (EG13C) records used in each case, as seen in Fig. 3.3, with the networks differing by at most four records.

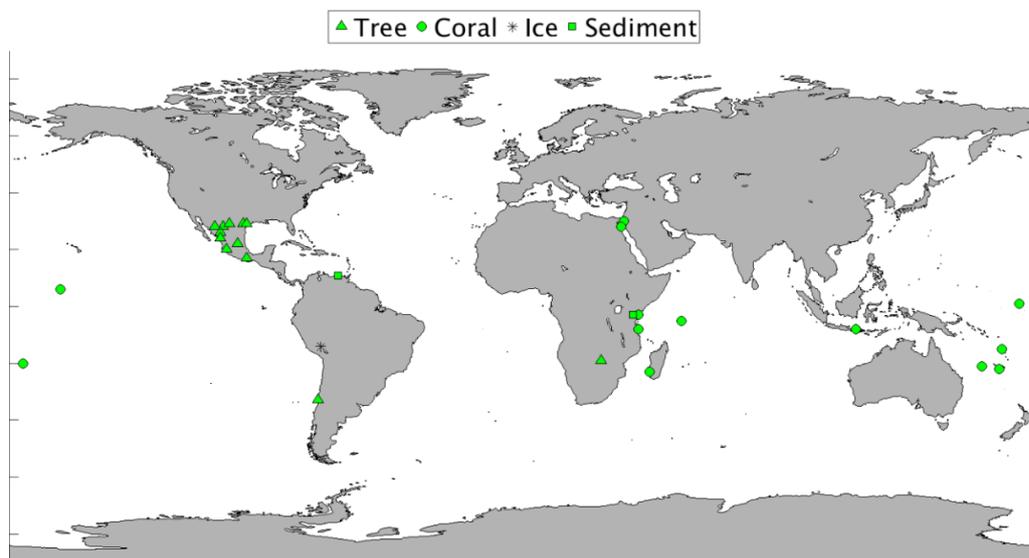


Fig. 3.3 Location of proxies used in ENSO reconstruction studies by EG13A (35 of these), EG13B (36) and EG13C (37). Shapes indicate different types of proxy records (triangle = Tree-rings, circle = corals, star = ice cores, square = sediment).

Li et al. (2013) also made use of a large network of proxy data sources by reconstructing a seven century long reconstruction of ENSO from 2,222 tropical and sub-tropical tree rings. They reconstructed the mean Niño 3.4 SST anomaly (November, December, January, NDJ). The data is clustered in seven main regions, as seen in Fig. 3.4: Central Asia (Monsoon Asia Drought Atlas (MADA PC1), Maritime Continent (MADA PC2), Southwest North America (North American Drought Atlas (NADA PC1), Pacific Northwest/Texas-Mexico, northern New Zealand, South American Altiplano and west-central Argentina.

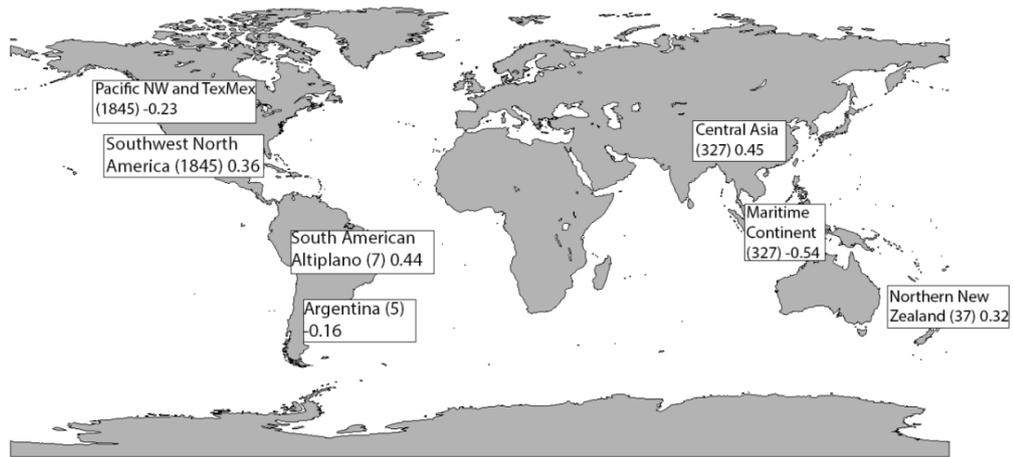


Fig. 3.4 The seven regions containing proxies used in the ENSO reconstruction by LI13. Number in parentheses indicates number of Tree rings in each region. Also shown is the weighting of the data from each region.

The eight multi-proxy reconstructions used in this chapter, and in Chapter 4, are summarised in Table 3.2 along with the abbreviations for each of these which are hereafter used to refer to these reconstructions. Datasets from these ENSO reconstructions are available from the US National Oceanic and Atmospheric Administration’s paleoclimatology archive (<https://www.ncdc.noaa.gov/paleo-search>). The period covered by all of these proxy data sources spans from 1706 to 1977, thus covering the logbook-based reconstruction period (1815-1854), and a period of overlap with the instrumental SOI (1876-1977).

Table 3.2 Proxy based ENSO reconstructions used in this thesis.

	Start year	End year	Index	Proxy types	Reference
ST98	1706	1977	DJF SOI	Tree-ring: Early-wood and total ring-width	Stahle et al. 1998
MANN00	1408	1987	Niño 3 SST (Oct-Mar)	Coral, ice-cores, tree-rings	Mann et al. 2000a
MCG10	1650	1977	Unified ENSO Proxy July-June	Corals, tree rings, ice core	McGregor et al. 2010
W10	1540	1998	Niño 3.4 (Nov-Dec)	Corals, ice-core	Wilson et al. 2010
EG13A/B /C	1150	1995	Niño 3.4 SST (DJF)	Corals, tree-rings, ice core, speleothems, sediment	Emile-Geay et al. 2013a
LI13	1301	2005	Niño 3.4 SST (NDJ)	Tree-rings	Li et al. 2013

The proxies used within these studies are sensitive to different climatic variables and are located in a range of teleconnection regions. Key variables driving proxy records include precipitation (tree rings, ice cores and corals), air temperature (tree rings and ice cores) and SST (Coral $\delta^{18}\text{O}$) (Mann et al. 2000a). The differences in these teleconnections during ENSO events could be influencing the ENSO reconstructions. Hereid et al. (2013) found corals used in a number of ENSO reconstructions originating from different regions have better or worse skill at reconstructing either El Niño or La Niña. The difference in skill at reconstructing El Niño and La Niña events is assessed further here using these ENSO reconstructions. Following on from this proxy ENSO skill assessment, the skill of the various reconstructions at capturing EP and CP events is assessed in Chapter 4.

3.2.3 Documentary ENSO reconstructions

A number of documentary-based ENSO reconstructions have been published over the past few decades, making use of archived documents that provide indications of weather conditions typical of ENSO events. Quinn et al. (1987) presented the first of these, with an extensive analysis of South American records extending back to AD 1525. This focused on El Niño events only and used key indicators such as heavy rainfall, flooding, the effects of winds and currents on ships, and changes in the productivity of coastal fisheries to infer El Niño events. Cross checks between sources were carried out to help establish the reliability of the source.

However, since its publication, Quinn's El Niño chronology has been subject to revisions and the reliability of some of the sources used have been placed under scrutiny, with reliability found to reduce prior to 1800 (Thompson et al. 1993). Ortlieb (2000) examined the chronology, questioning some of the sources used and the teleconnections on which they were based. For example, El Niño events that had been identified based on flooding events on the Rimac River, western Peru, were removed as there is no evidence of an association with floods along this river and El Niño events found within the instrumental record. Events relating to abundant rainfall in southeastern Peru were also removed as they are now thought to be indicative of La Niña events (Nash and Adamson 2014). Overall, 42 of the 86 El Niño events from Quinn's original chronology were questioned, 25 were suggested to be removed and seven previously unrecognized events were to be added (Ortlieb 2000). This highlights the problems and ambiguities associated with historical chronologies of ENSO using documentary sources. To improve this, Garcia-Herrera et al. (2008) carried out documentary analysis back to 1550 using only primary documentary sources from a local region of Northern Peru with previously established associations between ENSO events and heavy rainfall, with impacts such as large river discharge during past El Niño events.

Table 3.3 Documentary-based ENSO chronologies used in this thesis.

	Start year	End year	Index	Location of sources	Reference
QUINN87	1525	1987	El Niño events	Northern Peruvian coastal region and adjacent waters	Quinn et al. 1987
ORT00	1525	1900	El Niño events	Northern Peruvian coastal region and adjacent waters	Ortlieb 2000
GH08	1550	1900	El Niño events	Northern Peru	Garcia-Herrera et al. 2008

The three documentary-based El Niño reconstructions used for comparison to the logbook-based SOI reconstruction are listed in Table 3.3. These provide events-based chronologies of El Niño from 1525 onwards. A disadvantage is the focus only on El Niño events, with La Niña events not identified in these studies. Another limitation of these three historical reconstructions is their focus on only one main region, Peru and surrounding regions in South America. Collecting documentary evidence from this region is justified as it is a key region which is commonly impacted by ENSO, however previous studies have suggested a broader spatial coverage helps to fully capture ENSO events and diversity within reconstructions (Braganza et al. 2009). A number of the multi proxy reconstructions have very few records from South America, therefore the core data regions vary between the documentary (Table 3.3) and multi-proxy (Table 3.2) reconstructions.

The logbook-based reconstructions use data from the Indian Ocean, South of Africa and the Atlantic. Therefore, they are more representative of ENSO signals from

different regions than the previous documentary-based reconstructions. However, despite these differences, documentary sources often have fewer dating uncertainties associated with them than proxy records, so provide a good source of comparison for the logbook-based reconstructions as their dating errors are also minimal. Finally, Gergis and Fowler (2009), GF09, present a chronology of ENSO events since AD 1525, the same period as the Quinn chronology, which looked at both El Niño and La Niña events. They built upon an earlier attempt by Whetton et al. (1996) to document an historical record covering both phases of ENSO. They used a range of complementary proxy data and documentary evidence. This record provides an events-based chronology similar to that available from the documentary based records, and was therefore a useful addition to the comparison of events based chronologies. Events classified as weak events by GF09 were not used here due to conflicting signals which resulted in a given year been classified as both El Niño and La Niña. For all other documentary chronologies, the events, regardless of strength, were used for comparison.

A number of previous ENSO-multi-proxy comparison studies provide a comprehensive overview of this topic, notably Gergis et al. (2006). The comparisons within this chapter build upon this work, and provide an up-to-date ENSO comparison using a number of key multi-proxy reconstructions published over the past few years (MCG10, W10, EG13, LI13) as well as some which were included in previous comparisons. Thus, these provide the most up-to-date source of comparison for the new logbook-based reconstruction.

3.3 Methods

Comparisons were carried out between the different ENSO reconstructions over a number of periods in which data overlap (Fig. 3.5). Firstly, the fitting period 1979-2013 was used to assess the event-capture skill of the two reconstruction methods applied to the logbook data, PCR and CPS. This was tested with a full dataset and sub-sampled in order to provide the most realistic comparison to the logbook data. The sub-sampled dataset is a reconstructed SOI calculated using a sub-sample of the ERA-Interim data which replicates the average number of observations per DJF season in the period of the logbook-based reconstruction (Küttel et al. 2010). This is tested using

one iteration and the mean of five iterations, as in Küttel et al. (2010). Secondly, a period in which the multi-proxy reconstructions overlap with the instrumental DJF SOI (1876-1977) was assessed. Analysis using longer twentieth century reanalysis for a logbook-type reconstruction over this period is not carried out due to additional sources of uncertainty when using twentieth century reanalyses such as the Twentieth Century Reanalysis Project (20CRv2c) which spans back to 1851 but is informed from a smaller number of observations (Compo et al. 2011; Ferguson and Villarini 2014; Poli et al. 2016; Bett et al. 2017). This is discussed further in Chapter 5. The documentary chronologies have a more limited period of overlap with the instrumental SOI, with ORT00 and GH08 only providing an ENSO chronology up to 1900. As a result, only the QUINN87 record, which extends up to 1987, was compared to the instrumental SOI and events chronologies. The events-based chronology provided by GF09 was also used for comparison over this period. Finally, the logbook-based reconstruction period, 1815/6 to 1853/4 was used for comparison between the new logbook-based SOI reconstructions and the previously published ENSO reconstructions during this time period.

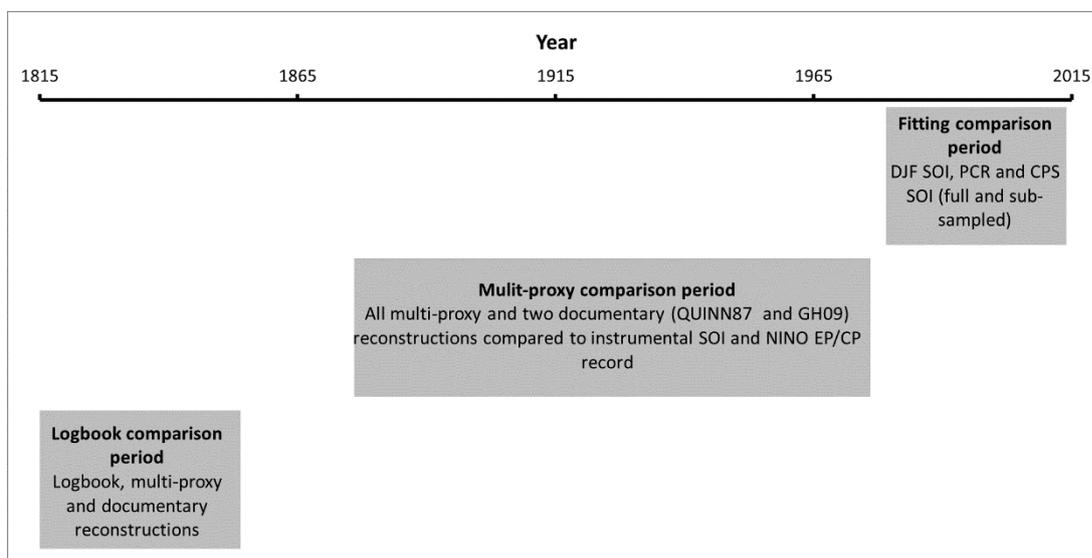


Fig. 3.5 Time periods over which comparisons are made between various ENSO reconstructions and records.

The multi-proxy reconstructions used in this comparison are based on a range of ENSO indices. ST98 provide a DJF SOI reconstruction, directly comparable to that obtained from the logbook-based reconstructions, whereas MANN00 reconstructs the

Niño 3 SST index over Oct-March and W10, EG13 and LI13 reconstruct the Niño 3.4 SST index, Nov-Dec, DJF and NDJ seasons respectively. Finally, MCG10 present a unified ENSO proxy from July-June. Although these all cover the majority of the DJF period analysed by the logbook-based reconstructions, some discrepancies may occur as a result of slightly different seasons used for reconstruction, most likely with W10 and MCG10 which cover the entire year. To compare the various reconstructions Pearson's correlation coefficients were calculated during the overlap period (1876-1977) and the logbook-based reconstruction period (1815-1854). The SST-based reconstructions were inverted to take into account the negative correlation between the SOI and the SST-based Niño indices during the instrumental record.

The multi-proxy and logbook-based reconstructions are continuous ENSO time series. In contrast, the documentary-based records provide an events-based chronology, listing El Niño and in some cases La Niña events. Therefore, El Niño and La Niña events had to be defined from the ENSO indices to enable comparison with the events-based chronologies. The common way of defining ENSO events from continuous data is through the use of standard deviation thresholds. In their reconstruction of the DJF SOI, Stahle et al. (1998) investigated both ± 1.0 and ± 2.0 standard deviations from the mean, as thresholds for defining ENSO events. When using Niño SST indicators of ENSO, it is typical to define an ENSO event based on a 5-month running mean of SST anomalies, using a threshold of $\pm 0.5^{\circ}\text{C}$, which is equal to roughly 0.5 standard deviations from the modern climatological Niño 3.4 mean (McGregor et al. 2010). However, the one standard deviation threshold is used in a number of previous studies (Mann et al. 2000a, Braganza et al. 2009). Therefore, here the one standard deviation threshold is used as an indicator of ENSO events, with means and standard deviations calculated per series.

Once event-based chronologies were constructed, hit rates were calculated to assess ENSO detection skill. This skill is defined as the percentage of correctly identified events in the target record by the index under comparison. The target record varies depending on the period of analysis. For example, the target record in the fitting period (1979-2013) and historical period (1876-1977) is the instrumental DJF SOI. Hit rates have been used in previous ENSO studies as an indication of skill when analysing ENSO on an event-by-event basis (Gergis and Fowler 2009; McGregor et al. 2010).

A hit rate of over 40% is suggested by Hereid et al. (2013) as one which indicates good skill. Year₀ of the ENSO event refers to that of the December of the DJF season, therefore assuming the El Niño impacts preceded the peak in the El Niño event. Year₁ analysis compares the DJF SOI with the year of January and Feb, therefore assuming the El Niño impacts followed the peak in the El Niño event. For multi-proxy analysis DJF 1815/1816 is compared to 1816 from the proxy reconstructions. For this period of logbook data, correspondence rates replace hit rates. These indicate how well the multi-proxy and documentary reconstructions correspond with the logbook-based reconstructions. Correspondence rates were used for this period as there is no ‘true’ (i.e. instrumental) ENSO index to compare to the reconstructions.

3.4 Results

3.4.1 Fitting period events analysis, 1979-2013

Firstly, the fitting reconstructions from both PCR and CPS methodologies were compared to the instrumental DJF SOI, assessing the event capture ability of these indices. Chapter 2 found that the two methods used in the logbook-based reconstructions produced an SOI which was strongly and significantly correlated with the observed DJF SOI, PCR ($r=0.89$) and CPS ($r=0.87$). The non-detrended DJF SOI values were used to create an events-based chronology, based on the one standard deviation threshold. The PCR (CPS) method identifies 8 (6) El Niño and 5 (4) La Niña events over the period 1979-2013, whilst the instrumental DJF SOI identifies 6 El Niño and 6 La Niña events (Fig. 3.6). Hit rates, indications of El Niño and La Niña skill, were calculated based on the number of events in the instrumental DJF SOI captured by the reconstruction indices for El Niño and La Niña events (Table 3.4). Hit rates are higher for El Niño (83%, 67%) than La Niña events (50% for both methods), and PCR hit rates for El Niño are higher than that for CPS. They are all higher than the hit rate of 40% used as a threshold for indicating a relatively high level of skill (Hereid et al. 2013). The better correspondence of the PCR-based reconstruction with the instrumental record was also found in the correlation analysis. Fig. 3.6 shows the ENSO events from the three different datasets and helps to illustrate the hit rates. It also shows that both reconstruction methods create ‘False alarm’ events, ones which appear in the reconstructed SOI but not in the observed record. The false alarm rate

indicates the percentage of events predicted by the PCR or CPS that were not events within the instrumental SOI. False alarm rates of 25% for El Niño and 40% for La Niña for PCR and 33% for El Niño and 50% for La Niña for CPS were found.

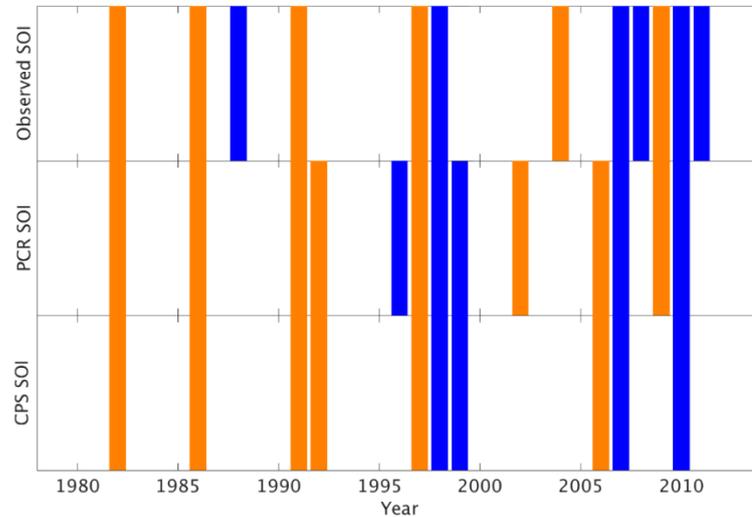


Fig. 3.6 El Niño (orange) and La Niña (blue) events from the instrumental DJF SOI and the full PCR and CPS fitting reconstructions, 1979-2013.

Table 3.4 Hit rates for El Niño and La Niña events from PCR and CPS fitting full and mean of five sub-sampled iterations compared to the instrumental DJF SOI.

	El Niño Skill %	La Niña Skill %
PCR fitting	83	50
PCR sub-sampled (5 it.)	83	33
CPS fitting	67	50
CPS sub-sampled (5 it.)	83	33

The sub-sampled PCR and CPS fitting reconstructions, which were constructed using a sample of ERA-Interim data that replicates the mean data availability over the reconstruction period, provide an indication of the hit rates which would be obtained for a logbook-based reconstruction over this modern period. For El Niño events the hit rates for the sub-sampled PCR are the same as that for the full PCR fitting, (83%), while for the sub-sampled CPS, the El Niño hit rates are higher (83%) than that from the fitting (67%). La Niña hit rates are lower for the sub-sampled than for the full fitting for both methodologies (Table 3.4). The performance of the two methodologies

for El Niño events is clearly better than their performance for La Niña events. Overall, the sub-sampled fitting produces an ENSO record that has less correspondence with the DJF instrumental SOI, but still captures events well, especially El Niño events, with hit rates not dropping below 67% for either reconstruction method.

It is possible that reconstruction methods could result in a bias towards El Niño rather than La Niña or vice versa, as previously found for individual coral records assessed across the Pacific (Hereid et al. 2013). Hereid et al. (2013) found that coral proxy records from different regions across the Pacific have varying skill at capturing El Niño and La Niña events, with some regions such as the Western Pacific Warm Pool favouring El Niño events and the South Pacific Convergence Zone more sensitive to recording La Niña events. Therefore, a non-symmetric response to El Niño and La Niña events could be causing a bias within various reconstructions. This is explored here, focusing on the various ENSO reconstructions assessed within this chapter. Further analysis focusing on the biases within reconstructions due to the differences in teleconnections associated with the two types of El Niño events is presented in Chapter 4.

Over the modern period, 1979-2013, hit rates are higher for El Niño than La Niña events for the two reconstruction methodologies (Table 3.4). To further investigate this potential bias, the spatial wind patterns of individual ENSO events were analysed, specifically within the nine predictor grid boxes. Composite plots of the DJF mean zonal wind anomalies for El Niño and La Niña events, defined using the one standard deviation of the instrumental DJF SOI, were constructed using ERA-Interim Reanalysis data based on 1979-2013 climatology (Fig. 3.7). Spatial patterns during these individual events are shown in Fig. 3.8 and Fig. 3.9, and demonstrate differences between individual El Niño events and between individual La Niña events respectively.

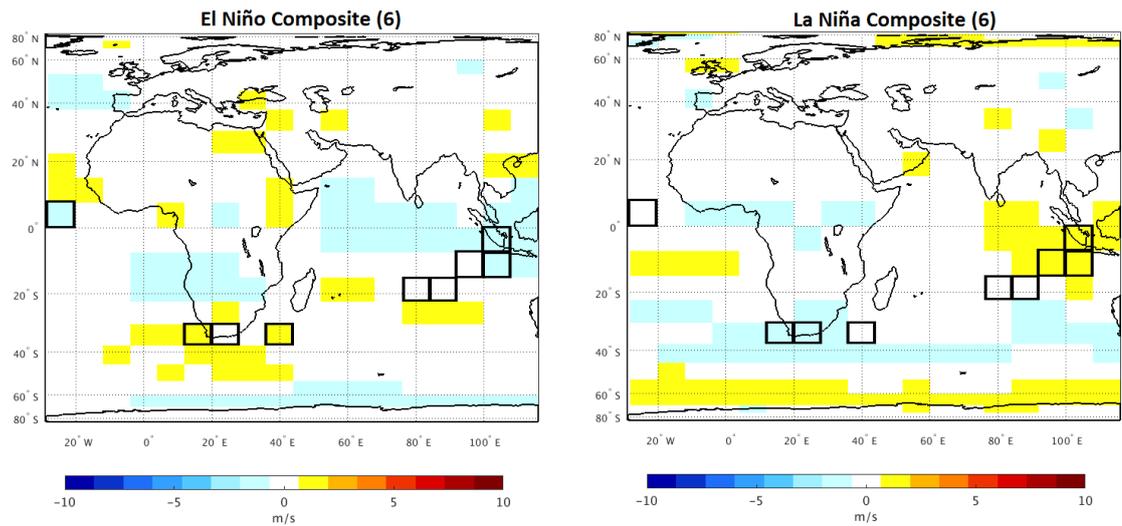


Fig. 3.7 DJF zonal wind composite anomalies (ms^{-1}) during El Niño events (1982/83, 1986/87, 1991/92, 1997/98, 2004/05 and 2009/10) and La Niña events (1988/89, 1998/99, 2007/08 2008/09, 2010/11, 2011/12), relative to the 1979-2013 climatology. Zonal wind data are taken from the ERA-Interim reanalysis. Predictor grid boxes used in the reconstruction are outlined in black, numbered 1-9 from west to east.

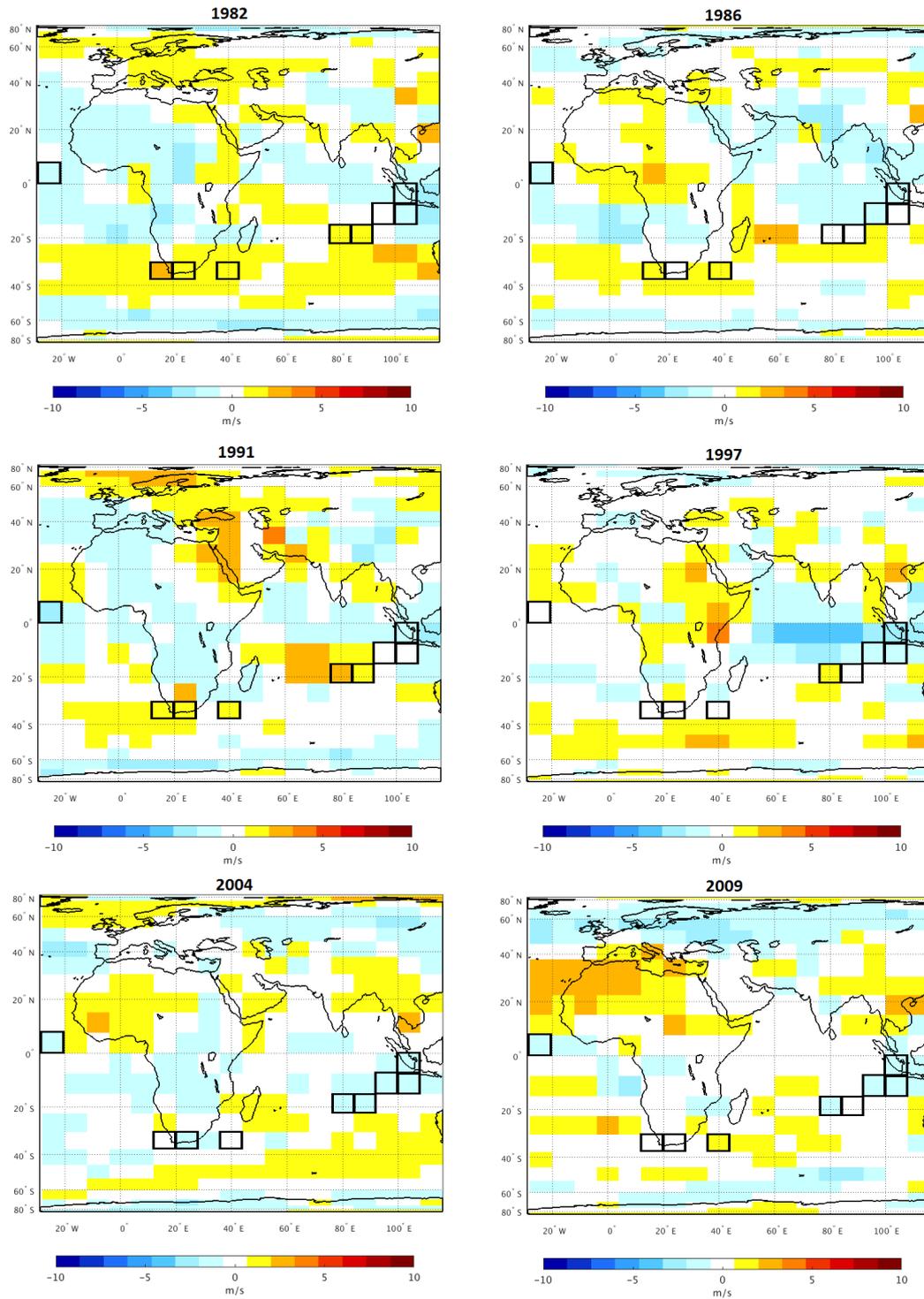


Fig. 3.8 DJF zonal wind composite anomalies (ms^{-1}) during El Niño events defined using the one standard deviation threshold of the instrumental SOI (1982/83, 1986/87, 1991/92, 1997/98, 2004/05 and 2009/10, relative to the 1979-2013 climatology). Zonal wind data taken from ERA-Interim reanalysis. Predictor grid boxes used in the reconstruction are outlined in black.

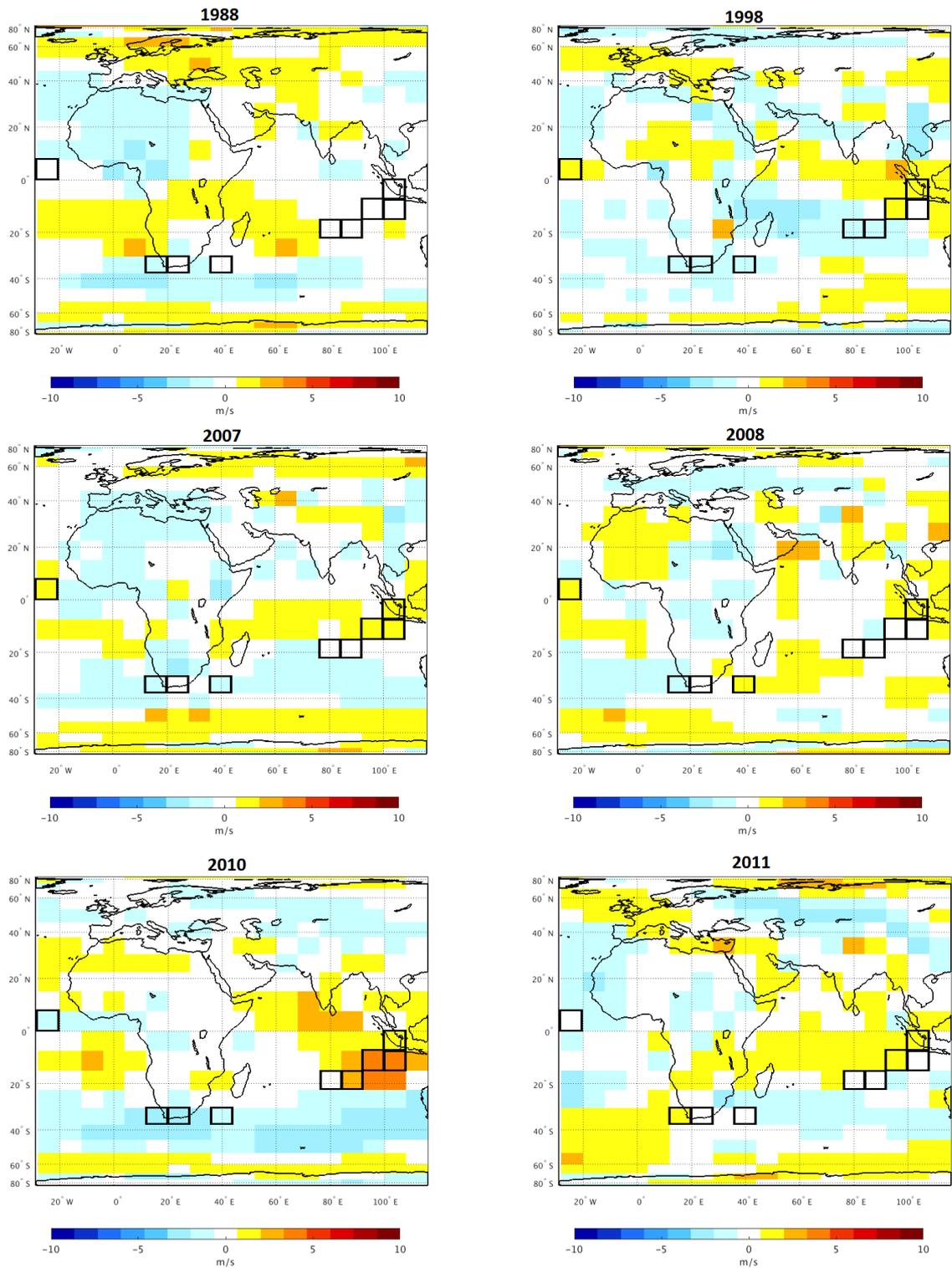


Fig. 3.9 DJF zonal wind anomalies (ms^{-1}) during La Niña events defined using the one standard deviation threshold of the instrumental SOI (1988/89, 1998/99, 2007/08, 2008/09, 2010/11 and 2011/12 relative to the 1979-2013 climatology). Predictor grid boxes used in the reconstruction are outlined in black.

Looking at the individual events (Fig. 3.8 and Fig. 3.9), the wind patterns are more varied due to the inter-event diversity of ENSO. Nevertheless, when considering the composite wind anomaly patterns (Fig. 3.7) both the El Niño and La Niña composites show zonal wind anomalies in five of the nine predictor grid boxes. For El Niño events, negative zonal wind anomalies are found in grid box 1, 8 and 9, and positive zonal wind anomalies in grid box 2 and 4. These zonal wind patterns match with expected sign of anomaly suggested by the DJF SOI and zonal wind correlation plot (Fig. 2.5) in these grid boxes, with positive correlations between these variables suggesting a negative zonal wind anomaly during negative SOI events (i.e. El Niño events) and positive zonal wind anomalies during positive SOI events (i.e. La Niña events). This is seen in grid boxes 1, 8 and 9 which have a positive correlation between the SOI and DJF zonal wind over the modern period, 1979-2013. The opposite is the case for grid boxes 2 and 4 which have negative correlation with the DJF SOI and zonal winds (Fig. 2.5). In the La Niña composites, the strong zonal wind anomalies are also as expected from (Fig. 2.5), with negative anomalies in grid box 2 and 3, and positive in grid box 7, 8 and 9. Only grid box 2, 8 and 9 have strong zonal wind anomalies in both the El Niño and La Niña composites over this period.

Overall the calculated hit rates suggest that El Niño events are more likely to be captured from reconstructions using these grid boxes compared to La Niña events (Table 3.4), however spatial zonal wind anomalies during El Niño and La Niña events show that there is variation in the spatial patterns of zonal wind anomalies during individual events (Fig. 3.8 and Fig. 3.9), with strong zonal wind anomalies within five of the nine grid boxes for both El Niño and La Niña composites (Fig. 3.7).

3.4.2 Multi-proxy comparison to instrumental DJF SOI, 1876-1977

The multi-proxy ENSO reconstructions were then compared to the instrumental DJF SOI. Pearson's correlation coefficients were calculated over the period of overlap 1876-1977. Fig. 3.10 shows a good agreement between the normalised ENSO reconstructions from these records, with reconstructions inverted on the figure to match the sign of the SOI for a more direct comparison.

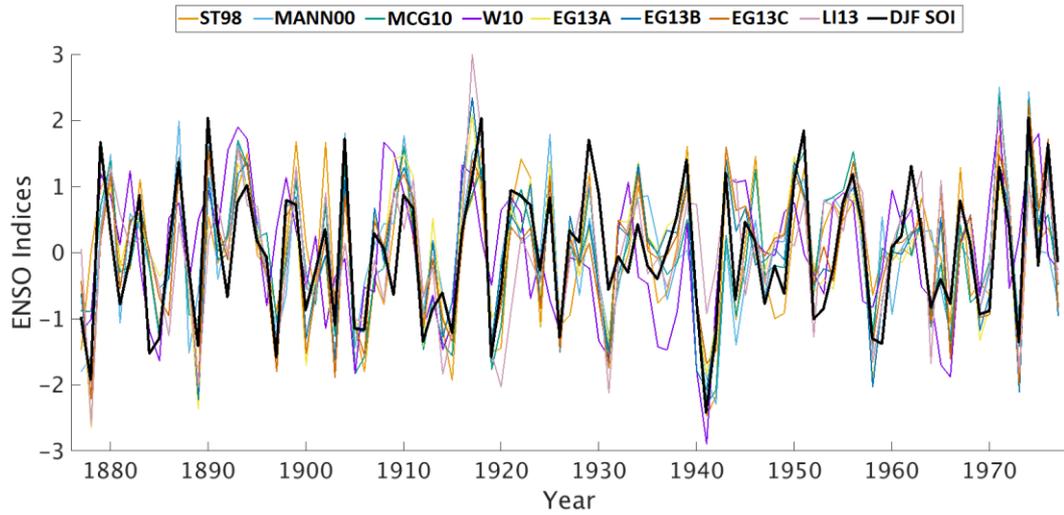


Fig. 3.10 Eight (ST98, MANN00, MCG10, W10, EG13A, EG13B, EG13C and LI13) reconstructed ENSO indices and the instrumental DJF SOI (black), 1876-1977. Non-SOI indices are inverted.

Table 3.5 Pearson’s correlation coefficients between the Instrumental SOI (DJF) and the Multi-Proxy ENSO reconstructions (Year₁ comparisons). All correlations are significant at the 99% confidence level.

Correlation coefficients (1876-1977)	
ST98	0.76
MANN00	-0.73
MCG10	-0.79
W10	-0.48
EG13A	-0.79
EG13B	-0.82
EG13C	-0.80
LI13	-0.65

A significant positive correlation is present between the instrumental DJF SOI and ST98. All others have negative correlations, as they are oceanic indices (e.g. Niño SST) which are negatively correlated with the SOI (Table 3.5). The strongest correlation ($r = -0.82$) exists between the instrumental SOI and the Niño 3.4 SST reconstruction from EG13B (Table 3.5). All correlations are statistically significant at the 99% confidence level, with the lowest correlation of $r = -0.48$ (W10). There is good agreement between the multi-proxy reconstructions and the instrumental DJF SOI, which suggest that the use of the DJF SOI as the predictand for the logbook-based reconstructions, as opposed to other ENSO indices, should not impact the correspondence between ENSO reconstructions significantly.

These multi-proxy reconstructions and the DJF SOI were then converted into discrete event chronologies. The mean and standard deviation are calculated separately for each reconstruction, over the period 1876-1977, and ENSO events from each reconstruction defined by the one standard deviation threshold. Between 1876/77 and 1977/78, 18 (17) El Niño (La Niña) events are found from the instrumental DJF SOI record. The El Niño and La Niña events from the instrumental SOI and the multi-proxy reconstructions are shown in Fig. 3.11.

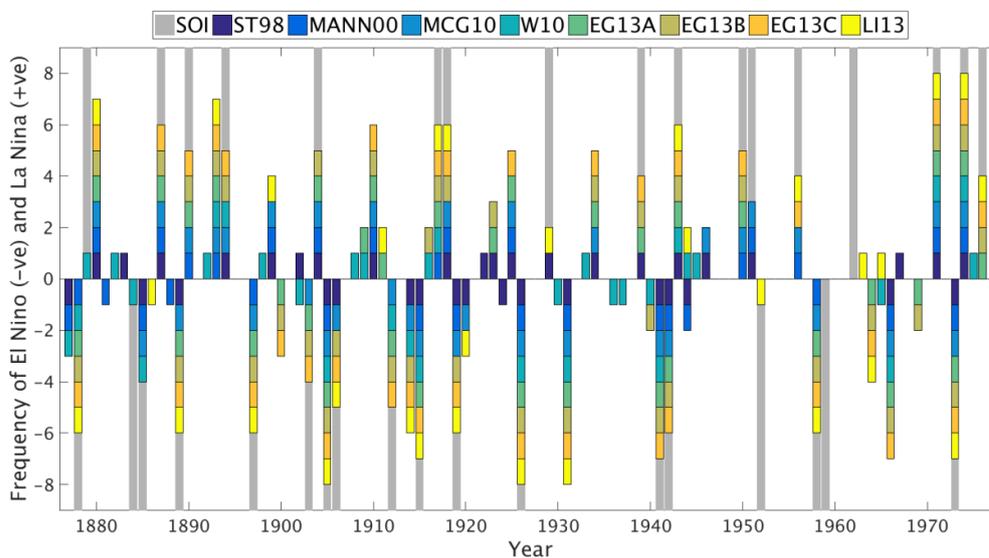


Fig. 3.11 El Niño (negative) and La Niña (positive) events, based on classification of ± 1.0 standard deviation from the mean (1876-1977) from all the multi-proxy reconstructions and the instrumental DJF SOI. Shown for Year₁ (year that contains Jan).

Table 3.6 Hit Rates (Year₁) between the instrumental DJF SOI and multi-proxy ENSO events with classification thresholds of ± 1.0 standard deviation 1876-1977 mean.

	El Niño Skill	La Niña Skill
	(%)	(%)
ST98	55	59
MANN00	61	59
MCG10	78	59
W10	44	35
EG13A	67	59
EG13B	72	71
EG13C	78	71
LI13	61	47

There is generally good agreement between the proxy reconstructions, however not all events are found within the entire set of reconstructions. Some years have as little as one reconstruction indicating an ENSO event is present, whereas in other years an event is found within all of the reconstructions. Over this period, the El Niño hit rates are higher than the La Niña hit rates for all but ST98 (Table 3.6). This is consistent with findings from previous analysis of ENSO multi-proxy reconstructions which stated that proxy reconstructions have better skill at capturing El Niño compared to La Niña (Braganza et al. 2009; McGregor et al. 2010). It is also consistent with findings from Section. 3.4.1, which found that the logbook-based reconstructions methods have better event capture skill for El Niño events compared to La Niña. EG13C has some of the highest hit rates for both El Niño and La Niña, suggesting this reconstruction best captures the events indicated by the instrumental DJF SOI. W10 performs worst, with only 35% of La Niña events captured, the only one of the reconstructions that is not above the 40% threshold suggested by Heried et al. (2013).

3.4.3 Documentary-based ENSO chronology comparison to instrumental DJF SOI, 1876-1977

The correspondence between the ENSO events captured by documentary evidence and by the instrumental DJF SOI was also assessed. OTR00 and GH08 only extend up to

1900, therefore only QUINN87 was compared over the period, 1876-1977. The hit rates when comparing to the instrumental DJF SOI are presented in Table 3.7, along with GF09's events-based reconstruction which also extends over this entire period. La Niña skill is available only for GF09, as the QUINN87 record only focuses on El Niño events. Comparisons are made for Year₀ and Year₁ to account for teleconnections that occur prior and post the ENSO-event peak (DJF).

Table 3.7 Year₀ and Year₁ (in parentheses) correspondence between instrumental DJF SOI ENSO and QUINN87 and GF09, 1876-1977.

	El Niño skill (%)	La Niña skill (%)
QUINN87	89 (61)	---
GF09	83 (61)	47 (64)

There are 18 El Niño events identified from the instrumental DJF SOI between 1876 and 1977. There are 27 El Niño events identified by QUINN87 during this same period, with 11 of these extending for more than one year. The majority of the events identified in the instrumental record are the same as those captured by QUINN87, with a hit rate of 89% (Table 3.7). This is based on comparisons in Year₀, whereas a lower hit rate is found when comparing to Year₁. This suggests that the El Niño impacts found in these documentary sources are those which precede the DJF season. GF09 has similar skill to QUINN87, with the highest hit rate, 83%, for El Niño in Year₀, and with La Niña in Year₁, 64%. The documentary records have El Niño hit rates which are higher than any of the multi-proxy reconstructions (Table 3.6), suggesting that the documentary records perform best over this period.

3.4.4 Multi-proxy comparison over logbook-based reconstruction period, 1815-1854

Comparisons were then made over the logbook-based reconstruction period, 1815-1854. Table 3.8 and Table 3.9 show the Pearson's correlation coefficients between the multi-proxy reconstructions during the two periods and there is reduced agreement between different multi-proxy ENSO reconstructions in 1815-1854 compared to 1876-1977. A number of the correlation coefficients are also no longer significant at the

95% confidence level during this earlier period, which also suggests a reduced agreement between the reconstructions during 1815-1854.

Table 3.8 Correlation between the normalised ENSO proxies 1876-1977. Bold highlights those which are significant at the 95% confidence level and non-SOI reconstructions are inverted to match the SOI.

	ST98	MANN	MGC10	W10	EG13A	EG13B	EG13C	LI13
ST98	--							
MANN00	0.77	--						
MGC10	0.89	0.89	--					
W10	0.44	0.55	0.56	--				
EG13A	0.69	0.74	0.78	0.52	--			
EG13B	0.72	0.76	0.81	0.55	0.97	--		
EG13C	0.73	0.76	0.83	0.53	0.95	0.95	--	
LI13	0.65	0.62	0.73	0.39	0.77	0.76	0.81	--

Table 3.9 Correlation between the normalised ENSO proxies 1816-1854. Bold highlights those which are significant at the 95% confidence level and non-SOI reconstructions are inverted to match the SOI.

	ST98	MANN	MGC10	W10	EG13A	EG13B	EG13C	LI13
ST98	--							
MANN00	0.49	--						
MGC10	0.74	0.77	--					
W10	0.22	0.09	0.17	--				
EG13A	0.27	0.49	0.55	0.03	--			
EG13A	0.22	0.49	0.54	0.08	0.80	--		
EG13C	0.29	0.49	0.53	0.15	0.94	0.86	--	
LI13	0.56	0.25	0.50	0.41	0.27	0.25	0.26	--

Fig. 3.12 shows the four logbook-based reconstructions and the eight multi-proxy ENSO reconstructions. The reduced agreement between the multi-proxy reconstructions during this period, 1815-1854, can be seen when compared to the agreement between the reconstructions in Fig. 3.10, 1876-1977. Pearson's correlation coefficients between the four logbook-based reconstructions and these multi-proxy indices, 1815-1854 are shown in Table 3.10. The best agreement is found with Emile-Geay et al.'s (2013a) three Niño 3.4 reconstructions, EG13A/B/C. All three of these reconstructions have significant correlations with both the CPS A and the CPS B reconstructions. For CPS B, correlations significant at the 95% confidence level are also found with ST98 and MANN00. Therefore, five of the eight proxy-based reconstructions are significantly correlated with CPS B, more than with any of the other logbook-based reconstructions. CPS A is significantly correlated with three of the multi-proxy reconstructions, EG13A/B/C. The better agreement with CPS B suggests that the adjustment in the mean carried out due to the shift in data availability

from the different logbook countries in Chapter 2 was justified. As hypothesised, CPS B is the most robust reconstruction of our four logbook-based reconstructions.

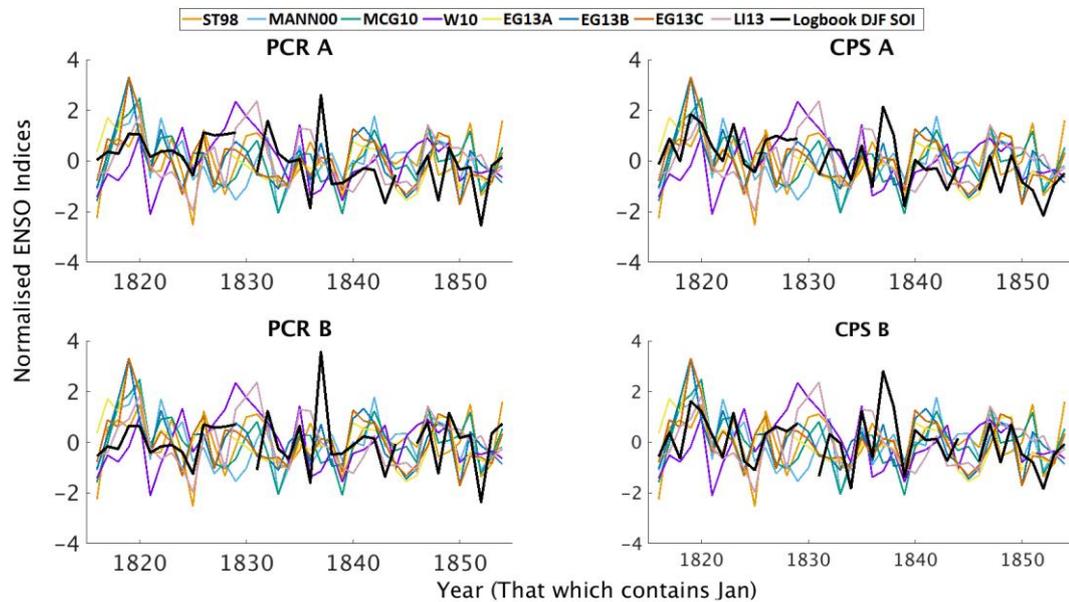


Fig. 3.12 Historical normalised ENSO indices from eight multi-proxy reconstructions (colour) and the four logbook-based SOI reconstructions (black) 1815/16 to 1853/54. The missing years in the logbook-based reconstruction are 1829/30 and 1844/45.

Table 3.10 Correlation coefficients between the logbook-based reconstructions and multi-proxy reconstructions 1815/16 – 1853/54. Those significant at the 95% confidence level are shown in bold. All indices were normalised over the period 1815-1853.

	PCR A	PCR B	CPS A	CPS B
ST98	0.20	0.29	0.25	0.32
MANN00	-0.18	-0.21	-0.25	-0.28
MCG10	-0.13	-0.12	-0.24	-0.23
W10	-0.24	-0.17	-0.23	-0.19
EG13A	-0.26	-0.16	-0.48	-0.42
EG13B	-0.27	-0.23	-0.36	-0.33
EG13C	-0.26	-0.22	-0.49	-0.48
LI13	-0.14	-0.15	-0.22	-0.26

The continuous proxy reconstructions were then used to identify El Niño and La Niña events over the period 1815 and 1853, using the one standard deviation threshold, shown in Fig. 3.13. There is good agreement between the logbook-based and the proxy-based reconstructions in 1818/19 and 1819/20, indicating a persistent La Niña event. However, there are also a number of events that are identified in only the logbook-based reconstructions, suggesting a degree of disagreement with the multi-proxy reconstructions. The percentage of the logbook El Niño and La Niña events found in the numerous multi-proxy reconstructions were analysed using correspondence rates (Table 3.11). The highest correspondence rate is 60%, between ST98 and CPS B El Niño events. CPS B has an El Niño correspondence of 40% for three reconstructions, MCG10, EG12C and LI13. Apart from this, the only other correspondence rates of 40% or more are three La Niña rates from CPS A compared to MANN00, MCG10 and EG13A. A number of correspondence rates of 0% were found, indicating inconsistency between some of the reconstructions. Overall, CPS B has the highest correspondence rates and the highest number of significant correlations during this period, further confirming that out of the four logbook-based reconstructions this has the best agreement with the multi-proxy reconstructions.

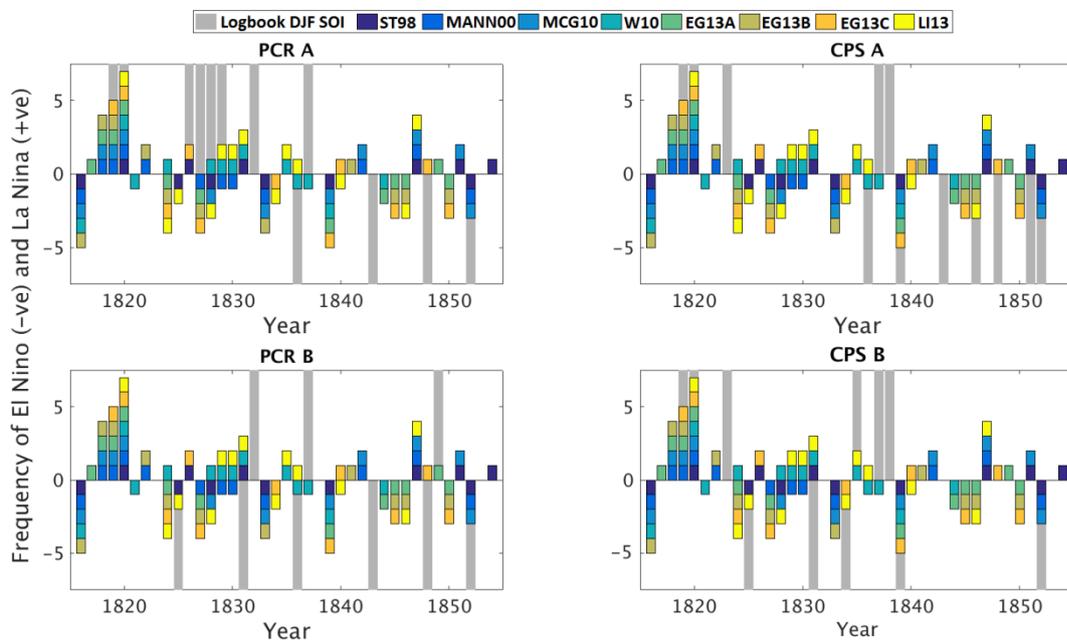


Fig. 3.13 Correspondence (Year₁) between logbook-based ENSO events and multi-proxy ENSO events with classification thresholds of ± 1.0 standard deviation 1815-1854 mean. Year₁ (that which contains Jan) shown.

Table 3.11 Correspondence rates (Year₁) for El Niño and La Niña events 1815-1854 between logbook and multi-proxy reconstructions. Bold highlights those which meet or exceed the 40% threshold.

El Niño correspondence (%)				
	PCR A	PCR B	CPS A	CPS B
ST98	25	33	29	60
MANN00	25	17	14	20
MCG10	0	17	29	40
W10	25	17	29	20
EG13A	0	0	29	20
EG13B	0	0	14	0
EG13C	0	0	14	40
LI13	0	17	14	40

La Niña correspondence (%)				
	PCR A	PCR B	CPS A	CPS B
ST98	25	0	20	17
MANN00	25	0	40	33
MCG10	25	0	40	33
W10	38	0	20	33
EG13A	25	33	40	33
EG13B	13	0	20	17
EG13C	38	0	20	33
LI13	13	0	20	33

3.4.5 Documentary chronology comparisons over the logbook-based reconstruction period, 1815-1854

The correspondence between the documentary chronologies and the logbook-based reconstructions was also assessed for the logbook-based reconstruction period, 1815/16 to 1853/54. Fig. 3.14 shows the timing of these El Niño events from the various records. A clear pattern is that there are more El Niño events in the earlier part of the period in the documentary chronologies, whereas El Niño events are more

prevalent in the later part of the logbook PCR reconstructions. There is a very limited correspondence between the logbook-based reconstructions and the documentary-based ENSO chronologies. The highest correspondence rate is 20%, found between CPS A and QUINN87 (Year₀). The cluster of El Niño events identified in the documentary-based reconstructions at the start of the record is, in fact, opposite from what is seen in both the logbook-based and multi-proxy reconstructions, with these suggesting a persistent La Niña (Fig. 3.13).

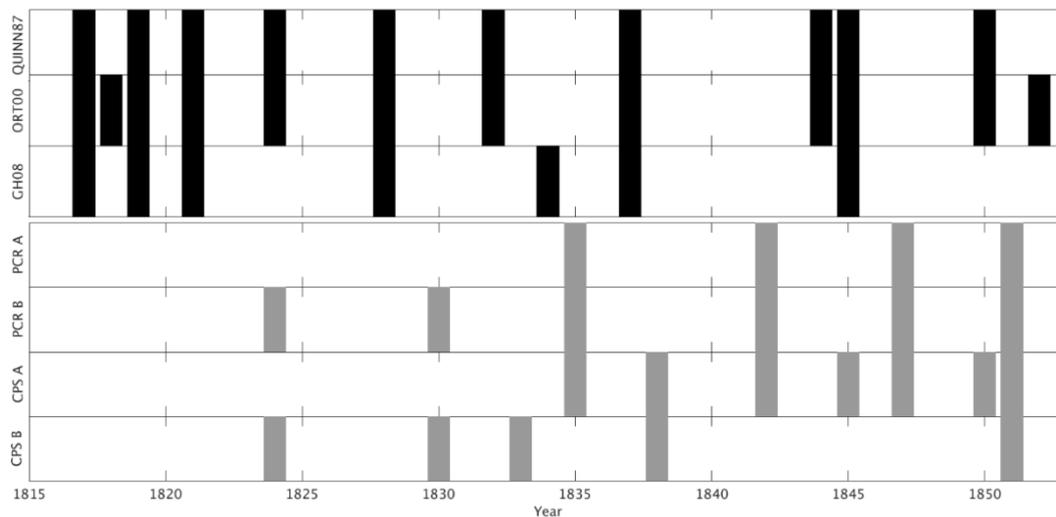


Fig. 3.14 El Niño chronologies from documentary sources, Year₀ (black) and logbook-based reconstructions (grey) 1815/16 – 1853/54. Bars indicate El Niño event.

3.4.6 Combined multi-proxy and documentary reconstruction, 1815-1854

Finally, a comparison to the reconstructions of Gergis and Fowler (2009) was carried out. GF09 identify seven moderate to extreme El Niño and seven moderate to extreme La Niña events in the period 1815 to 1854. Correspondence rates for these events are shown in Table 3.12. They are much lower than the high hit rates found for GF09 over the more modern period, 1876-1977.

Table 3.12 Correspondence rates Year₀ (Year₁ in parenthesis) for El Niño and La Niña events 1815-1854 between logbook and GF09 ENSO reconstruction

El Niño Correspondence (%)				
	PCR A	PCR B	CPS A	CPS B
GF09	0 (0)	17 (0)	14 (0)	40 (0)
La Niña Correspondence (%)				
	PCR A	PCR B	CPS A	CPS B
GF09	25 (25)	0 (33)	20 (60)	17 (50)

For CPS B, the correspondence rate of GF09 for Year₀ is 40% for El Niño and 17% for La Niña, and when comparing to Year₁ this becomes 0% for El Niño, and 50% for La Niña. Therefore, GF09 correspond better to CPS B El Niño events better in Year₀ but to CPS B La Niña events in Year₁. This corresponds with findings in Table 3.7 which suggested La Niña events are better captured by GF09 Year₁.

Overall it is clear that there are many disagreements between the numerous ENSO reconstructions covering the period of the logbook-based reconstruction, 1815 - 1854. The agreement between multi-proxy reconstructions reduces during this period compared to the more modern period analysed, 1876-1977, with a drop in the number of significant correlations and the lower correspondence rates. All correlation coefficients are significant at the 95% confidence level during the latter period, whereas a number of weak correlations in the earlier period drop below the 95% confidence threshold (Table 3.8; Table 3.9). The documentary chronologies do not show an ENSO signal similar to the logbook-based reconstructions over this period. Overall, CPS B has the best agreement with the multiple ENSO reconstructions.

The logbook-based SOI reconstructions built upon the initial work carried out by Jones and Salmon (2005), who first attempted a DJF SOI reconstruction using the observations from ships' logbooks. They found four major El Niño events during their study period, 1750-1854. Two of these are within the study period of the reconstructions presented in Chapter 2: 1833 and 1834. From the four new logbook-based reconstructions, only CPS B find 1833 to be El Niño year, and none of them

highlight 1834 as an ENSO year. Jones and Salmon (2005) correlated their reconstructed Oct-Mar SOI with two early multi-proxy studies. Low correlation coefficients were obtained, $r=0.10$ with ST98 and $r=0.08$ with MANN00. These were calculated over a longer period, 1750-1850, than the new logbook-based SOI reconstructions and therefore the strength of the correlation coefficients are not directly comparable. However, considering the new logbook-based reconstructions, correlations significant at the 90% confidence level are found between the CPS B reconstruction and both ST98 ($r= 0.32$) and MANN00 ($r= -0.28$). Therefore, there is better agreement of the new logbook-based reconstructions with the alternative ENSO reconstructions, compared to that found in earlier work from Jones and Salmon (2005). This suggests that the additional data and revised methodology has improved the ability of ENSO reconstruction using the data from ships' logbooks.

3.5 Discussion

This discussion is firstly driven by the differences in the skill of the various reconstructions in El Niño and La Niña events. Secondly, the strong La Niña signal from the logbook and multi-proxy reconstructions during 1818 to 1820 is investigated. This event occurs in the years following the 1815 Tambora eruption, which is known to have had widespread effects on global climate. Here, the influence of this non-ENSO signal on the ENSO reconstructions during these few years is explored. Finally, a number of reasons for lack of agreement between the various datasets are assessed, focusing on the differences between the documentary chronologies compared to the other reconstructions of ENSO. The results and discussion here also link to the discussion within Chapter 4, which follows a similar methodology to assess the skill of the reconstructions at representing ENSO diversity.

3.5.1 Skill of reconstructions in representing El Niño and La Niña events

Firstly, it was found that the logbook-based reconstructions (Table 3.4) and most of the multi-proxy reconstructions (Table 3.6) have higher hit rates for El Niño events compared to La Niña events. One reason for differences in the ENSO chronologies overall could be the use of the one standard deviation as a threshold to define ENSO events. The capturing of events is sensitive to this, somewhat arbitrary, but commonly used, threshold (Braganza et al. 2009). Slight differences in reconstruction method and

parameters can lead to quite different event classifications, as seen with the differences in events from the logbook-based reconstruction methodologies (Fig. 3.13). A limitation of this methodology is the non-linearity of ENSO teleconnections, with El Niño teleconnections often stronger than La Niña teleconnections, and thus more likely to provide a stronger El Niño signal in the proxy records compared to the signal from a La Niña event (Hoerling et al. 1997; Batehup et al. 2015). This in part could explain the higher hit rates for El Niño compared to La Niña found here. However, the one standard deviation threshold is most suitable for use here with the logbook-based reconstructions as it is more likely to capture real events than incorrectly interpreting noise, and is used within a number of the multi-proxy reconstructions compared here (Stahle et al. 1998; Mann et al. 2000).

Hereid et al. (2013) found that coral proxies from different regions have varying El Niño skill compared to La Niña skill depending on a number of factors including the location of the proxy records and sensitivity of the record to different signals associated with ENSO events. Therefore, the location of proxies used within multi-proxy ENSO reconstruction, and the locations of the predictor grid boxes used for the logbook-based reconstructions could bias the ENSO reconstruction from these records. This will be discussed further in Chapter 4 within the context of ENSO diversity, and the locations of the proxy records used in the multi-proxy reconstructions will be explored in relation to the spatial teleconnection patterns of variables to which they are sensitive to.

3.5.2 Influence of non-ENSO factors on reconstruction: Case study 1818-1820 La Niña

Another key finding was that compared to the poor agreement between records over much of the logbook-based reconstruction period, there is a striking degree of agreement during the years 1818-1820. The logbook-based reconstructions indicate La Niña during 1818/19 and 1819/1820, with 1818/19 the second strongest event La Niña event in CPS A and CPS B. The zonal wind patterns from the logbook data are largely consistent with expected La Niña patterns for these two years (Fig. 3.15). The logbook mean DJF zonal wind is negative in grid boxes 2-4 in both years, which is the same direction as that from modern La Niña composites (Fig. 3.7) and from the sign expected from the correlation plots of zonal wind and SOI (Fig. 2.5). The zonal wind

in all the predictor grid boxes for these two DJF seasons are not all consistent with that from the modern day La Niña composite, but as seen from individual La Niña events (Fig. 3.9), the zonal wind patterns are not entirely consistent between La Niña events.

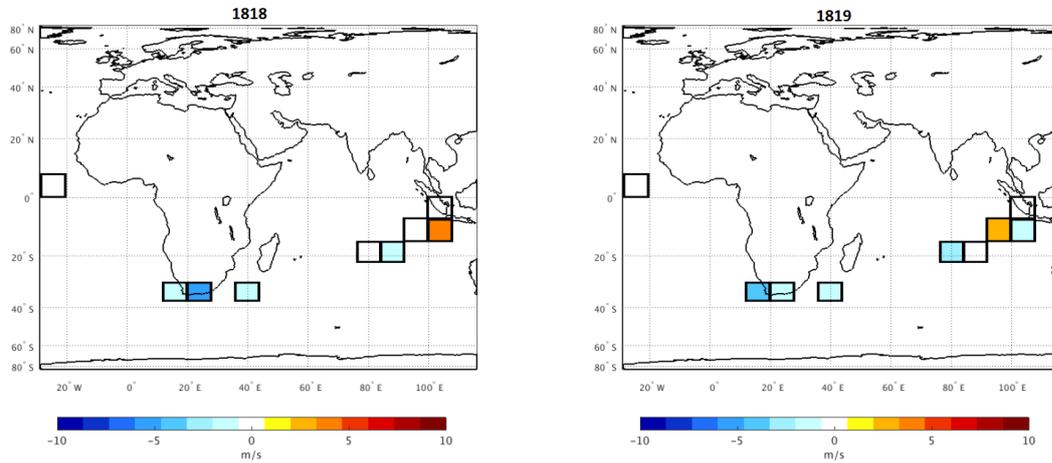


Fig. 3.15 Zonal wind anomalies (ms^{-1}) in 1818/19 and 1819/20, relative to the 1815/16 - 1833/34 DJF climatology.

Gergis and Fowler (2009) suggest a strong La Niña event in the years 1819 and 1820. The strongest La Niña values of the entire logbook-based reconstruction period are found in 1818/19 for EG13A, EG13B and EG13C. The peak of the La Niña event in the other proxy reconstructions falls in 1819/20 (ST98, MANN00, MCG10 and LI13). Only W10 does not identify this as a La Niña event. The high level of agreement between the multi-proxy and logbook-based reconstructions during these few years is interesting, as such good agreement is not seen for any period during the rest of the logbook-based reconstruction period.

A key feature of global climate during the early nineteenth century was the eruption of Tambora in 1815. The volcanic eruption on the Indonesian island of Sumbawa led to the ‘Year without a summer’ in 1816 (Chenworth 2001) and a reduction in global temperature (Raible et al. 2016). It has been suggested that large volcanic eruptions could be associated with specific ENSO behaviour. Adams et al. (2003) used the ENSO reconstruction of ST98 and MANN00 to investigate the relationship between ENSO and volcanic activity beyond the range of the instrumental record. It was found that the probability of an El Niño event doubled in the winter following a tropical

volcanic eruption. Emile-Geay et al. (2008) built upon this previous work and found that large volcanic eruptions appear to ‘load the dice’ in favour of El Niño events. Only the largest eruptions are thought to have a noticeable effect on ENSO behaviour (Emile-Geay et al. 2008). A number of the multi-proxy reconstructions indicate an El Niño event in 1816, supporting this ENSO-volcanic link (Fig. 3.13). Although the CPS B logbook-based reconstruction does not indicate an El Niño event 1815/16, the SOI index is negative, just not at a magnitude that would flag this year as an El Niño event. Recently, McGregor et al. (2010) concluded that there is also an increased likelihood of La Niña three years after a volcanic eruption. This three-year association would cover the La Niña event of 1818/19. This suggests that it is possible that the eruption of Tambora in 1815 increased the likelihood of this strong La Niña event that is indicated so clearly in the ENSO reconstructions.

Additional volcanic eruptions during the reconstruction period include 1831 and 1835 (Mann et al. 2005). The logbook-based reconstruction CPS B identified an El Niño event 1835/36 followed by a La Niña event 1836/37. However, this is picked up in very few of the multi-proxy reconstructions, suggesting an association between the volcanic eruption of 1835 and ENSO is not strong in this case. A main reason for this could be that the 1835 eruption was not as large as the Tambora eruption, and the association between ENSO and volcanic eruptions was found to be most evident during only the largest eruptions (Emile-Geay et al. 2008). It is also possible that the ENSO signal might be masked by the climatic effects of volcanic eruptions in the predictor region used in the reconstructions, or that an eruption causes an ENSO type response when an ENSO event is not actually present. Destructive interference between these different signals would result in an unclear indication of ENSO variability during, and following, a volcanic eruption. This is a possible factor masking a volcanic-ENSO response, but is not explored further here. Overall, in agreement with previous studies, it appears likely that there is an ENSO-volcano association for the Tambora eruption.

3.5.3 Disagreement with documentary chronologies

Despite the strong La Niña signal from the multi-proxy and logbook-based reconstructions, the documentary ENSO-chronologies suggest an El Niño event

during this period, spanning 1818-1819 according to QUINN87, 1817-1819 according to ORT00 and a persistent El Niño from 1816-1819 recorded by GH08. These historical records are based on conditions in South America, such as flooding and water abundance in Northern Peru (Garcia-Herrera et al. 2008). The multi-proxy reconstructions, which have a larger spatial coverage, suggest a La Niña event, in agreement with the signal apparent in the logbook-based reconstructions. This conflict between the documentary chronologies and the other ENSO reconstructions is found throughout the period of the logbook-based reconstruction.

For example, QUINN87 and ORT00 identify a ‘Very Strong’ El Niño event occurring in 1828. This was based on documentary evidence of flooding and exceptional rainfall in Northern Peru and Ecuador (Ortlieb 2000). GH08 reported 1827/28 as an El Niño year based upon additional records from Northern Peru including flood damage to bridges and roads (Garcia-Herrera et al. 2008). However, none of the logbook-based reconstruction identified 1827/8 or 1828/9 as an El Niño event, and PCR A even suggested 1827/28 and 1828/29 was a La Niña event. Of the multi-proxy reconstructions, three suggest 1827/28 as an El Niño year, and only one classifies 1828/29 as El Niño and one as La Niña (Fig. 3.13). Therefore, the signals based on historical documentary evidence, largely of rainfall and flooding from South America, appear to be inconsistent with multi-proxy reconstructions, based on records from multiple teleconnection regions. The latter are more likely to provide a more robust ENSO signal. This was also found when comparing the timing of El Niño events in the documentary chronologies and the logbook-based reconstructions (Figs. 3.14). Despite the Quinn El Niño chronology being commonly used as a reference for historical events (Gergis and Fowler 2009), it appears to have poor agreement with both the multi-proxy and logbook-based reconstructions. Results suggest that further work incorporating records from multiple regions should be carried out in order to provide a more consistent chronology of past ENSO events.

An additional reason for disagreement between the ENSO reconstructions is the differences in the scale of dating uncertainty between the proxy records and the documentary records. Documentary records, including the ships’ logbooks, are often dated to the day, and yearly summaries are based on a collection of individual weather events. In contrast, proxy records often represent seasonal or annual climate

conditions. Dating uncertainty in proxy records depends on the sources used, with tree rings having lower uncertainty than corals (Gergis and Fowler 2009). One of the criteria used for proxy selection by Emile-Geay et al. (2013), was an absolute dating uncertainty of less than ± 5 years. This is poor compared to the daily data from logbooks and historical documents, and indeed the typical duration of ENSO events. Therefore, in terms of dates the documentary and logbook data are more accurate than the multi-proxy reconstructions. This could be one of the key reasons for lack of agreement between different ENSO reconstructions, especially when comparing on an event by event basis. However, this cannot account for the differences between the logbook-based reconstructions and the documentary chronologies.

3.6 Summary

Here, a comparison of a number of reconstructions of the El Niño Southern Oscillation is presented. Particular focus is given to the DJF SOI reconstruction carried out using meteorological data from ships' logbook recently presented in Chapter 2. This logbook-based reconstruction built upon initial work by Jones and Salmon (2005), and better agreement was found between the improved logbook-based reconstruction and multi-proxy ENSO reconstructions compared to the agreement of the Jones and Salmon (2005) reconstruction with the multi-proxy reconstructions. Therefore, the use of additional logbook data, the methodological changes and the focus on strong teleconnection regions in the logbook-SOI reconstruction has shown that the use of observations from ships' logbooks to reconstruct an SOI is a viable approach to ENSO reconstructions. The logbook data therefore can provide an additional indication of the behaviour of ENSO during the pre-instrumental period. Further digitisation of wind observations from ships' logbooks can increase the temporal extent and robustness of our reconstruction significantly.

The impact of the 1815 Tambora volcanic eruption on ENSO behaviour was also investigated, within the background of previous research. McGregor et al. (2010) suggested an increased likelihood of El Niño in the year of a volcanic event. This is supported by the El Niño event suggested in 1815/16 by a number of the multi-proxy reconstructions. The likelihood of La Niña event doubled three years after a large volcanic eruption (McGregor et al. 2010). A strong La Niña was suggested in the

logbook-based reconstruction three years after the Tambora eruption and was also found in the multi-proxy reconstructions. The link between volcanic eruptions and ENSO is still contested and further work is needed to explore this fully (McGregor et al. 2010). However, here evidence is presented which suggests that this link was found for the Tambora eruption, which was one of the strongest eruption of the last millennium (Emile-Geay et al. 2008).

Overall, it is clear that there is still disagreement between reconstructions of ENSO in the early to mid-nineteenth century. The logbook-based reconstruction adds an additional, well-dated indication of ENSO behaviour during this period, being based on direct weather observations from the pre-instrumental period. It also benefits from not being subject to some of the range of uncertainties and limitations that often surround proxy records. The ENSO reconstructions compared within this chapter are now assessed in terms of their representation of ENSO diversity (Eastern Pacific and Central Pacific El Niño events) in Chapter 4.

4. ENSO diversity in historical and paleo-reconstructions: The detection of Eastern Pacific and Central Pacific El Niño events from ships' logbooks, documentary chronologies and multi-proxy reconstructions

This chapter is based upon the publication 'Barrett HG, Jones JM, Bigg GR (2017) Reconstructing El Niño Southern Oscillation using data from ships' logbooks. Part II: Comparisons with existing ENSO reconstructions and implications for reconstructing ENSO diversity. Clim Dyn DOI 10.1007/s00382-017-3797-4.' and includes text and figures from this publication.

4.1 Introduction

Following on from the ENSO reconstruction comparisons in Chapter 3, this chapter aims to assess how well the various ENSO reconstructions capture ENSO diversity. Here, ENSO diversity describes the difference in the characteristics and teleconnections of El Niño events. It is widely acknowledged that no two El Niño events are the same and that different flavours of ENSO exist (Schollaen et al. 2015). As explained in Chapter 1, El Niño events can be classified as either East Pacific (EP) or Central Pacific (CP), based upon the spatial patterns of sea surface temperature (SST) anomalies across the equatorial Pacific (Wang et al. 2012). Chapter 3 found that the majority of ENSO reconstructions have a higher skill at capturing El Niño events compared to La Niña events. This chapter assess how well EP and CP El Niño events are captured within the multiple ENSO reconstructions from ships' logbooks, presented in Chapter 2, and documentary and multi-proxy reconstructions, introduced in Chapter 3. Assessing the EP and CP El Niño event capture skill over the modern period enables an assessment of the bias within the reconstructions. The implications for the historical reconstructions, 1815-1854, are then assessed.

The past decade has seen a vast increase in research into the different types of El Niño. The 2004/05 El Niño was an igniting force behind these investigations as the typical SST anomalies associated with more classic type El Niño's were not evident, as the

centre of warming was located further towards the central and western Pacific (Ashok et al. 2007). In addition, the 2009/10 event was one of the strongest CP El Niño events in recent decades, bringing further attention to this type of event (Wang et al. 2012). CP events have become more common over recent decades (Ashok et al. 2007) and some studies suggest that they will occur more frequently in the future due to anthropogenic climate change (Yeh et al. 2009). However, there is still little known about the frequency of CP events during the pre-instrumental era. Understanding this in the past will help to distinguish between natural ENSO variability and that forced by anthropogenic climate change in the future, as well as providing an historical context for current trends. Focus is given to El Niño events rather than La Niña, as differences between La Niña events are subtler than between El Niño events, and the existence of two types of La Niña is still a controversial topic (Capotondi et al. 2015; Song et al. 2016).

There is large event-by-event diversity between El Niño events, however there are two main spatial patterns of SST anomalies (SSTA) across the Pacific which classify El Niño events into two categories: Eastern Pacific (EP) and Central Pacific (CP). EP events are defined by the typical SST pattern associated with El Niño, as explained in Fig. 1.2, with the strongest positive anomalies shifted eastwards towards the Eastern equatorial Pacific. On the other hand, CP events have a zonal tripole pattern of SSTA, with strongest positive SSTA in the Central Pacific, and negative SSTA in the western and eastern Pacific (Ashok et al. 2007). As well as the differences in SST patterns during different events, recent studies have identified some key differences in the teleconnections between EP and CP events. Central Pacific El Niño events can have unconventional teleconnections (Kug et al. 2009), which can confuse the ENSO signal within proxy records.

For example, differences exist in the spatial patterns of temperature and precipitation anomalies over the USA, variables which influence ENSO reconstructions from tree-rings (Infanti and Kirtman 2016). The relationship with the western North Pacific monsoon and ENSO is found to be weaker for EP events than CP events (Weng et al. 2011), and Australian rainfall and Indian Monsoon rainfall are more sensitive to CP events than EP events (Wang and Hendon 2007; Kumar et al. 2006). Precipitation anomalies in South America are strong during EP events but not during CP events

(Andreoli et al. 2016). Therefore, there are distinct climatic impacts of EP and CP El Niño which may leave different signals in the range of paleoclimate proxies used for ENSO reconstruction (Wang et al. 2012).

Studies into river discharge relating to El Niño events were also based around ‘typical’ El Niño until a recent study explicitly looking into the effects of the different types of El Niño (Liang et al. 2016). Some regions have an asymmetric response to the two types of El Niño, with discharge anomalies of opposite signs in the same locations. On the other hand, some regions have a symmetric response in which anomalies of the same sign are found for both types of El Niño events. By grouping El Niño events of different types together, the ENSO from regions with an asymmetric response may not be evident due to destructive interference (Hereid et al. 2013; Liang et al. 2016). Therefore, separate analysis of EP and CP El Niño events are required, and should be made a priority of future ENSO reconstructions.

The differences in teleconnections are vitally important for ENSO reconstructions which often use proxy networks from a number of remote regions (Braganza et al. 2009). As seen in Chapter 3, disagreements in the historical reconstructions of ENSO exist between the multiple reconstructions compared. Carré et al. (2014) suggested that differences between individual ENSO records from the western and eastern Pacific may be due to the spatial patterns associated with EP and CP events. Here, this idea is expanded further to include changes in spatial patterns outside of this Pacific region, using multi-proxy reconstructions which include proxy records from a number of more remote ENSO regions. A number of the ENSO reconstructions (e.g. ST98, MANN00) were carried out before the differences of EP and CP events was known. As the understanding of ENSO diversity has improved fairly recently, there are still few proxy studies which address this.

Schollaen et al. (2015) were the first to test the relationship between tree-ring proxies and the two ENSO flavours. Oxygen isotope analysis of Javanese tree-rings was used to distinguish the differences in the signal between EP and CP El Niño events. Instrumental and reanalysis data show that El Niño events are associated with drought and La Niña to be associated with excess precipitation in Java, Indonesia (Aldrain and Susanto 2003; Jourdain et al. 2013). However, Schollaen et al. (2015) show that a CP

El Niño index significantly correlated with drought, whereas EP relationship is much weaker and non-significant, because Java lies on the nodal line of influence. This difference in precipitation anomalies was used to distinguish between EP and CP events from 1900 (Schollaen et al. 2015). Karamperidou et al. (2015) noted that precipitation sensitive proxies in eastern Pacific and South America would only capture EP events, as CP El Niño don't produce significant precipitation anomalies in this region. This would therefore also be the case for precipitation based El Niño chronologies based on documentary evidence of heavy rainfall and flooding (QUINN89, ORT00 and GH08). Java and equatorial South America are just two regions in which the EP and CP teleconnections differ; Sections 4.4.3 and 4.4.4 explore this further, looking into the locations of the data used in the various ENSO reconstructions.

Here, the skill of ENSO reconstructions at capturing East Pacific and Central Pacific ENSO events is assessed. This will give an indication of any bias within the various reconstructions, and help to interpret the historical ENSO chronologies they provide. Previous studies have assessed the El Niño skill versus the La Niña skill (Gergis and Fowler 2009; Hereid et al. 2013), and this was discussed in Chapter 3. Here, the analysis of ENSO reconstructions is developed further, following similar methods to those employed in the El Niño and La Niña comparisons. This chapter investigates the spatial patterns of the ENSO teleconnections of EP and CP events along with the location of proxy records in order to investigate any potential bias towards the different types of ENSO. This is the first time that a number of recent multi-proxy ENSO reconstructions have been investigated for their EP and CP reconstruction skill.

This chapter is organised as follows: *Section 4.2* introduces the datasets used to distinguish between EP and CP events along with those used to analyse their teleconnection patterns. *Section 4.3* briefly describes the methods used. *Section 4.4* presents the results of ENSO reconstruction skill and bias assessment. It also investigates the spatial patterns of teleconnections relating to variables and regions relevant for the logbook-based, documentary and multi-proxy reconstructions. *Section 4.5* summarises this chapter and presents suggestions for areas of further research surrounding this topic. The sections within this chapter, along with those in Chapter 3, address Aim 2 as outlined in the Chapter 1

4.2. Data

4.2.1 Modern records of EP and CP El Niño events

Traditional ENSO indices such as the SOI and the Niño 3.4 index, introduced in Chapter 1, were not intended for, and do not provide a distinction between EP and CP events. There are a number of methods, indices, datasets and time periods over which classification of EP and CP events has been carried out. They use a range of variables such as SST, subsurface ocean temperature, sea surface salinity (SSS) or outgoing longwave radiation (OLR) across the equatorial Pacific (Yu et al. 2012; Capotondi et al. 2015). There currently remains a lack of consensus over which produces the best chronology (Pascolini-Campbell et al. 2015).

Three commonly used methods, and the differences between them, were discussed by Yu et al. (2012). Firstly, there is the El Niño Modoki Index (EMI), which uses the difference in area-averaged SSTA in the central, western and eastern Pacific (Ashok et al. 2007). Secondly is the EP/CP Index, which takes the DJF-averaged values of an EOF-defined CP index, compared to an EP index (Kao and Yu 2009). Finally, there is the Niño 3/Niño 4 method which involves comparison of SSTA in the Niño 4 region in the central Pacific and the Niño 3 region further towards the eastern Pacific (see Fig. 1.6 for locations) (Yeh et al. 2009). Yu et al. (2012) aimed to identify the consensus from these methods however, looking at event chronologies from 1950 onwards over a third of the El Niño years were not being classified as either all EP or all CP by the three methods, showing a degree of disagreement from the various methods.

To obtain a long record of EP and CP El Niño events, which corresponds better with the temporal coverage of the multi-proxy reconstructions, here a method is used which makes use of the Niño 3 and Niño 4 indices which can be obtained from NOAA, which extend back to 1870 (NOAA, 2017a; NOAA 2017b), and were calculated from the Hadley Centre Sea Ice and Sea Surface Temperature dataset (HadISST1). Defined by Kug et al. (2009), and similar to a method used within Yu et al. (2012), this method referred to here as the NINO method, uses the Niño 3 (5°N-5°S 150°W-120°W) and the Niño 4 (5°N-5°S 160°E-150°W) indices (Kug et al. 2009; Yeh et al. 2009). EP (CP) El Niño events are defined by the normalised Niño 3 (Niño 4) SSTA being greater than 1K and being greater than the Niño 4 (Niño 3) SSTA (Ham and Kug 2012). This

index is chosen as it provides a record back to 1870 which can be compared to the multi-proxy reconstructions and is one of the commonly used methods (Yu et al. 2012; Capotondi et al. 2015). In addition, to provide an alternative method for comparison a record of EP and CP El Niño events from a recently published proxy-based record by Schollaen et al (2015) is used here. The Schollaen et al. (2015) record of EP, CP and La Niña events, available from 1900 onwards, is based on Indonesian tree-rings which are influenced by distinctly different teleconnection patterns between ENSO flavours (Schollaen et al. 2015).

4.2.2 Existing ENSO reconstructions

The ENSO reconstructions compared in Chapter 3 are used again here to assess ENSO diversity. These include the logbook-based DJF SOI reconstructions, the documentary ENSO-chronologies (QUINN87, ORT00, GH08, and GF09) and the multi-proxy reconstructions (ST98, MANN00, MCG10, W10, EGA13A/B/C and LI13). The instrumental DJF SOI is also used over the modern period in place of the logbook DJF SOI reconstruction. The locations and weightings of the multi-proxy networks used in the various reconstructions were presented in Chapter 3 (Fig. 3.1-3.4), and are taken from the various publications.

A number of the ENSO reconstructions analysed in this study were produced before the appreciation of the differences of the types of ENSO events, so did not explicitly consider the effect of this on their reconstruction. Therefore, it is possible that the ENSO reconstructions could portray slightly different flavours of past ENSO and be biased towards reconstructing one type over another (Wilson et al. 2010; Karamperidou et al. 2015). The Niño 3.4 reconstruction by Li et al. (2013) was stated to be a reconstruction of canonical (EP) El Niño. Thus the Niño 3.4 is used here to identify EP El Niño events. However, other reconstructions using this index do not explicitly state that their reconstruction is solely for this one type of El Niño event (e.g. Emile-Geay et al. 2013a). Inconsistencies between different ENSO reconstructions, as seen in Chapter 3, mean that the true nature of ENSO variability of the past is still unclear (Emile-Geay et al. 2013). The effect of EP and CP El Niño events and their teleconnections are explored as a potential source of the poor consistency amongst

previous ENSO reconstructions and in comparison to the logbook-based reconstruction.

4.2.3 Teleconnection data

In order to explore the teleconnections associated with EP and CP El Niño events a number of datasets are used to create spatial composite anomaly plots. Firstly, the DJF zonal wind from ERA-Interim reanalysis is used, as in Chapters 2 and 3, 1979/80 – 2013/14 (Dee et al. 2011). This is used to assess the strength and direction of zonal wind patterns during EP and CP El Niño events within the logbook predictor grid boxes. The wind data from ships' logbooks is also used in order to analyse the wind patterns in individual ENSO years, defined from this record, during the historical period 1815-1854.

The Global Precipitation Climatology Project (GPCP; Adler et al. 2003) is a global gridded precipitation dataset used here to assess precipitation anomalies. The data was downloaded from the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, at <http://www.esrl.noaa.gov/psd/>. This monthly data set provides average precipitation rate (mm/day) from which, EP and CP Composites are plotted using events defined from the NINO method (1979-2013). Surface temperatures from ERA-Interim are used to create composite anomalies for temperature over the North American tree-ring region. The SST anomaly composites for EP and CP El Niño years were extracted from HadISST, the Hadley Centre Sea Ice and Sea Surface Temperature dataset (Rayner et al. 2013).

4.3 Methods

Firstly, the logbook-based reconstructions were assessed for their ability to capture ENSO diversity. The two reconstructions methods, PCR and CPS, were investigated over the fitting period, 1979-2013. Their EP and CP skill was calculated, based on hit rates, by comparison to the events classified by the NINO method (Kug et al. 2012). The spatial pattern of wind anomalies from EP and CP El Niño events were also analysed to further investigate any potential bias to a specific flavour of ENSO in the nine predictor grid boxes. Using the spatial patterns of zonal wind anomalies from ERA-Interim in the predictor grid boxes during El Niño events between 1979-2013,

an experimental attempt to categorise the historical logbook events into EP and CP El Niño was carried out. A simple agreement methodology was used in which the spatial wind patterns from modern composites were compared to those in historical individual events. The percentage of the predictor grid boxes in the logbook-defined events with wind anomalies of the same sign as seen in the modern EP and CP composites was calculated. For individual events within the period of the logbook-based reconstruction, if a higher percentage of EP (CP) grid boxes have the modern zonal wind anomaly pattern than CP (EP), then the event is labelled as EP (CP). This provides an indication of the frequency of CP and EP events during the reconstruction period. This is carried out for El Niño events defined from the CPS B logbook-based reconstruction and then the PCR B reconstruction in order to assess the differences in events identified from the two methodologies.

Following this, the documentary records are assessed, looking at both hit rates and at the differences in teleconnections between the types of El Niño event in the locations of these records. Focus is given to precipitation in coastal South America, centred around the Peruvian coast due to the abundance of documentary records in this region (QUINN87, ORT00, GH08). Finally, the multi-proxy reconstructions are investigated focusing on hit rates for EP and CP El Niño events. The teleconnections associated with EP and CP El Niño events are analysed with respect to the locations of the proxies and the variables to which they are sensitive. Focus is given to temperature and precipitation teleconnections over the North American tree ring region, and in the tropical oceans, which contain the majority of the coral records used in multi-proxy ENSO reconstructions. Looking into hydroclimate composites from the two types of events, and based upon previous work of Hereid et al. (2013) and Schollaen et al. (2015), discussion will focus on the influence of the difference in teleconnections of the two types of events on the reconstruction methods and outputs.

4.4 Results

4.4.1 ENSO diversity in the logbook-based reconstruction methods and predictor grid boxes, 1979-2013

Firstly, hit rates were calculated using the same methodology as explained in Chapter 3. The EP and CP El Niño events from the NINO method using HadISST-derived

Niño indices were compared to those identified from the PCR and CPS fitting reconstructions (Table 4.1). This was carried out using the full fitting data, and the sub-sampled data which aims to more directly replicate the logbook data availability within ERA-Interim reanalysis as explained in Chapter 3 (Section 3.2.1).

Table 4.1 Hit rates for EP and CP El Niño events from PCR, CPS and the instrumental DJF SOI, compared to the NINO method, 1979-2013. PCR and CPS full fitting and the mean of five sub-sampled iterations are presented.

	EP Skill %	CP Skill %
Instrumental DJF SOI	100	33
PCR fitting	100	66
PCR sub-sampled (5 it.)	66	66
CPS fitting	100	33
CPS sub-sampled (5 it.)	100	33

The NINO method identifies three EP events (1982/83, 1991/1992 and 1997/98) and three CP events (1994/95, 2002/03 and 2009/10) since 1979. All three EP events were identified as El Niño events in the instrumental DJF SOI, whereas only one of the three CP events were found in this record, a hit rate of 33% (Table 4.1). This immediately suggests that the instrumental DJF SOI classification may identify EP events better than CP events. The PCR reconstruction identified 100% of the EP events, and two of the three (66%) CP events, while the CPS reconstruction identified all three (100%) EP events, but only one of the three (33%) CP events (Table 4.1). This suggests that the two methods used in the logbook-based reconstruction also capture EP events more successfully than CP events, with a stronger bias towards EP events in the CPS-based reconstruction compared to PCR-based. These conclusions are also drawn from the sub-sampled reconstructions. However, the hit rates for the PCR five iterations is 66% for both CP and EP events (Table 4.1).

To further investigate this potential bias, the spatial wind patterns of individual ENSO events were analysed, specifically within the nine predictor grid boxes, highlighted in Fig. 4.1 (see Appendix Table A.1 and Fig. A.2 for labels and exact locations). Composite plots of the DJF mean seasonal zonal wind anomalies for EP and CP events

were constructed using ERA-Interim Reanalysis data based on the 1979-2013 climatology (Fig. 4.1) as well as anomaly plots for individual EP and CP El Niño events (Fig. 4.2 and Fig. 4.3 respectively). Stronger anomalies are seen in a higher number of the predictor grid boxes during EP events compared to CP events. Positive wind anomalies are found during EP events in grid boxes 2, 3, 4 (South African grid boxes), 5 and 6 (Indian Ocean) and negative anomalies in grid boxes 1, 8 and 9. In CP events, the only visible signal in the predictor grid boxes is a negative one in grid boxes 3, 4 and 9. It should be noted that the sign of the anomaly in grid boxes 3 and 4 is the opposite in EP compared to CP events. Therefore, the differences in the ENSO-wind teleconnection in this region around South Africa could be a useful indicator of ENSO flavours within future studies using wind as a key variable to inform ENSO reconstructions.

The El Niño composite presented in Chapter 3 (Fig. 3.7) was calculated based upon events defined from the one standard deviation of the DJF SOI threshold, and thus is informed by a different collection of events compared to the EP and CP composites presented here (Fig. 4.1). The zonal wind anomalies within the predictor grid boxes in Fig. 3.7 has higher similarity to that in the EP composites compared to the CP composites. This suggests that EP El Niño events have stronger similarities to more traditionally defined El Niño events (i.e. using the SOI) than CP El Niño events.

Looking at the individual events (Fig. 4.2; Fig. 4.3), the wind patterns are more varied due to the inter-event diversity of ENSO. Nevertheless, when considering the composite wind anomaly patterns (Fig. 4.1) the higher number of predictor grid boxes with stronger wind anomalies during EP compared to CP events, and the increased skill of the two methodologies at capturing EP events over CP events suggests that the logbook-based reconstructions would favour reconstruction of EP over CP events. This has implications for the ability of the historical reconstruction to capture CP events and is considered in section 4.4.2 when interpreting the historical logbook-based reconstructions. The zonal wind patterns associated with EP and CP El Niño events globally and within the predictor grid boxes are also investigated further in Chapter 7 using a range of historical reanalysis and historical climate model simulations.

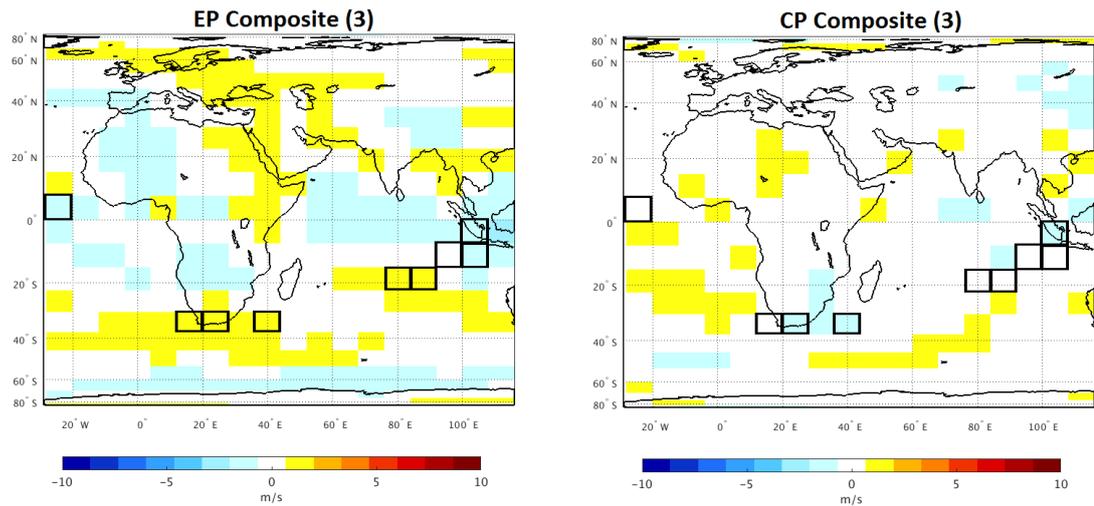


Fig. 4.1 DJF zonal wind composite anomalies (ms^{-1}) during EP El Niño events (1982/83, 1991/92 and 1997/98) and CP El Niño events (1994/95, 2002/03 and 2009/10), relative to the 1979-2013 climatology. Zonal wind data are taken from the ERA-Interim reanalysis. Predictor grid boxes used in the reconstruction are outlined in black, numbered 1-9 from west to east. Brackets indicate the number of events used to calculate each composite.

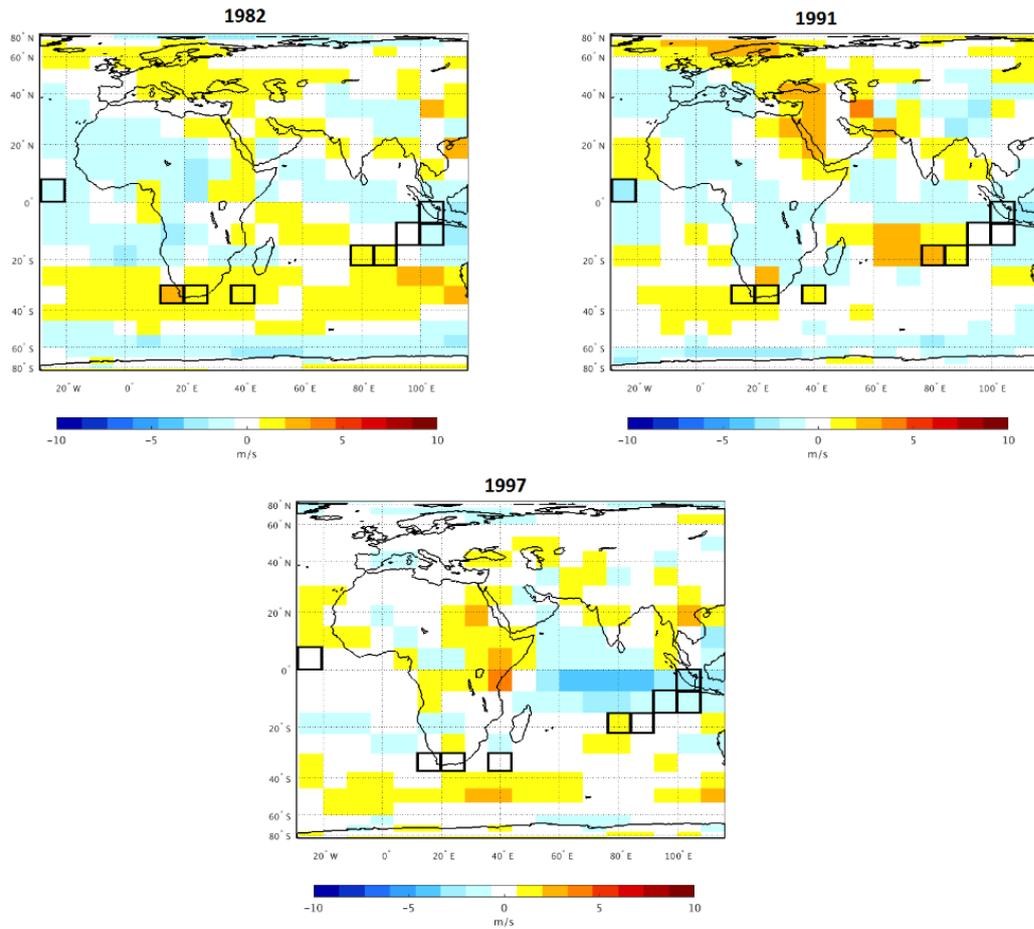


Fig. 4.2 DJF zonal wind anomalies (ms^{-1}) during EP El Niño events 1982/83, 1991/92 and 1997/98 relative to the 1979-2013 climatology. Predictor grid boxes used in the reconstruction are outlined in black.

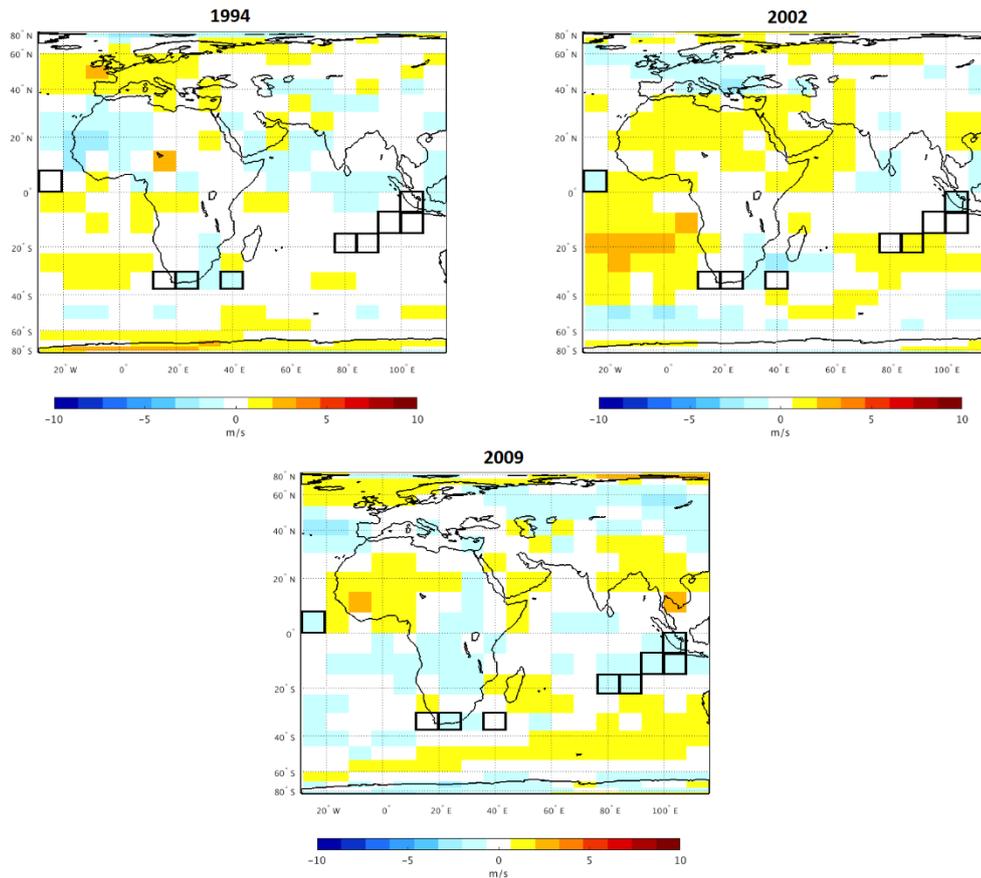


Fig. 4.3 DJF zonal wind anomalies (ms^{-1}) during CP El Niño events 1994/95, 2002/03 and 2009/10, relative to the 1979-2013 climatology. Predictor grid boxes used in the reconstruction are outlined in black.

4.4.2 ENSO diversity within the logbook-based reconstructions, 1815-1854

The zonal wind anomalies in the predictor grid boxes were stronger in a higher number of grid boxes for EP events than CP events (Fig. 4.1). Using these different wind patterns, speculative analysis of El Niño events in the logbook-reconstruction was carried out, in an attempt to categorise the historical events into EP and CP El Niño. PCR B and CPS B reconstructions are analysed in this section. Chapter 3 found CPS B to have the best agreement with the multi-proxy ENSO reconstructions (Table 3.11). PCR B is also analysed as the PCR method was found to have a lower EP-bias than the CPS method (Table 4.1).

4.4.2.1 CPS logbook-based reconstruction, 1815-1854

An attempt to categorise the El Niño events identified by the logbook-based reconstructions was then carried out. In the ERA-Interim composites (Fig. 4.1), eight of the nine grid boxes show negative (grid box 1, 8, 9) or positive (grid box 2-6) wind anomalies in EP events, whereas only three (3, 4, 9) display visible negative zonal wind anomalies in CP events. Fig. 4.4 shows the DJF zonal wind anomalies from logbook data for the five years defined as El Niño events in the CPS B reconstruction, 1824/25, 1830/31, 1833/34, 1838/39 and 1851/52, as identified in Fig. 3.14. Compared to the El Niño zonal wind anomalies in the fitting period, there is a less consistent pattern in the direction and strength of the anomalies in the logbook-derived winds. An attempt to classify these events as EP and CP is very much an experimental one, however a simple agreement methodology was attempted in which the spatial wind patterns from modern composites were compared to those in historical individual events.

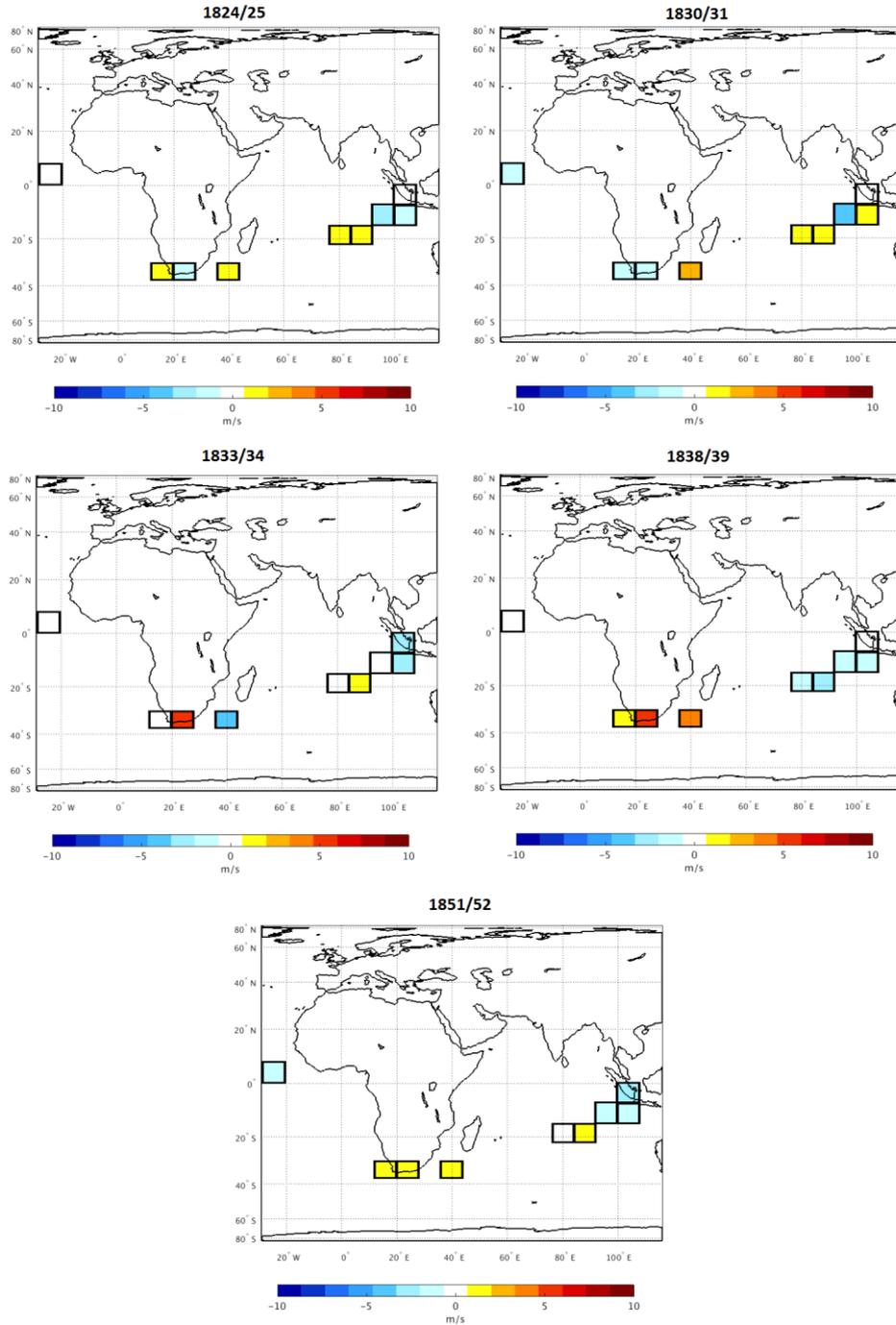


Fig. 4.4 DJF Zonal wind anomalies (ms^{-1}) from logbook observations for CPS B El Niño years with 1824/25 and 1830/31 anomalies relative to the mean of 1815-1833 and 1833/34, 1838/39 and 1851/52 relative to the mean of 1834-1854.

Table 4.2 Classification of CPS B El Niño events into EP and CP El Niño.

	% EP Grid boxes	% CP Grid boxes	Type of Event	Multi-proxies indicating El Niño
1824/25	63	33	EP	ST98, LI13
1830/31	50	33	EP	None
1833/34	38	66	CP	EG13C, LI13
1838/39	50	0	EP	ST98, MCG10, W10, EG13A, EG13C
1851/52	88	33	EP	ST98, MANN00, MCG10

Four of five the historical events from CPS B are suggested to be EP El Niño events and only one to be a CP event (Table 4.2). As discussed in Chapter 3, some of the multi-proxy reconstructions also identify El Niño events within four of these five years, however there is far from universal agreement upon these within the records. The event chronology from this reconstruction suggests a lower frequency of CP El Niño events in this period compared to the reported rate of increase in CP El Niño in recent decades (Lee and McPhaden 2010).

4.4.2.2 PCR logbook-based reconstruction, 1815-1854

The PCR method was found to have a lower EP-bias than the CPS method over the modern period, 1979-2013 (Table 4.1), with both methods capturing all the EP events in this period, but PCR capturing two of three CP events while CPS only captured one. Therefore, it is suggested that the CPS method has a stronger bias towards EP than the PCR method. This would suggest that the CPS method would identify less CP events. During the logbook period, PCR B reconstruction found six El Niño events. Three of these events (1824/25, 1830/31, 1851/52) are also identified by CPS B reconstruction. These are events are all classified as EP events (Table 4.2). The three remaining events from PCR B which are not found in CPS B, are in the years 1835/36, 1842/43 and 1847/48 and are all found in both PCR A and CPS A, as well as PCR B. The zonal wind anomalies for these additional events are shown in Fig. 4.5, and again are compared to the modern patterns from EP and CP El Niño composite analysis (Fig. 4.1).

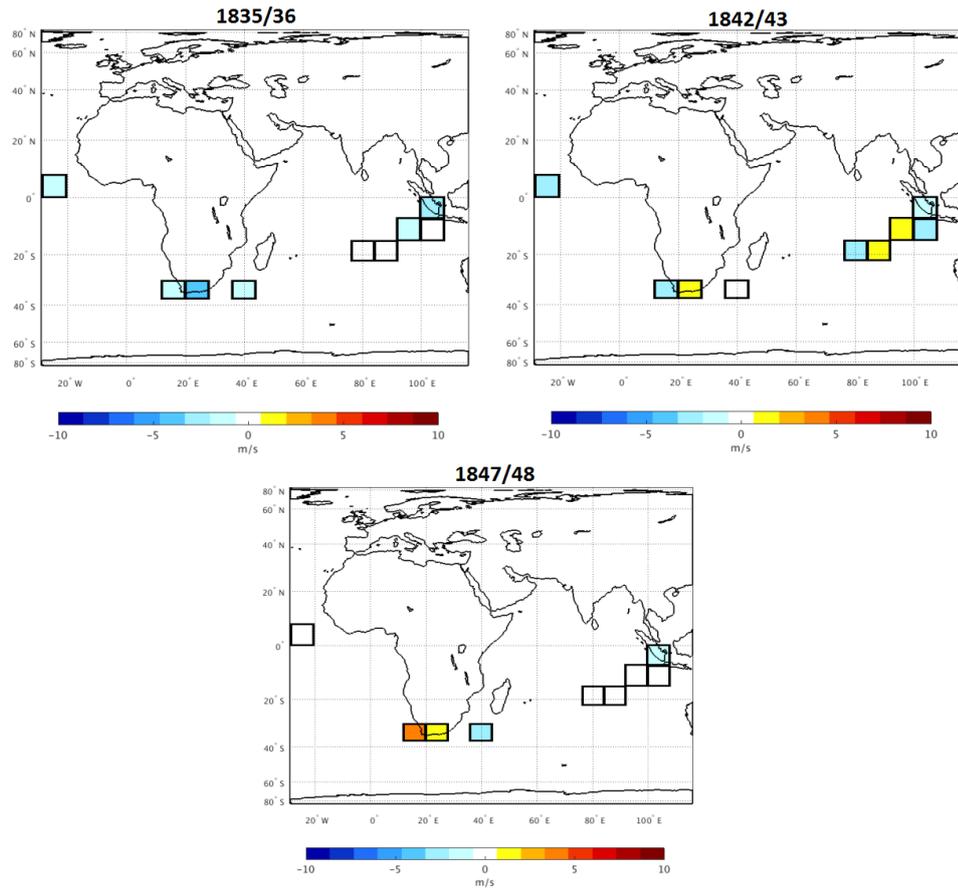


Fig. 4.5 DJF Zonal wind anomalies from logbook observations for PCR B El Niño years, 1835/36, 1842/43 and 1847/48 relative to the 1834-1854 mean. See Fig. 4.3 for the additional three PCR B events, which are also found in CPS B (1824/25, 1830/31, 1851/52).

Table 4.3 Classification of PCR B El Niño events into EP and CP El Niño, Bold indicates events unique to PCR B, thus not already presented in Table 4.2 for CPS B.

	% EP Grid boxes	% CP Grid boxes	Type of Event	Multi-proxies indicating El Niño
1824/25	63	33	EP	ST98, LI13
1830/31	50	33	EP	None
1835/36	25	100	CP	W10
1842/43	63	33	EP	None
1847/48	38	66	CP	None
1851/52	88	33	EP	ST98, MANN00, MCG10

Of the three additional El Niño events from PCR B, two of these are suggested CP El Niño years. Only W10 from the multi-proxy reconstructions indicate an El Niño event in 1835/36, whereas none of the multi-proxies indicate an El Niño during the remaining two years (1842/43 and 1847/48). Overall, four out of the six El Niño events identified by PCR B are classified as EP events, suggesting more EP than CP El Niño events during this historical period (1815-1854), with a CP-EP ratio of 0.5. This is also supported by the higher frequency of EP than CP events in CPS B-defined events, with a CP-EP ratio of 0.25. Further work is needed in order to fully understand the reasons for the different event-capture skill of the logbook-based reconstructions. This is discussed further in Section 4.5.

4.4.3 Documentary records: Event chronology comparison to EP and CP El Niño events, 1876-1977

ENSO diversity within the documentary event chronologies was then investigated. Hit rates were calculated for the El Niño chronologies of QUINN87 and GF09 during the period of overlap with the multi-proxies and instrumental record, 1876-1977, as used in Chapter 3. Table 4.4 shows the EP and CP El Niño hit rates for QUINN87 and GF09 for both Year₀ and Year₁. Both records have higher skill at capturing EP events compared to CP events in Year₀, with QUINN87 having an EP skill of 93% and a CP

skill of 56%, while GF09 has a 100% EP skill and a 77% CP skill. A higher EP skill compared to CP skill is also found in QUINN87 for Year₁ of events (Table 4.4).

Table 4.4 EP and CP skill of documentary records compared to target record of events from the NINO method, 1876-1977.

	EP skill (%)	CP skill (%)
QUINN87	93 (86)	56 (33)
GF09	100 (71)	77 (77)

ENSO reconstructions from documentary records, discussed Chapter 3, are largely from records located in South America and are based upon observations of heavy precipitation and flooding. The two records were found to have poor agreement with the logbook-based reconstruction in Chapter 3 (Fig. 3.14, Table 3.12). Questions were raised over the reliability of the documentary sources including the ability of individual events to represent a season and the use of sources from different ENSO regions. The documentary-chronologies are based on the link between El Niño events and the increased incidence of heavy rainfall and flooding in equatorial coastal South America, assuming the conventional ENSO teleconnection of precipitation surplus in El Niño events and precipitation deficit during La Niña events (Quinn et al. 1987).

However, previous studies of precipitation anomalies in South America suggest that the signal is strong during EP events but not during CP events (Andreoli et al. 2016). CP El Niño events do not generate rainfall anomalies on the Peruvian Coast or off the western coast of South America (Carré et al. 2014; Karamperidou et al. 2015). This suggests that the event-chronologies would fail to identify CP events. These precipitation anomalies can also be seen in Fig. 4.6, with stronger precipitation anomalies clearly seen for EP events compared to CP in this northern Peru coastal region, which covers the location of the majority of the documentary records used by QUINN87, ORT00 and GH08. This supports findings from previous studies looking at differences in precipitation patterns during EP and CP El Niño events (Karamperidou et al. 2015).

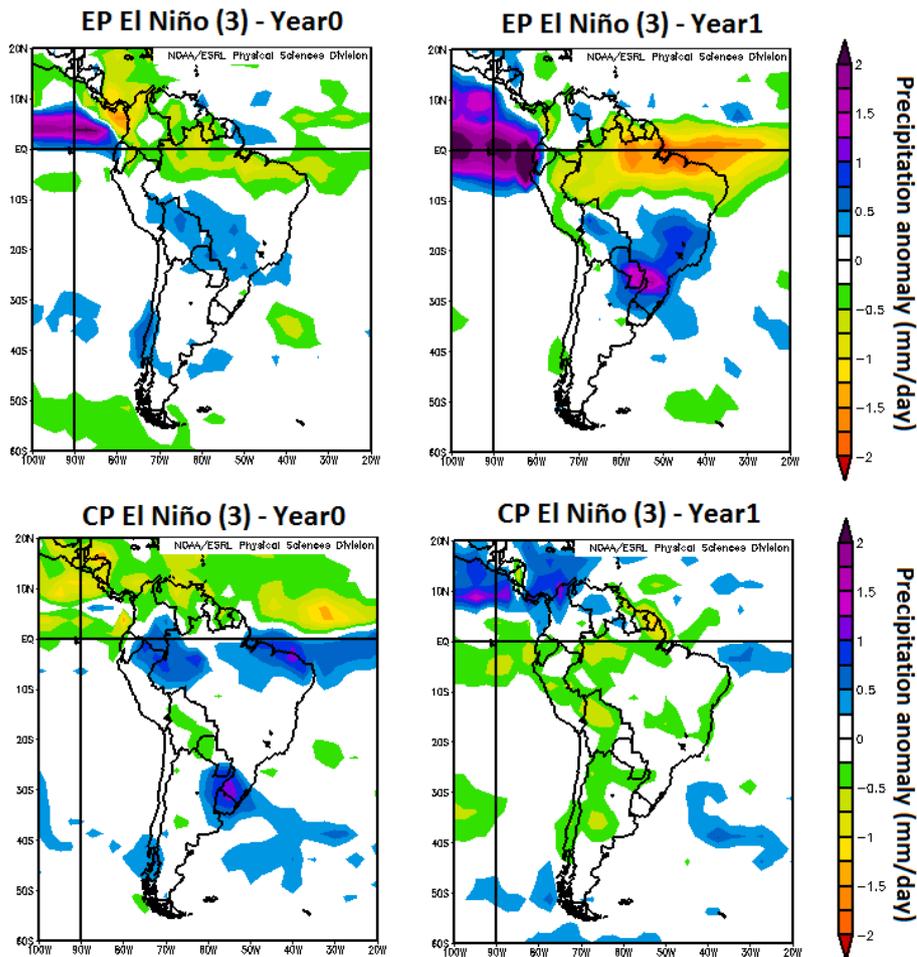


Fig. 4.6 Jan-Dec precipitation Composite anomaly plots for EP El Niño (1982/83, 1991/92 and 1997/98) and CP El Niño (1994/95, 2002/03 and 2009/2010) events during Year₀ and Year₁ of the events, using data from GPCP and made using NOAA composite plotting tool. Precipitation (mm/day) composite anomaly 1981-2010 climatology (NOAA 2017g).

Overall, this suggests that disagreement between the El Niño chronologies of the documentary record and ENSO reconstructions may in part be a result of the differences in precipitation patterns associated with the two types of El Niño events. Typical documentary-based studies that link observations of heavy rain and flooding with El Niño events, are based on the established teleconnections during classical, EP El Niño events, whereas, the precipitation patterns during CP events do not show anomalous rainfall to the extent seen in EP El Niño events, and are thus likely to be events missed within the documentary-based ENSO reconstructions.

4.4.4 Multi-proxy ENSO reconstructions

4.4.4.1 Multi-proxy comparison to EP and CP El Niño, 1876-1977

The eight multi-proxy reconstructions were then assessed for their skill in reconstructing ENSO flavours. Using the NINO method with HadISST Niño 3 and Niño 4 indices, 14 EP events and 9 CP events are identified across 1876-1977. Table 4.5 shows the EP and CP skill of the multi-proxy reconstructions when compared to this record. EG13B has the highest EP (92%) and joint highest CP skill score of the multi-proxies. All of the multi-proxy reconstructions have better skill at capturing EP events than CP, except EG13A which has a CP skill score 3% higher than its EP skill score. With this exception, these results show the same bias towards EP events as was found for the instrumental DJF SOI, the logbook-based reconstruction methodologies (Table 4.1) and the documentary ENSO chronologies (Table 4.4).

Table 4.5 EP and CP skill of multi-proxy reconstructions compared to target record of events from the NINO method, 1876-1977.

	EP Skill (%)	CP Skill (%)	Difference between EP and CP skill (%)
ST98	57	33	24
MANN00	50	44	6
MCG10	71	55	16
W10	50	33	17
EG13A	64	67	-3
EG13B	92	67	25
EG13C	85	56	29
LI13	64	22	20

4.4.4.2 Multi-proxy comparison to Schollaen et al. (2105) defined EP and CP events, 1900-1977

In addition to comparison of the NINO method event chronology, the multi-proxy reconstructions were compared to the EP and CP events identified by Schollaen et al. (2015). This reconstruction identifies 11 EP events and 12 CP events between 1900

and 1977, the period of overlap with the multi-proxy reconstructions. All of the multi-proxy reconstructions have better skill at capturing EP events than CP when comparing to this reconstruction (Table 4.6). This also suggests that the multi-proxy reconstructions are more sensitive to conventional ENSO events, rather than CP events, and supports the results found when comparing to the NINO method defined EP and CP events, in Table 4.5.

Table 4.6 EP and CP skill of multi-proxy reconstructions compared to target record of events from Schollaen et al. (2015), 1900-1977.

	EP Skill (%)	CP Skill (%)	Difference between EP and CP skill (%)
ST98	82	25	57
MANN00	55	25	30
MCG10	82	42	40
EG13A	55	42	13
EG13B	64	58	6
EG13C	73	42	31
LI13	55	25	30

The largest difference between the EP and CP skill score of multi-proxy reconstruction is found for ST98, with an EP score of 82% but a CP score of only 25% (Table 4.6). The smallest difference between skill scores is found for EG13 reconstructions, with EG13B having only 6% difference between an EP skill score of 64% and a CP skill score of 58%. This would suggest that, using this method, ST98 has the biggest EP bias of the multi-proxy reconstructions, and EG13B has the smallest EP-bias. This differs from the results found in Section 4.4.4.1 which finds EG13C having the largest EP bias, but EG13A having a slight CP bias, when comparing to the NINO method chronology. However, when comparing to two independent records of EP and CP El Niño (NINO method and Schollaen et al. 2015) over two overlapping periods (1876-1977), the same general conclusion that the multi-proxy reconstructions have a bias towards EP events compared to CP events can be drawn.

There is no clear reason for the higher skill of the multi-proxy reconstructions at capturing EP rather than CP events. One factor that might influence their skill in capturing ENSO diversity are the locations and the weightings of the proxy records used. The multi-proxy reconstructions compared in this study contain large proxy data networks which include ENSO signals from a number of teleconnection regions. Regions with higher weightings which have a consistent teleconnection in both EP and CP events will be most likely to produce an ENSO reconstruction that can capture both types of events. MCG10 noted that their reconstruction might be influenced by an over-representation of North American tree rings due to the common use of these tree rings within the multiple reconstructions they used, as well as their higher weightings. This leads to a potential source of bias within their ENSO reconstruction, as the temperature and precipitation teleconnections during EP and CP events differ in this highly weighted North American teleconnection region (McGregor et al. 2010; Infanti and Kirtman 2016). This is investigated further in section 4.4.5.

Emile-Geay et al. (2013b) showed that during the period coincident with the logbook-based reconstruction lower weightings were given to the tree-ring networks of North America, with highest weightings given to corals in the western Pacific, eastern Indian Ocean, Red Sea and a South American ice core. Therefore, EG13 reconstructions are more influenced by the Indo-Pacific region, a less remote teleconnection region. LI13 used a tree ring network that was divided into seven regions. The highest weightings were given to those networks from the Maritime Continent (EOF loading = -0.54), South American Altiplano (0.44) and Central Asia (0.45), see Fig. 3.4. Therefore, the LI13 reconstruction had high weightings on either side of the Pacific. These multi-proxy reconstructions, despite using different networks, methods and weightings, were all found to capture EP events better than CP events. LI13 addressed ENSO diversity in their study stating that they carry out a reconstruction of canonical (EP) El Niño. This in part explains the reason for the difference between EP and CP skill in this record. A number of these ENSO reconstructions were produced before the appreciation of the different of the types of ENSO events so did not explicitly consider the effect of this on their reconstruction. Therefore, it is possible that the ENSO reconstructions could portray slightly different flavours of past ENSO, and be biased towards reconstructing one type over another (Karamperidou et al. 2015). However,

as for the logbook-based reconstruction, it is suggested that the higher EP skill of the multi-proxy reconstructions could be a result of the index they are calibrated to, the DJF SOI which is shown to be biased towards EP events (Table 4.1). Therefore, it is clear that reconstructions with records located in different regions and with different weightings have varying skill at capturing ENSO diversity.

Here, the locations of the proxy records used in the multi-proxy reconstructions are explored in order to provide an initial investigation into the climate during recent EP and CP events to highlight the signals from the proxies located in different regions. Focus is given to tree-rings from North America and corals from across the Pacific as these are commonly used within the multi-proxy reconstructions. The spatial anomaly patterns of variables to which the proxies are sensitive are investigated to highlight any possible regions of differences between EP and CP events which may be influencing the ENSO reconstructions.

4.4.5 North American tree rings and teleconnections during EP and CP El Niño events

A number of multi-proxy ENSO reconstructions contain a proxy network which includes tree rings from North America (ST98; MANN00, MCG10; EG13). A portion of the reconstruction variability will be due to background climatic influences, not directly linked to ENSO forcing. However, this region is thought to contain some of the strongest ENSO signals from tree-rings (Stahle et al. 1989). The conventional understanding of the ENSO-precipitation signal over south-eastern North America is of enhanced (reduced) winter and early spring precipitation during El Niño (La Niña) events (Ropelewski and Halpert 1987; Stahle et al. 1998). The enhanced precipitation during El Niño events lead to a recharge in the soil moisture and enhanced tree ring growth (Stahle et al. 1998). This relationship was the basis for tree-ring-based ENSO reconstructions in this region (e.g Stahle et al. 1998). ST98, MANN00, MGC10, EG13 all use tree rings in their multi-proxy reconstructions, the locations of the North American tree-ring records used are shown in Fig. 4.7, spanning western North America, southern central and east coast USA. A number of recent studies have investigated the ENSO-precipitation relationship in this region further, in light of the increased understanding of El Niño diversity. During EP and CP El Niño events a difference in the spatial precipitation patterns, along with a difference in the strength

of precipitation anomalies over North America has been found (Capotondi et al. 2015; Infanti and Kirtman 2016).

The spatial precipitation anomaly patterns for EP and CP El Niño events across the North American tree ring region, and the difference between them, are shown in Fig. 4.7, for Oct-March, the season in which the early wood tree rings are most strongly correlated (Stahle et al. 1998). Stronger positive precipitation anomalies centred on the west coast of America around 40°N are found for EP events compared to CP events. In addition, in the Gulf of Mexico the positive precipitation anomalies extend over a wider region during EP events compared to CP events. In relation to the sign of the anomaly in the locations of the tree-rings they are fairly consistent between EP and CP events over large areas of North America. However, differences in the spatial patterns between EP and CP events, depend upon the season, precipitation dataset and significance value used, as seen by Infanti and Kirtman (2016) and Capotondi et al. (2015), who found differences in their conclusions at local scales.

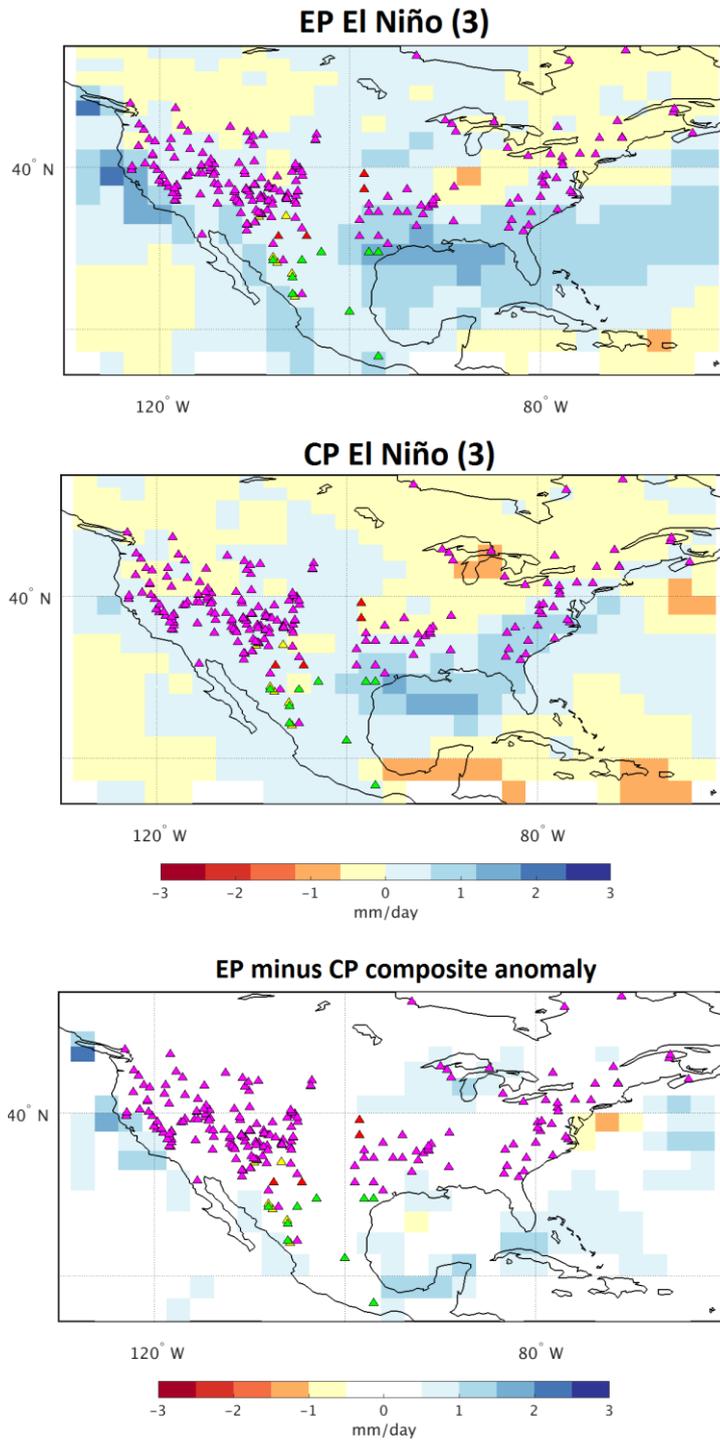


Fig. 4.7 Oct-March Precipitation anomalies from GPCP from NOAA composite plotting tool for EP events defined from NINO method (1982/82, 1991/92 and 1997/98) and CP events (1994/95, 2002/03 and 2009/10), and locations of tree ring proxies used from STH98=yellow, MANN00=red, MCG10=red and pink, EG13B=green. The bottom plot shows the EP El Niño anomalies minus the CP El Niño

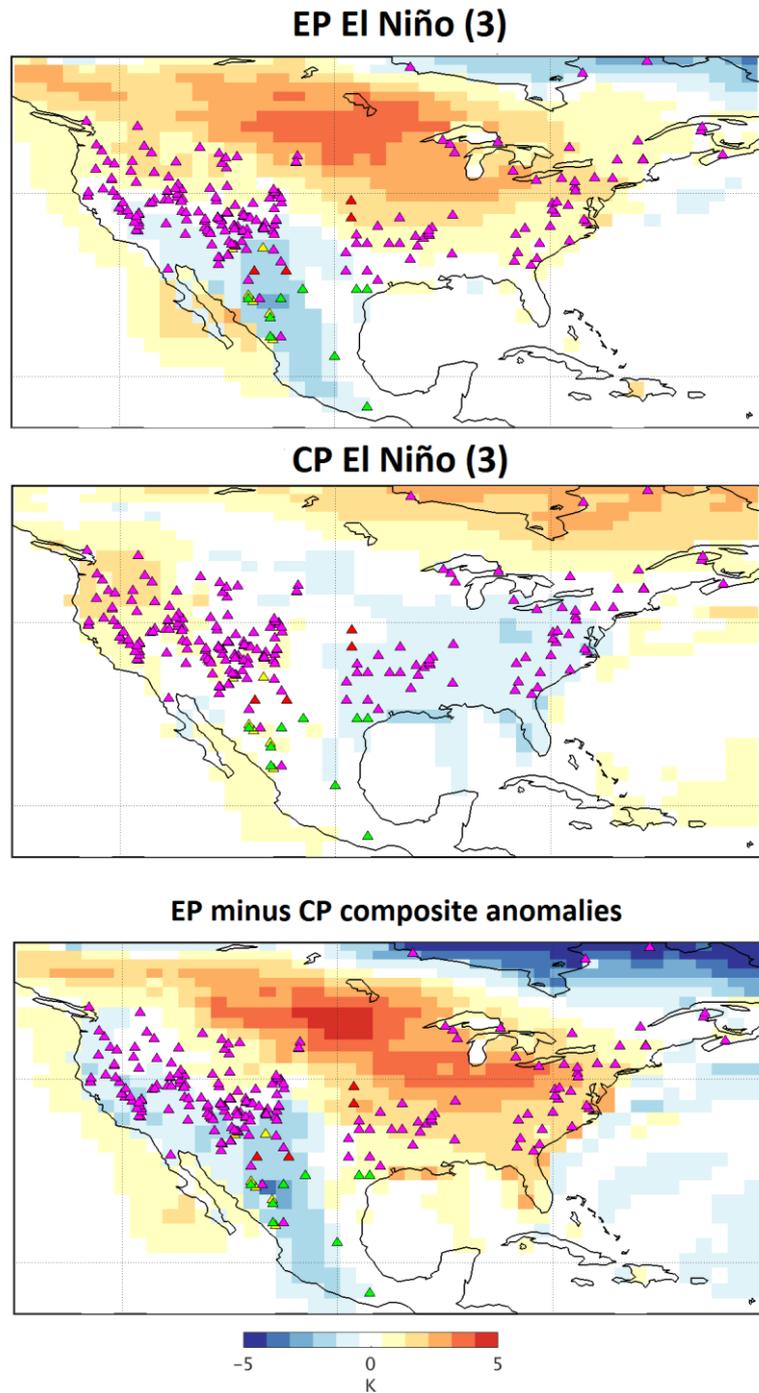


Fig. 4.8 DJF surface temperature anomalies from ERA-Interim for EP events defined from NINO method (1982/82, 1991/92 and 1997/98) and CP events (1994/95, 2002/03 and 2009/10), and locations of tree ring proxies used from STH98=yellow, MANN00=red, MCG10=red and pink, EG13B=green. The bottom plot shows the EP El Niño anomalies minus the CP El Niño

All of the tree rings used by ST98 are primarily sensitive to precipitation, whereas some of the tree rings used MANN00 and MCG10 over North America are more sensitive to air temperature. The DJF air temperature anomalies over North America are more wide spread and stronger during EP events compared to CP events (Fig. 4.8). During EP events there are positive temperature anomalies along the east coast of the USA, whereas negative temperature anomalies are found in this tree-ring region during CP events (Fig. 4.8). Thus, the temperature signal in this region is contrasting during EP and CP El Niño events. Much of the south-western North America experiences negative temperature anomalies in EP El Niño compared to CP El Niño, whereas the north-west experiences positive temperature anomalies in both types of El Niño event. Therefore, there is a difference in the spatial pattern of temperature anomalies during the two types of El Niño events over a number of tree-ring proxy record regions in North America. This results in implications for interpreting proxy records from this region when using conventional ENSO teleconnections (i.e. those more typically associated with EP El Niño). It also presents an avenue for future research to use the differences in EP- and CP- signal in this region to distinguish between paleo-ENSO events.

Due to the interaction of temperature and precipitation in their influence of tree ring growth, variations in signal could result in constructive or destructive interference between the signals from precipitation and temperature (warmer/wetter or cooler/drier) (Heried et al 2013). Constructive interference between variables describes when climatic conditions are warmer and wetter or cooler and drier, with both temperature and precipitation providing the same signal, enhancing or reducing tree growth respectively. However, in cases of destructive interference one variable promotes growth whereas the other contradicts this signal (i.e. in warmer but drier conditions). Destructive interference would result in the signal from the proxy record not indicating anomalous behaviour and thus not identifying an ENSO event due to an asymmetric response (Liang et al. 2016). Therefore, multiple-competing signals between temperature and precipitation in a region results in extra caution needed when interpreting any proxy which is sensitive to more than one climatic parameter.

4.4.6 Pacific corals and teleconnections during EP and CP El Niño events

Corals were used by a number of the multi-proxy ENSO reconstructions: MANN00, MCG10, W10 and EG13A/B/C. They are located across the Pacific and Indian Oceans as well as near to Australia, New Zealand, Indonesia and Africa. Corals are sensitive to local environmental conditions, but are also affected by larger scale oceanic and atmospheric patterns. They are used as an indicator of past ENSO behaviour as they are sensitive to changes in SST, precipitation and Sea Surface Salinity (SSS), all of which are subject to interannual variability due to ENSO (Brown et al. 2008). A number of techniques are used to analyse corals due to their sensitivity to these oceanic and atmospheric variables. Oxygen Isotope analysis ($\delta^{18}\text{O}$) of corals is primarily an indicator of SST, but can also be influenced by precipitation and SSS at local levels (Heiss 1994; Linsley et al. 1994; Brown et al. 2008). The Strontium to Calcium ratio can also be analysed in corals to be indicative of SST (Linsley et al. 2000), luminescence is indicative of precipitation and river runoff (Hendy et al. 2003; used in MCG10) while C^{13} is indicative of ocean circulation (Heiss 1994 in MANN00). Therefore, the ENSO-signal from corals can come from a range of interlinking factors which can vary from region to region.

The SST and precipitation spatial anomaly patterns during EP and CP El Niño events are investigated to increase understanding of the factors affecting corals from different regions. The locations of the coral records used by the multi-proxy ENSO reconstructions and the SST and precipitation composite anomalies from NINO method defined EP and CP El Niño events are shown in Fig. 4.9. The difference between EP and CP events can also be seen. Of the corals used in the ENSO reconstructions, only those which contain data within the logbook period, 1815-1850 are investigated here. The April-March year is used following the annual mean SST used by Evans et al. (2002), who state this as the natural climate year for the tropical phenomenon that are resolved within the coral data (Ropelewski and Halpert 1987). For the same reason, precipitation is also analysed over the April-March year (Fig. 4.10).

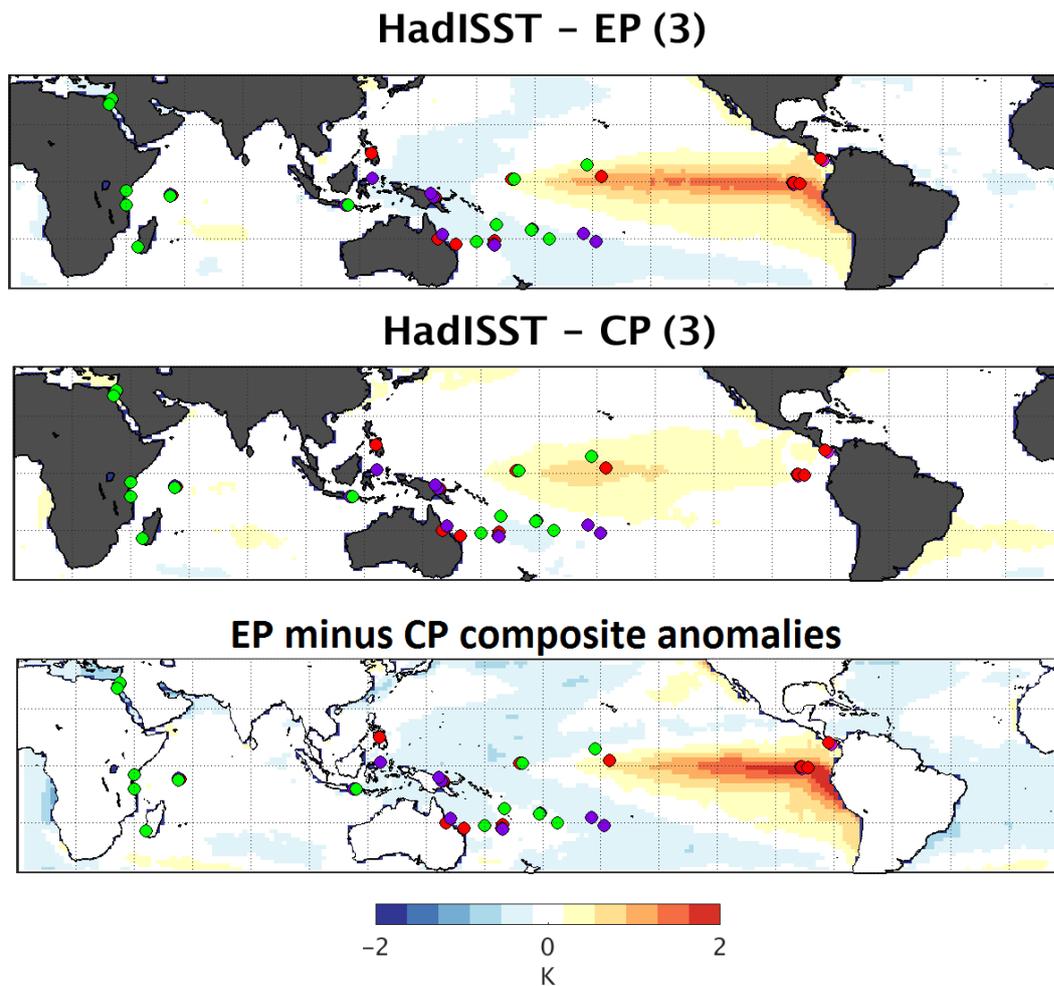


Fig. 4.9 Composite April-March SSTA (HaDISSTv4) for EP (above) and CP (below) El Niño years defined using NINO method. Dots indicate the location of coral records used by multi-proxy ENSO reconstructions (MCG10=Red; W10=purple and EGB13=green). The bottom plot shows the EP El Niño anomalies minus the CP El Niño anomalies.

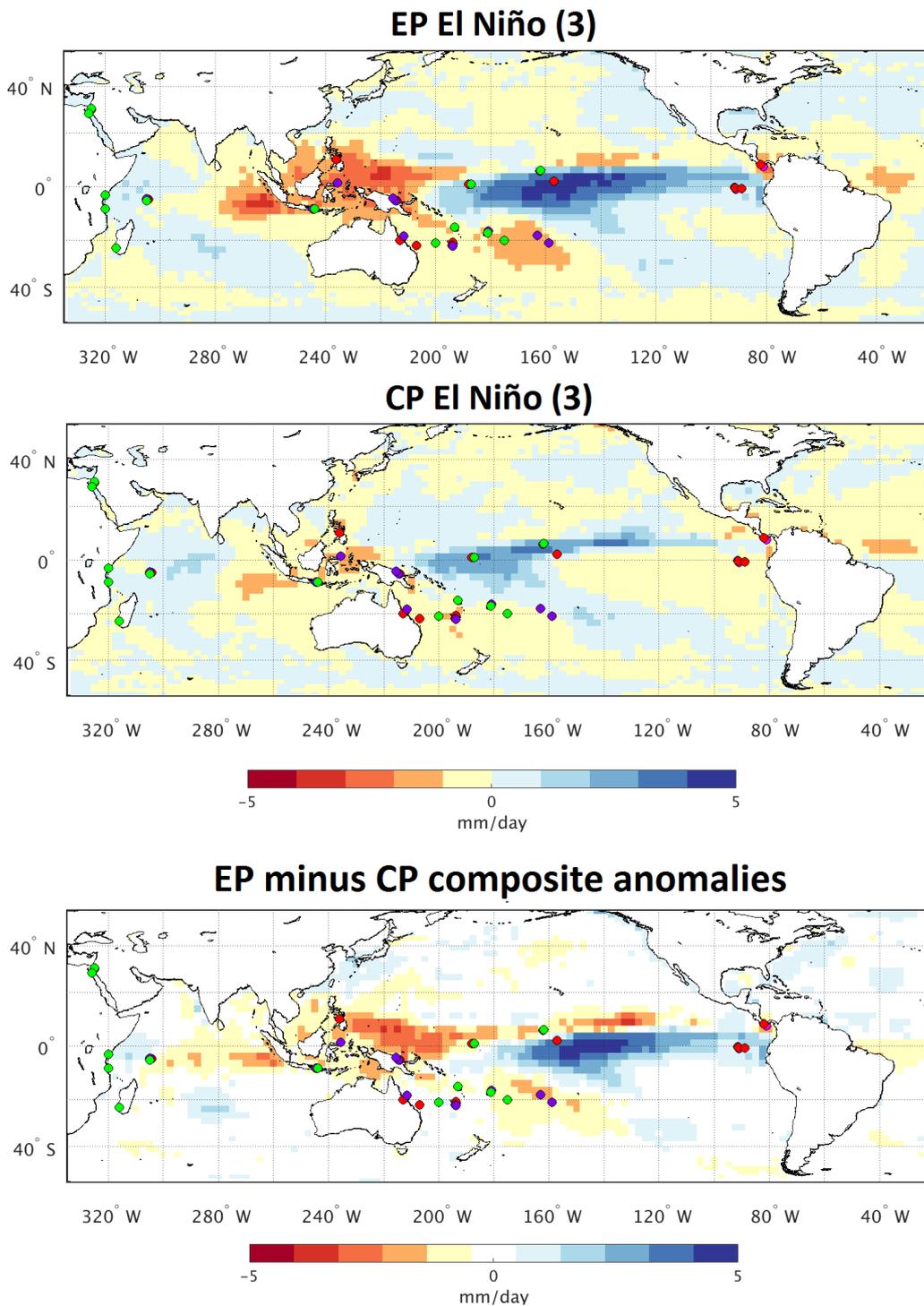


Fig. 4.10 April-March GPCP precipitation (mm/day) composite anomaly (1981-2010) climatology for the three HadISST NINO method defined EP El Niño events (1982/83, 1991/92, 1997/98) and the three CP El Niño events (1994/5, 2002/3, 2009/10). The bottom plot shows the EP El Niño anomalies minus the CP El Niño anomalies. Dots

indicate the location of coral records used by multi-proxy ENSO reconstructions (MCG10=Red; W10=purple and EGB13= green).

Strong positive SSTA are found in the eastern equatorial Pacific during EP events, whereas positive SSTA are shifted westward towards the Date Line during CP events (Fig. 4.9). The magnitude in the central region within the positive SSTA is higher during EP events than CP events. During EP events there are cool SSTA both north and south of the warm SSTA, in the mid-latitudes. These SSTA patterns are commonly discussed within the literature when assessing differences between EP and CP events (Ashok et al. 2007; Kug et al. 2009).

The differences in strength and location of the SSTA during the two types of El Niño have implications for the coral records used to reconstruct ENSO. For example, in the Eastern Pacific, corals from Urvina Bay, Galapagos Islands (Dunbar et al. 1994); used by MCG10) and Secas Island, Panama (Linsley et al. 1994) are in regions with a different SST response to ENSO flavours with SSTAs much stronger during EP than CP events. Located in the centre of the SSTA warm anomalies, it is expected that corals in this region would have a strong ENSO signal. However, previous work found that Eastern Pacific records have a low ENSO sensitivity (McGregor et al. 2010). The Urvina Bay corals were found to have zero to weak correlations with the ENSO index used by MCG10 and were therefore given a weighting of approximately zero. This poor ENSO signal was supported by Hereid et al. (2013) who found these Galapagos corals to have an El Niño skill of 6% and La Niña skill of 13%. Secas Island (Linsley et al. 1994) in East Pacific also has low El Niño (4%) and La Niña (4%) skill (Hereid et al. 2013). A relatively low number of proxies in East Pacific hinder understanding of the low skill in this region, however it is suggested that these corals are subject to non-climatic influences. Therefore, further research is needed in order to fully exploit ENSO information from these corals, located in a region of differences in the strength of the SST-ENSO teleconnection.

There is a much weaker SST pattern in the western Pacific, especially during CP events, and therefore precipitation dominates the ENSO signature in this region. Fig. 4.10 shows the precipitation anomalies associated with EP and CP El Niño events. The strongest precipitation anomalies are found in the central and western Pacific and over the Maritime Continent, as opposed to the strongest SSTA in central and eastern

Pacific. This explains the dominance of precipitation in the coral records over this region. The EP El Niño precipitation anomalies are stronger and wider spread than those for CP El Niño events (Fig. 4.10).

In the western tropical Pacific, the regional signature of ENSO is dominated by changes in rainfall in association with the migration of the Indonesian Low convective system (Cole et al. 1993). Therefore, corals from Tarawa Atoll, Kiribati are sensitive to regional convection due to precipitation-induced salinity changes. Papua New Guinea corals are also stated to be indicative of rainfall due to the abundance of, and isotopically light, rainfall with only small variations in temperature (Tudhope et al. 1995). For the corals from Espiritu Santo Island in the tropical south Pacific, the $\delta^{18}\text{O}$ signal is influenced by both SST ($r=0.77$) and rainfall-induced changes in SSS ($r=0.82$) (Quinn et al. 1993). New Caledonian corals (Quinn et al. 1998) are influenced by both SST and precipitation, however the region lies near the ‘hinge line’ separating regions of opposite response in SOI and weak correlations with SOI are found in this region. Therefore, the strength of the correlations with SOI vary depending on the location of the corals used in the reconstruction and their associated weighting.

Further work is needed in order to fully understand the differences in the proxy signals during EP and CP El Niño events. Schollaen et al. (2015) were the first to use oxygen isotope records from tree rings to detect the different flavour of ENSO. Using proxy records from regions which have uniquely different response during EP and CP events can help to make this differentiation. Only once these proxies have been fully exploited will we be able to gain a full understanding of the behaviour of ENSO and its diversity during the pre-instrumental period (Karamperidou et al. 2015). This would also help inform future ENSO reconstructions using data from ships’ logbooks.

4.5 Summary

An important finding from this chapter is the varied ability of the ENSO reconstructions to capture the different flavours of ENSO. This is summarised, along with a summary of the differences in the teleconnections between EP and CP El Niño events and the implications of this for El Niño reconstructions.

Different classification schemes can produce contrasting classifications of individual El Niño events. Almost half of the El Niño events between 1951 and 2010 have not been firmly classified (Wiedermann et al. 2016), due to the variations in methods used by the numerous classification schemes. There is a need for a more consistent and universally adopted classification procedure, as currently studies and conclusions are limited due to the variations in classification schemes (Pascolini-Campbell et al. 2015). In addition, for a more consistent list of El Niño years and their types since 1870, more improvements are needed to the reanalysis datasets (Yu and Kim 2013).

The instrumental DJF SOI, multi-proxy reconstructions and documentary chronologies were found to have better skill at capturing EP events than CP events over the common period, 1876-1977. The logbook-based reconstruction methodology and the predictor grid boxes used also appear to capture EP events better than CP events during the fitting period (Table 4.1). This would suggest that the historical logbook-based reconstructions may also contain an EP bias. It was found that of the five El Niño events identified by CPS B logbook-based reconstruction, four are suggested to be EP events (Table 4.2). The PCR method was also found to be biased towards EP events, but to a lesser extent (Table 4.1). Four of the six El Niño events found in the historical logbook-based reconstruction using this method were suggested to be EP events (Table 4.3). One reason for this is that in the fitting and calibration of the logbook methodologies, the SOI was used as the target reconstruction. The instrumental DJF SOI has been shown to have an EP bias (Table 4.1). Thus, the choice of calibration target could be an influence on the EP bias within the reconstructions. Another reason for the EP bias within the logbook-based reconstruction could be the spatial distribution of predictor grid boxes. Overall, it remains unclear as to the main cause of the EP bias within the logbook-based ENSO reconstructions

An EP bias was also found within the documentary based El Niño chronologies (Table 4.4). The documentary records are heavily reliant on records from South America, where precipitation anomalies have been found to differ between EP and CP El Niño events (Fig. 4.6). The precipitation teleconnections which these records are based upon are more comparable to those found in the EP El Niño composites compared to the CP El Niño composite. Therefore, the records based upon heavy rainfall in these regions will be biased towards identifying EP El Niño events. A difference in the precipitation

anomalies during EP and CP El Niño events is also found within previous work in this region, supporting the conclusions drawn here (Carré et al. 2014; Karamperidou et al. 2015). Therefore, it is suggested that the documentary-based records are based on an ENSO signal biased towards capturing EP El Niño events.

A difference in the skill of the multi-proxy reconstructions in capturing the different flavours of ENSO was also identified. All but one of the ENSO reconstructions capture EP events better than CP events when using the NINO method. Analysis of the teleconnections associated with tree-ring growth in North America and corals in the equatorial Pacific found that there are variations in the signal that would be obtained from these proxies during EP and CP El Niño events. In some cases, conflicting signals from multiple competing teleconnections could result in a lack of ENSO signal from the various proxy records in different locations. A further complication in the interpretation is introduced in regions in which the CP El Niño signal is similar to that of the La Niña signal for that given region. In order to obtain the best possible ENSO reconstruction from proxy records it is suggested that various proxy regions should be used in order to capture the most common signal from the proxy network. Further work should be done to address the differences in proxy signal between EP and CP events and their representation in proxy records. There is also large scope for an increased coverage of proxy networks. An increase in the number of corals from the Pacific Ocean would help to improve the reconstruction of the different flavours of ENSO (Wilson et al. 2010). Recent work has looked at using proxies from regions where the EP and CP signals are distinctly different, such as Java Indonesia, in order to distinguish between ENSO flavours (Schollaen et al. 2015). The results from this chapter suggest further regions where this is the case, such as around southern Africa when using wind proxies (Fig. 4.1), central USA for precipitation-based proxies (Fig. 4.7), eastern USA for temperature based proxies (Fig. 4.8) and corals from various regions across the equatorial Pacific (Fig. 4.9; Fig. 4.10). This demands more attention, in order to fully capture ENSO behaviour using proxy records.

Overall, it is clear that there is still disagreement between the various ENSO reconstructions which cover the early-mid eighteenth century, prior to the introduction of the instrumental record of ENSO. It is suggested that further work is still needed in digitisation of additional logbook data and collection of additional multi-proxy

reconstructions in order to produce an ENSO history based on records with a stronger agreement than currently attainable. With increased digitisation the observations from ships' logbooks has the potential to be a key data source for this future work. Such work should focus on expanding the possibilities of distinguishing the different flavours of El Niño through the use of differences in teleconnections associated with EP and CP El Niño events. Here it is suggested that there is a difference in the zonal wind anomalies during EP and CP El Niño events around South Africa, which lies on a key shipping route. Further investigation into regions with this differentiating signal will help to inform future digitisation of records and further ENSO reconstructions using this data source.

5. El Niño Southern Oscillation from instrumental records and historical reanalysis products: Mean climatology and ENSO indices

5.1 Introduction

The aim of this chapter is to assess ENSO within a number of atmospheric reanalyses. Attention is given to reanalysis products which span the twentieth century: ERA-20C (Poli et al. 2016) and the Twentieth Century Reanalysis (20CR) (Compo et al. 2011). Using both the Southern Oscillation Index (SOI) and the Niño 3.4 index as key indicators of ENSO, these historical reanalyses are compared to ERA-Interim, a reanalysis which is produced using a higher number of observations, including upper atmospheric data, from 1979 onwards. Poli et al. (2016) compared the SOI and the Niño 3.4 index from these three reanalyses and found a generally good agreement over their period of overlap. Here, this is expanded further by comparing the reanalyses with each other but also with the instrumental SOI and the Niño 3.4 index. As well as comparisons of indices, composite plots help assess the representation of ENSO and ENSO-teleconnections spatially within the reanalyses for key variables such as Mean Sea Level Pressure (MSLP), surface temperature and zonal wind. A key objective is to determine which of the reanalyses best captures ENSO over the twentieth century. This reanalysis can then be used to compare to historical climate model simulations which span over longer time periods, in Chapters 6 and 7.

Chapter 2 introduced ERA-Interim reanalysis which was used for the logbook-based reconstructions of ENSO. Here, reanalysis datasets are explained in more detail to understand the differences between the various reanalysis products. Reanalyses are commonly used in climatic studies as they provide a useful global dataset, of a number of variables, which is a coherent representation of global atmospheric circulation (Dee et al. 2011). Since the first attempt at a reanalysis, the First Global Experiment of the Global Atmospheric Research Programme in 1979, developments in observations, data assimilation and forecast models have resulted in a number of generations of reanalyses being produced by various institutes (Dee et al. 2011). The scope for historical reanalysis data to extend back in time is limited by the quantity and quality

of observations available to be assimilated. The first upper air observations were taken from 1940 onwards, and the modern rawinsonde network was developed from 1958 onwards providing further information on the state of the atmosphere above the surface (Mooney et al. 2010). However, it is not until 1979, with the start of the modern satellite era, that global observations are available to assimilate into reanalyses (Poli et al. 2016). The development of upper air observations over the twentieth century has resulted in an increase in the amount of observational data from several thousand surface observations per month in 1900 to a few million observations per month by 2010 (Poli et al. 2016). Here, the representation of ENSO in a reanalysis product which includes these satellite observations (ERA-Interim) is compared to the two historical reanalyses, which only use surface observations (ERA-20C and 20CR).

As discussed in Chapter 2, ERA-Interim reanalysis spans from 1979 to present and is produced by the European Centre for Medium-Range Weather Forecasts (ECMWF). It addressed some of the data assimilation problems which were highlighted during the production of the ECMWF's previous reanalysis product, ERA-40 (Dee et al. 2011). ERA-Interim covers the modern satellite era, from 1979 to modern day, and therefore assimilates a large number of observations throughout its entire time coverage. The forecast model used is ECMWF's Integrated Forecast Model with three fully coupled components for the atmosphere, land surface and ocean waves, and it advances forward using 12-hourly analysis cycles. Observations are assimilated using 4D-var analysis, a major advance from previous reanalyses (Dee et al. 2011). One of the boundary conditions for the atmospheric forecast model is a global estimate of Sea Surface Temperatures (SSTs). Prior to 2002, ERA-Interim used the same input as ERA-40, the weekly Optimal Interpolation Sea Surface Temperature (OISSTV1) (Fiorino 2004). However, since 2002, the SSTs used in daily operational forecasting were used (Dee et al. 2011).

In addition to ERA-Interim, two reanalyses which span the entire twentieth century are used, referred to here as historical reanalysis. Firstly, ERA-20C, another ECMWF product, is a reanalysis which spans from 1900 to 2010. Observations assimilated into ERA-20C are from surface pressure and marine wind observations only. HadISST (Hadley Centre Sea Ice and Sea Surface Temperature dataset) version 2.1.0 provides time-evolving SST and sea ice cover while the atmospheric composition changes and

solar forcing are those prescribed by the CMIP5 experiment (Poli et al. 2016). The observational surface pressure comes from the International Surface Pressure Databank (ISPD) version 3.2.6 (Cram et al. 2015) and the International Comprehensive Ocean-Atmosphere Data Set (ICOADS) v2.5.1. (Woodruff et al. 2011). In addition, marine wind reports for ICOADS are also used. The number of surface observations assimilated into ERA-20C per month increases from around 30,000 to 3.6 Million between 1900 and 2010 (Poli et al. 2016). As with ERA-Interim, the ECMWF Integrated Forecast Model is used to create the reanalysis and a 4D-Var data assimilation system combined the observations with the background state, however in 24-hour cycles, rather than 12-hour (Poli et al. 2016). ERA-Interim is preferred to ERA-20C post-1979 as ERA-20C doesn't include the upper air and satellite observations which are assimilated into ERA-Interim. Additionally, ERA-Interim has a shorter time step and produces a higher resolution output (Compo et al. 2011; Poli et al 2016).

The Twentieth Century Reanalysis (20CR) v2c provides an additional reanalysis dataset which covers the entire twentieth century, spanning back to 1851. Produced by the US-National Oceanic and Atmospheric Administration and the Cooperative Institute for Research in Environmental Sciences, University of Colorado, (NOAA-CIRES), 20CRv2c uses surface pressure observations from the ISPD version 3.2.9, which were assimilated using a 4-times daily ensemble Kalman Filter, based on an atmospheric model constrained with prescribed SST from Simple Ocean Data Assimilation with Sparse Input (SODAsi) version 2 prior to 2013 (Giese et al. 2015). Earlier versions of 20CR were forced by HadISST, however this updated version uses the new SODAsi as the SST input. HadISST2.1.0 and SODAsi differ in that SODAsi is informed by both observations of SST and the use of data assimilation to provide model-generated estimates of variables which are combined together (Giese et al. 2015). This process merges sparse observational data with estimates generated from data assimilation. On the other hand, HadISST is a combination of SST and sea ice observations using methods of optimal interpolation. Therefore, the model and data assimilation scheme used by 20CR differs from that used by ERA-Interim and ERA-20C, and it assimilates only surface pressure observations, not winds. Therefore, differences in model, data assimilation and observations will result in differences

between the reanalyses (Mooney et al. 2010). The resolution of 20CR ($2^\circ \times 2^\circ$) is also the coarsest of those used here. For each reanalysis, monthly mean values of MSLP, surface temperature and zonal wind are investigated. A DJF SOI is extracted from the MSLP values and a Niño 3.4 index from the temperature fields. The data are re-gridded to a common $1.5^\circ \times 1.5^\circ$ degree resolution to carry out direct comparisons.

Reanalysis products are commonly used as observations within climate studies. A wide range of studies have compared reanalyses to each other and to observations for various climatic indices and phenomena. A number of studies have investigated ENSO indices within reanalyses (Compo et al. 2011; Poli et al. 2016). Compo et al. (2011) compare the Sept-Jan Pacific Walker Circulation (PWC) Index in 20CR to ERA-40, ERA-Interim and NCEP/NCAR reanalyses. The Walker Oscillation is a primary atmospheric component of ENSO (Oort and Yienger 1996). Under normal conditions the Walker Circulation is characterised by rising motion over the western equatorial Pacific and Indian Oceans and sinking motions over the eastern equatorial Pacific. During El Niño conditions, the Walker Circulation is weakened and sometimes reversed. The PWC index is the difference in area-averaged vertical velocities at 500mb between the eastern equatorial Pacific ($10^\circ\text{S}-10^\circ\text{N}$, $180-100^\circ\text{W}$) and the western equatorial Pacific ($10^\circ\text{S}-10^\circ\text{N}$, $100-150^\circ\text{E}$) (Compo et al. 2011). This index captures the strength of the Walker Circulation, as the vertical pressure velocity indicates the rate of rising or sinking motion of the atmosphere. Compo et al. (2011) found that there was good agreement between the PWC Index from various reanalysis products, with strong positive correlations found during their periods of overlap (Fig. 5.1). The 20CRv2 PWC has significant correlation coefficients ($p < 0.01$) of $r=0.97$ with ERA-Interim (1989-2007), $r=0.95$ with ERA-40 (1958-2001) and $r=0.90$ with NCEP-NCAR (1948-2007). Therefore, over the second half of the twentieth century the 20CRv2 is in good agreement with alternative reanalyses in the interannual variability of the PWC, a useful index of the atmospheric circulation associated with ENSO.

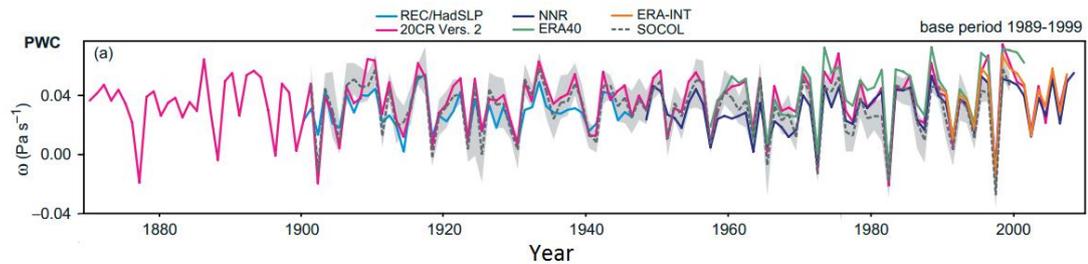


Fig. 5.1 Pacific Walker Circulation in HadSLP2 (The Hadley Centre’s Sea Level Pressure reanalysis), 20CR, NNR (NCEP-NCAR), ERA40, ERA-Interim and SOCOL (a chemistry-climate GCM) ensemble mean (grey shading is ensemble SOCOL max and min) (Adapted from Compo et al. 2011).

ERA-20C is one of the most newly available reanalysis datasets, therefore only a limited amount of comparison has been carried out using this dataset. Poli et al. (2016) compared the Niño 3.4 SST index and the SOI extracted from reanalysis datasets including ERA-Interim, ERA-20C and 20CRv2c (Fig. 5.2). In the early twentieth century there are some years in which the Niño 3.4 differs between ERA-20C and 20CR, however overall they agree well. There is good agreement between the SOI from the reanalysis datasets, in the later part of the twentieth century, but prior to 1940 ERA-20C tends to have more extreme values than 20CR (Poli et al. 2016). There is scope to expand on this initial analysis of ENSO from reanalyses spanning the twentieth century by Poli et al. (2016). Therefore, here, this centennial reanalysis data is compared to instrumental records of ENSO in order to assess which reanalysis dataset produces an SOI and a Niño 3.4 index closest to that which has been observed in the instrumental record.

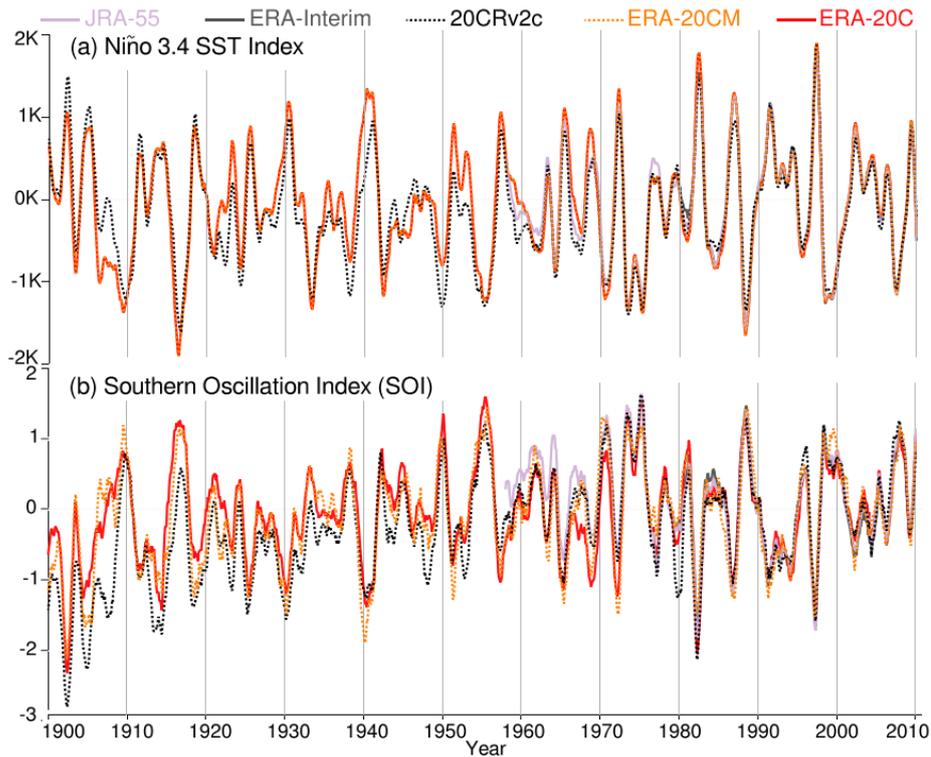


Fig. 5.2 Comparison of the Niño 3.4 index (above) and the SOI (below) from a number of reanalyses JRA-55 (Japanese 55-year reanalysis) ERA-Interim, 20CR, ERA-20CM (ensemble run with no observations assimilated) and ERA-20C (Adapted from Poli et al. 2016).

This chapter is organised as follows: *Section 5.2* describes the various datasets used to compare to the reanalysis products discussed in *Section 5.1*. *Section 5.3* summarises the methods used in this chapter, while *Section 5.4* presents the results of the climatology and ENSO index comparisons. *Section 5.5* summarises the key findings of this chapter and how they are used to inform the work of following chapters. This work within this chapter addresses Aim 3, specifically Objective 3A, as outlined in Chapter 1.

5.2 Data

In addition to the three reanalysis products discussed in Section 5.1, the instrumental record of the SOI and the Niño 3.4 were used. The instrumental SOI is available from the Australian Bureau of Meteorology (BoM), (<http://www.bom.gov.au/climate/current/soi2.shtml>) and is the difference in MSLP

between Tahiti and Darwin. Also available from BoM are the monthly MSLP from Tahiti and Darwin which were used in the calculation of the SOI. These are useful in comparing the MSLP in the reanalyses from these two key ENSO areas.

Measurements of SST over the Niño 3.4 region (see Fig 1.6 for location) are more complex than the SOI, which only requires station data records from two locations. The Niño 3.4 region spans 10° latitude and 50° longitude over the eastern central Pacific. A monthly index from NOAA-CPC, derived from the Extended Reconstructed Sea Surface Temperature V4 (ERSSTv4), extends as far back as 1950. Here, in order to provide consistency with this over the entire twentieth century, ERSSTv4 was used directly to create a Niño 3.4 index that spans the entire twentieth century. ERSSTv4 is provided by the NOAA/OAR/ESRL PSD (<http://www.esrl.noaa.gov/psd/>) and was derived from observations from ICOADS. Monthly mean SST is available on a 2.0° latitude x 2.0° longitude global grid from Jan 1854 to present (Huang et al. 2015). Monthly data were averaged into a DJF seasonal mean. They were then converted from degrees Celsius to degrees Kelvin to match the output of the other datasets.

Additionally, the Hadley Centre Sea Ice and Sea Surface Temperature dataset (HadISST), provides monthly global analyses of SST at a 1° latitude-longitude grid from 1870 to present (Rayner et al. 2013). The SST values from HadISST are based on in-situ SST observations from ICOADS, and from radiometers: Advanced Very High Resolution Radiometer (AVHRR) data from the Pathfinder project and Along-Track Scanning Radiometer (ATSR) reprocessing (Hersbach et al. 2015). Both HadISST and ERSST are commonly used as SST observational data to compare to reanalysis and climate model simulations (Michael et al. 2013; Li et al. 2016). Table 5.1 shows the original resolution of these reanalysis and SST analyses, and summarises their temporal coverage. All datasets were re-gridded to 1.5° x 1.5° resolution to enable a more direct comparison.

Table 5.1 Reanalysis and SST analyses used in this thesis: original resolution, time span and reference.

	Resolution (lon*lat)	Time Span	Reference
ERA-Interim	480*241	1979 - present	Dee et al. 2011
ERA-20C	360*181	1900 - present	Poli et al. 2016
20CRv2c	180*91	1851 - 2014	Compo et al. 2011
HadISSTv1	360*180	1870 - present	Rayner et al. 2013
ERSSTv4	180*89	1854 - present	Huang et al. 2015

5.3 Methodology

Firstly, the climatology of the reanalysis was assessed over the period 1979-2004. Spatial patterns of the mean DJF values of MSLP and surface temperature were compared over the tropical Pacific and Indo-Pacific region. Similarities in reanalyses were measured by comparison of climatological means and standard deviations. Two-sample t-tests were used to assess the similarity of the various reanalysis products to the instrumental record. For MSLP, comparisons were made with the instrumental MSLP time series from Tahiti in the central Pacific, and Darwin in Northern Australia, two key regions of MSLP variations during ENSO. An SOI was then constructed and comparisons of reanalysis-derived SOI with the instrumental SOI were made. For surface temperature, comparisons were made over the Niño 3.4 region, testing the similarity of the Niño 3.4 index from the reanalysis to that from ERSSTv4. ERSSTv4 was used as an independent data source for comparison, as it is not used to force any of the reanalyses. Pearson's correlation coefficient was used to assess the similarity of the indices over time. In order to assess ENSO representation further, the one standard deviation threshold was used to identify the number and frequency of El Niño and La Niña events (as in previous chapters). These comparisons were initially made over the period 1979-2004, which includes ERA-Interim. Following this, analysis over the entire twentieth century was carried out using the ERA-20C and 20CR. For both periods the mean and standard deviation over the 1979-2004 reference period was used in the calculation of climatic anomalies and for normalisation.

To calculate the SOI from the reanalysis products, the MSLP from regions covering Tahiti and Darwin were used. Firstly, monthly MSLP was calculated by extracting data from a western grid box (125°E-135°E, 17°S-7°S) which centres around Darwin, northern Australia, and an eastern grid box (205°E-215°E, 22°S-12°S) which centres around Tahiti, south-eastern Pacific (AchutaRao and Sperber 2002; Ji et al. 2014). This larger region is used for reanalysis and models compared to the point-source for observations due to the spatial resolution of the reanalysis (AchutaRao and Sperber 2002). Data were averaged to create a monthly MSLP for the western and eastern grid boxes spanning 1900-2010 (for ERA-20C). Fig. 5.3 shows the monthly MSLP from the Tahiti and Darwin grid boxes over the boreal winter months (DJF) from ERA-20C for example. The MSLP is higher over Tahiti compared to Darwin and there are clear interannual fluctuations in MSLP found in the two regions.

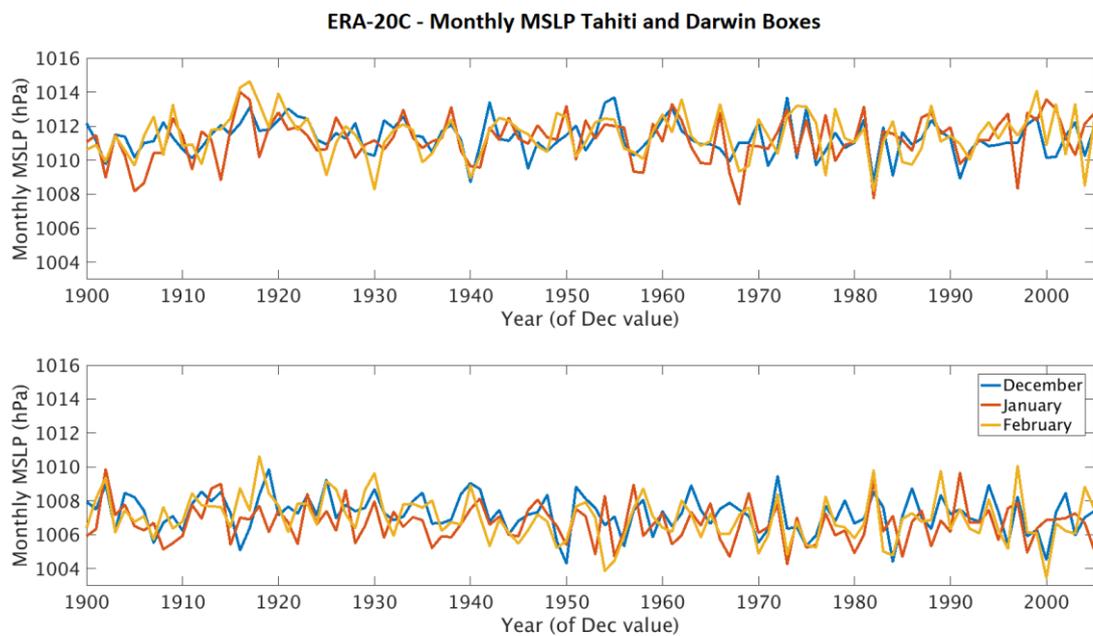


Fig. 5.3 Mean Sea Level Pressure over Tahiti (upper) and Darwin (lower) boxes, December, January and February, 1900-2004, from ERA-20C reanalysis.

The following formula was then used to calculate an SOI, using the Troup method (Troup 1965) as explained in Power and Kociuba, (2011):

$$SOI = 10(T - D - \overline{T - D})/\sigma_{T-D}$$

where here, T = monthly MSLP in the Tahiti box, and D = monthly MSLP in the Darwin box, $\overline{T - D}$ is the long term average of the difference between the Tahiti and Darwin box per calendar month, and σ_{T-D} is the standard deviation of the difference between them for the same calendar month. The climatology and standard deviations were calculated using the 1979-2004 analysis period, the period in which the historical model simulations overlap with all of the reanalysis datasets. Once the monthly SOI was calculated, the DJF values were averaged to create a mean seasonal index for the focus season of this analysis. The Niño 3.4 index was extracted from the re-gridded reanalysis data, using the skin temperature variable, creating a monthly and then a DJF average for the area covering 5°N to 5°S, from 170°W to 120°W. The correlation coefficient between the SOI and the Niño 3.4 index over the two periods from the various datasets were also calculated.

5.4 Results

5.4.1 Mean Sea Level Pressure climatology

Firstly, DJF seasonal mean values were assessed over the period 1979 - 2004. Fig. 5.4 shows the MSLP from ERA-Interim, ERA-20C and 20CR reanalysis. The three reanalysis datasets have broadly similar spatial patterns of MSLP across the tropical oceans. In the Pacific, there is relatively low pressure in the equatorial region, and higher pressure regions to the north and south between 20°N/S - 40°N/S. In the equatorial region, there is lowest pressure in the far western Pacific and across the south of the Maritime Continent, thus creating a pressure gradient across the equatorial Pacific. There are differences between the reanalysis datasets in the exact position of the low and high pressure centres, with the low pressure over a larger area in ERA-20C and 20CR compared to ERA-Interim and extending further north, across the equator, in 20CR.

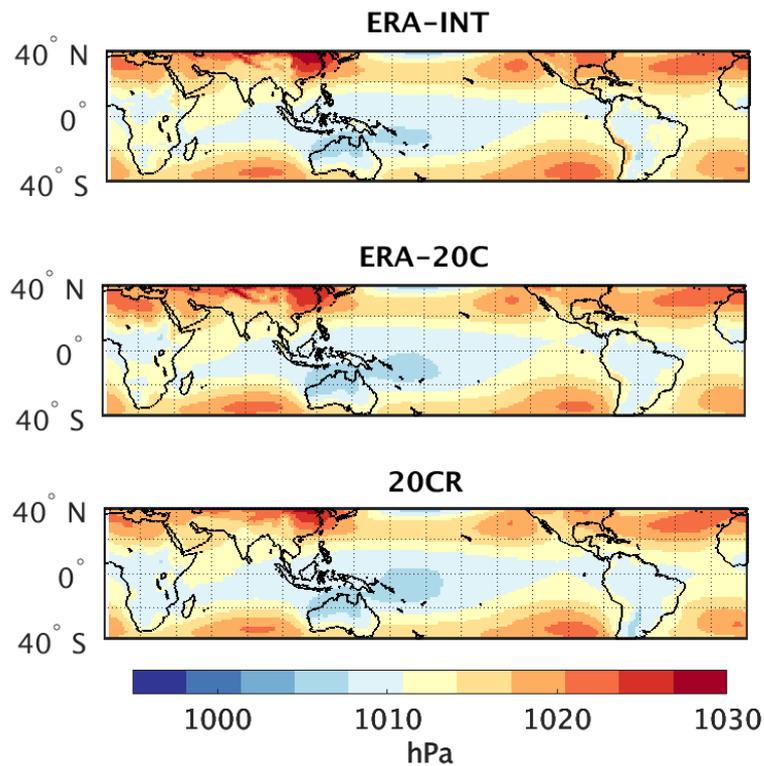


Fig. 5.4 Mean sea level pressure (hPa) DJF climatology 1979/80 – 2004/05 from reanalysis datasets (all $1.5^\circ \times 1.5^\circ$ latitude-longitude).

In terms of ENSO, the SOI is the most common atmospheric index used to identify ENSO events. Fig. 5.5 highlights the two regions from which the reanalysis-SOI data were obtained, with lower MSLP found in the Darwin grid box and higher pressure in the Tahiti grid box. ERA-Interim reanalysis data has a MSLP of 1007.5 hPa in the Darwin region, compared to 1011.6 hPa in Tahiti, a MSLP gradient of 4.10 hPa. This pressure gradient is a key component of the Walker Circulation. The ascending limb of the Walker Circulation is located over the far western Pacific and the descending limb in the eastern Pacific (Wang et al. 2012). This matches to the areas of low and high pressure in Fig. 5.5. All three of the reanalyses overestimate the MSLP in Tahiti (Table 5.2 and Fig. 5.5) and have a higher MSLP gradient than the instrumental record. One reason for the differences between the instrumental and the reanalysis is the use of a box SOI method within the reanalysis compared to the point observations of the instrumental record. The Tahiti grid box covers a region with more spatially variable MSLP than the more uniform MSLP within the Darwin grid box, as seen in Fig. 5.5. The location of the Tahiti observation station in the middle of the grid box represents

a more local area than the $10^\circ \times 10^\circ$ grid box used in the reanalysis, and Tahiti is located in a region of a higher spatial pressure gradient than Darwin. However, this box-SOI is commonly used by reanalysis and climate models studies (AchutaRao and Sperber 2002; Ji et al. 2014).

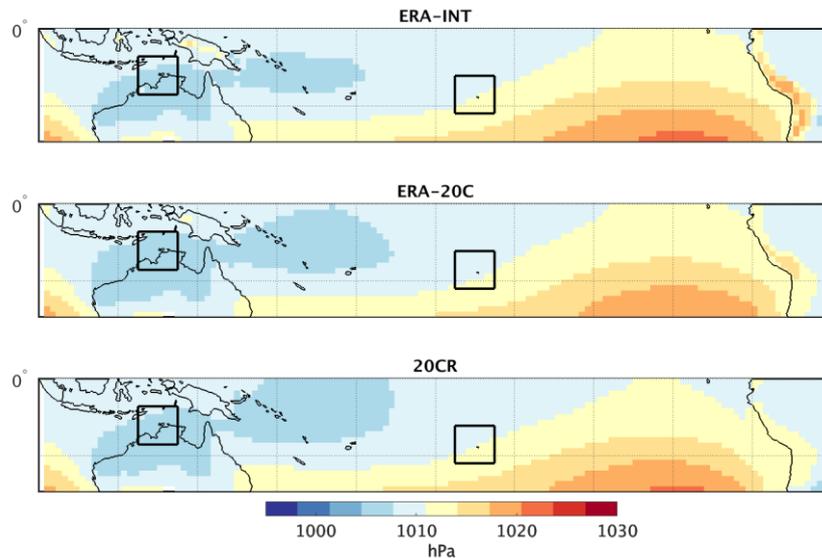


Fig. 5.5 DJF MSLP from reanalyses, 1979-2004. Black rectangles indicate the western and eastern regions used to define SOI from the reanalysis and model simulations.

Table 5.2 DJF MSLP from Darwin and Tahiti grid box (hPa), and MSLP gradient, 1979-2004 and 1900-1999.

	1979-2004			1900-1999		
	Darwin	Tahiti	MSLP	Darwin	Tahiti	MSLP
	MSLP	MSLP	gradient	MSLP	MSLP	gradient
	(hPa)	(hPa)	(hPa)	(hPa)	(hPa)	(hPa)
Instrumental	1007.1	1010.8	3.7	1006.7	1011.0	4.3
ERA-INT	1007.5	1011.6	4.1	---	---	---
ERA-20C	1006.8	1011.3	4.5	1007.0	1011.3	4.3
20CR	1007.1	1011.7	4.6	1007.0	1011.3	4.3

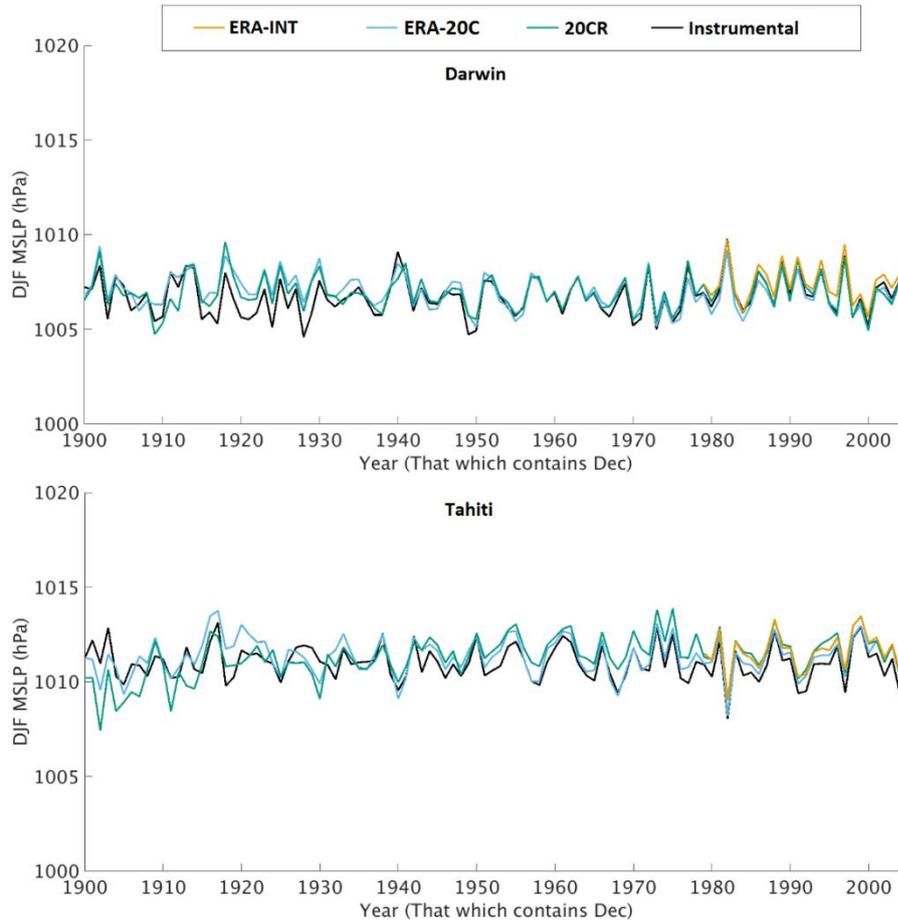


Fig. 5.6 DJF MSLP from reanalyses and instrumental records Darwin (upper) and Tahiti (lower), 1900-2004.

The year-to-year variability in the instrumental DJF MSLP in these two regions is well captured by the three reanalyses (Fig. 5.6). During El Niño years such as 1982 and 1997, there are high pressure anomalies in Darwin and low pressure anomalies in Tahiti, thus the MSLP gradient over the equatorial Pacific is reduced in El Niño years. Thus, the reanalysis data capture the interannual variability in MSLP in these regions well.

A two-sample t-test to test the null hypothesis that the DJF MSLP from the reanalyses and instrumental record are from datasets with equal means and equal but unknown variances was carried out for the Darwin region, and then the Tahiti region. This tests if the regional MSLP from the reanalyses are statistically different from those found in the instrumental record. A t-test result in which $H=0$ states that the null hypothesis cannot be rejected at the 95% confidence level, therefore suggesting similarity

between the two datasets. This is the case for all three reanalyses over the Darwin region, whereas only for ERA-20C in the Tahiti region when compared to the instrumental record, 1979-1999. This suggests that ERA-20C is not statically significantly different from the instrumental MSLP in Tahiti, 1979-2004. This is supported by Fig. 5.6 and Table 5.2, which shows that in Tahiti the reanalyses MSLP values are higher than the instrumental, and the Tahiti DJF MSLP climatology is 0.8 hPa (ERA-Interim) and 0.9 hPa (ERA-20C) higher than the instrumental MSLP, 1979-2004, whereas the 20CR MSLP is only 0.5 hPa higher than the instrumental record.

As discussed, this is due to the location of Tahiti within a region of spatially non-uniform MSLP. The reanalysis MSLP is more variable in the Tahiti grid box than it is over Darwin region. Over the twentieth century the MSLP gradient from the two reanalysis is the same as that from the instrumental record (4.3 hPa), although the instrumental record is slightly lower absolute values. Fig. 5.6 suggests that in the earlier part of the twentieth century there is also less agreement over Tahiti, with lower mean values in the reanalysis than in the instrumental record.

5.4.2 The Southern Oscillation Index

The DJF SOI was then calculated from the MSLP from the reanalyses, as shown in Fig. 5.7, 1900-2004. The instrumental DJF SOI was also normalised over the 1979-2004 period in order to provide direct comparison to that available from ERA-Interim and those from the other datasets over this period.

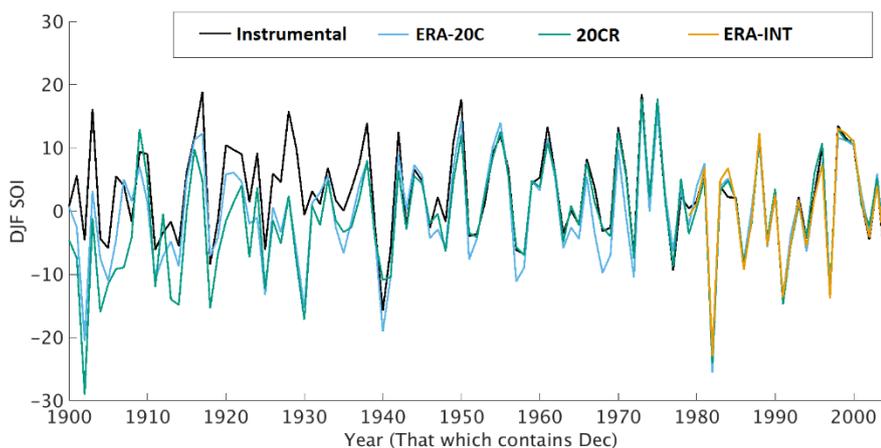


Fig. 5.7 The DJF SOI from instrumental record and reanalyses, 1900/01 -2004/5, with 1979/2004 reference period for all records.

Table 5.3 Pearson’s correlation coefficient DJF SOI from the instrumental and reanalysis data over different analysis periods, all significant at the 99% confidence level.

Pearson’s correlation coefficient between the DJF SOI and the reanalysis				
	1979-2004	1900-1949	1950-1999	1900-1999
ERA-INT	0.99	---	---	---
ERA-20C	0.99	0.81	0.96	0.88
20CR	0.99	0.72	0.98	0.80

The DJF SOI from all three reanalysis datasets (Fig. 5.7) are significantly correlated ($p < 0.01$) to the instrumental SOI over the 1979-2004 period ($r = 0.99$ for all three), as seen in Table 5.3. This replicates the strong agreement between SOI from reanalyses found by Poli et al. (2016) and builds upon their analysis by addition of comparison to the instrumental SOI. A two-sample t-test compared the SOI from the reanalyses and the instrumental record and found that the difference between the datasets over the period, 1979-2004, were not significantly different over this period ($H = 0$, $p = 1.0$ for all three reanalysis). Therefore, although differences in the Tahiti box MSLP in ERA-Interim and 20CR were found compared to the instrumental record, once the SOI is calculated this difference between datasets is reduced. One reason for the strong agreement is that the data are normalised over this period, 1979-2004. However, initial investigations using a different reference period (1933-1992 as in BoM), found that strong agreement still existed over this period, with correlation coefficients of $r = 0.98$ for all datasets. Here, for consistency the reanalysis, climate models and the instrumental record use the same 1979-2004 base period as this enables the inclusion of initial comparisons to ERA-Interim.

Over the period 1900 to 1999 the DJF SOI from ERA-20C and 20CR are significantly correlated ($p < 0.01$) to the instrumental DJF SOI, $r = 0.88$ and $r = 0.80$ respectively (Table 5.3). The reanalyses generally agree well with the instrumental record, with lowest agreement found in the earlier decades of the twentieth century (Fig. 5.7). There is a decrease in the strength of the correlation between the datasets in the earlier half of the twentieth century (1900-1949), compared to the latter half (1950-1999) (Table

5.3). The largest drop in magnitude is found when correlating 20CR and the instrumental record, with a positive correlation of $r=0.98$ over 1950-1999, which drops to $r=0.72$ over 1900-1949. Poli et al. (2016) also found the largest disagreement between reanalyses during the first decades of the twentieth century. The mean of the normalised SOI from ERA-20C (-1.39) and 20CR (-3.48) reanalysis are both more negative than the observed SOI (+3.14) during this period and the 20CR more negative than ERA-20C.

A two-sample t-test found that the DJF SOI from 20CR and ERA-20C are both statistically different from the instrumental DJF SOI, 1900-1950 whereas the null hypothesis cannot be rejected when comparing ERA-20C and 20CR to the instrumental DJF SOI over the latter half of the twentieth century (Appendix B, Table B1). Differences in the datasets over this earlier period could be due to the differences in the number and type of observations input into ERA-20C and 20CR. The inclusion of both wind and pressure observations in ERA-20C, compared to pressure only inputs into 20CR results in a higher number of observations. The data assimilation schemes and the way in which data-sparse regions are dealt with could also be a factor resulting in this difference, as observations from the Pacific are known to be limited prior to 1940 (Poli et al. 2016). This suggests that differences exist between datasets earlier in the twentieth century and therefore should be treated with additional caution.

The number of El Niño and La Niña events in each dataset was assessed by using the one standard deviation threshold (Table 5.4) as used in previous chapters. Previous studies analysing the Power Spectral Density found ENSO had a peak around 3 to 8 years (Hope et al. 2016). Allan et al. (1996) found a frequency of ENSO events from the instrumental record to be in the range of 3 to 7 years. From 1979 to 2004 the frequency of El Niño events from the instrumental record is four years, and five years for La Niña events. Due to the slight differences in the mean and standard deviations between the MSLP reanalysis datasets used to construct the SOI, differences in the frequency of ENSO events are found (Table 5.4). This in part highlights some limitations to the strict threshold definitions of ENSO events.

Table 5.4 Number of El Niño and La Niña events from instrumental and reanalysis DJF SOI, 1979-2004 and 1900-1999.

1979-2004				
	El Niño		La Niña	
	Number of events	Frequency (Years)	Number of events	Frequency (Years)
Instrumental	3	8.7	5	5.2
ERA-INT	4	6.5	4	6.5
ERA-20C	3	8.7	4	6.5
20CR	3	8.7	5	5.2
1900-1999				
Instrumental	15	6.67	17	5.88
ERA-20C	15	6.67	13	7.69
20CR	14	7.14	15	6.67

Over the entire twentieth century there are fewer events defined in the reanalyses compared to the instrumental record. ERA-20C only detects 13 La Niña events compared to the 17 events found within the instrumental record (Table 5.4). Over the twentieth century ERA-20C identified the number of El Niño events closer to that of the instrumental, whereas 20CR identifies the number of La Niña events closer to that of the instrumental out of the reanalyses. In Appendix B, Table B2, shows the years of the individual events from the various records over the twentieth century. Although there are 15 El Niño events identified in both the instrumental record and ERA-20, not all the events are the same with only 12 events found in both records. The same is the case for the 20CR events, which identifies a number of events prior to 1940 which don't match the instrumental record. This is a result of the selection method and the increased uncertainty in the datasets in the earlier part of the twentieth century. Differences in the number and years of events are commonly found using threshold methodologies for defining ENSO (e.g. Yeh et al. 2009).

However, overall the historical reanalyses capture the SOI well throughout the twentieth century when compared to the instrumental record. The largest disagreement is found in the early decades of the twentieth century, in which more negative SOI

values are found in the reanalysis. Overall ERA-20C captures the SOI better than 20CR during this period (i.e. with a higher correlation and slightly closer frequency of events compared with the instrumental record, Table 5.3; Table 5.4) and is therefore used as the best comparison to climate models over this period for SOI in Chapter 6.

5.4.3 Surface temperature climatology

In addition to using the MSLP and SOI to analyse ENSO, surface temperature and the Niño 3.4 index can be used, providing a measure of ENSO more related to the oceanic component of ENSO than the atmospheric component. When looking at surface temperature, in addition to using output from the three reanalyses, SST analyses ERSSTv4 and HadISST were used. Fig. 5.8 shows that there is a western Pacific Warm Pool, and a SST gradient across the equatorial Pacific in both observational analyses. Cooler SST values are found in the eastern Pacific due to upwelling of coastal waters, the Humboldt current. ERA-20C has a larger area of relatively cooler temperatures in the eastern portion of the Niño 3.4 box, expanding further west than in the other datasets, although 20CR, ERA-Interim and HadISST have this relatively cooler region extended further west than ERSSTv4. Overall they capture the spatial climatology of SST across the Pacific well.

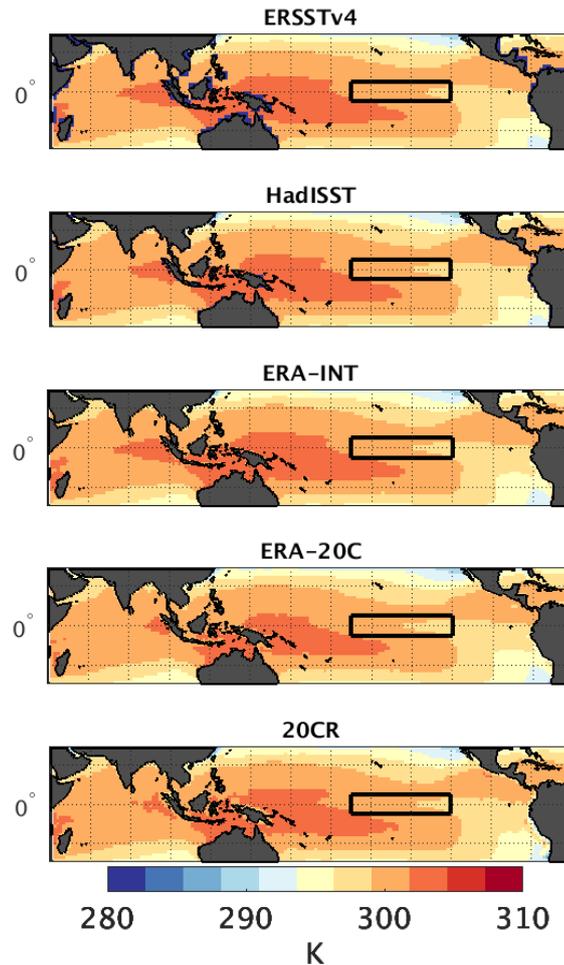


Fig. 5.8 DJF surface temperature from SST analysis and reanalysis, 1979-2004. Black rectangle indicates the Niño 3.4 region.

5.4.4 The Niño 3.4 index

The ERSSTv4 was used to construct a Niño 3.4 index directly from this data which extends back to 1854. Initial investigations showed little difference in the index from HadISST and ERSSTv4 for Niño 3.4 therefore the ERSSTv4 index was used henceforth and is referred to as the instrumental record, in order to provide consistency with the instrumental record available from NOAA-CPC (<http://www.cpc.ncep.noaa.gov/data/indices/>). To match the SOI normalisation period the DJF Niño 3.4 temperature anomalies were calculated relative to the 1979-2004 climatology. The Niño 3.4 index from the various reanalyses were compared to the ERSST record (Fig. 5.9), with the anomaly index seen in Appendix B, Fig. B1.

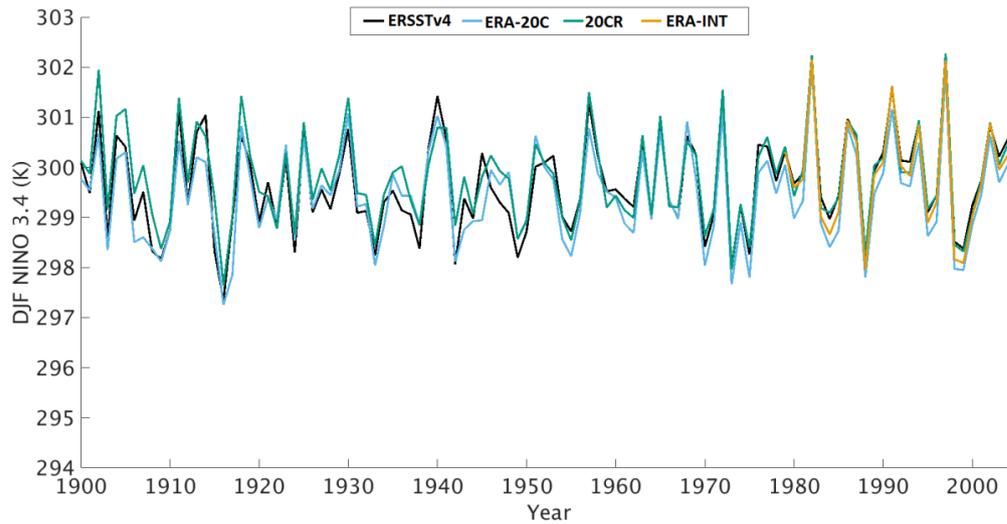


Fig. 5.9 The DJF Niño 3.4 from the three reanalyses (ERA-20C, 20CR and ERA-Interim) and ERSSTv4, 1900-2004 (Anomaly in Appendix, Fig. B1)

Table 5.5 The Pearson’s correlation coefficient between the DJF Niño 3.4 indices from the ERSST record and reanalysis data over different analysis periods, all significant at the 99% confidence level.

	1979-2004	1900-1949	1950-1999	1900-1999
ERA-INT	1.00	---	----	---
ERA-20C	0.99	0.90	0.97	0.93
20CR	0.99	0.93	0.99	0.95

Table 5.6 The DJF Niño 3.4 region mean surface temperature and standard deviation (std) over 1979-2004, 1900-1949, 1950-1999 and 1900-1999.

	1979-2004		1900-1949		1950-1999		1900-1999	
	std	mean	std	mean	std	mean	std	mean
	(K)	(K)	(K)	(K)	(K)	(K)	(K)	(K)
ERSSTv4	1.01	300.02	0.96	299.48	0.97	299.80	0.97	299.64
ERA-INT	1.13	299.87	--	--	--	--	--	--
ERA-20C	1.13	299.60	0.89	299.38	1.06	299.49	0.98	299.43
20CR	1.07	300.00	0.91	299.80	1.02	299.80	0.96	299.80

The reanalysis data have strong agreement with the instrumental Niño 3.4 index (Fig. 5.9). All three have strong significant ($p < 0.01$) correlations with the instrumental record. From 1979-2004 the Pearson's correlation coefficient between the instrumental record and ERA-Interim is 1.00, and is 0.99 when comparing ERA-20C and 20CR to the instrumental record (Table 5.5). As with the SOI the correlation coefficients between the reanalyses and the instrumental record decreases in the earlier half of the twentieth century (1900-1949) compared to the latter half (1950-1999). However, the magnitude of the reductions in strength is lower for the Niño 3.4 index than for the SOI and the correlations coefficients are higher for the Niño 3.4 than for the SOI over all the analysis periods (Table 5.5). This suggests that there is a larger agreement between the surface temperature variables in the Niño 3.4 region compared to the surface pressure variables over SOI regions. This may in part be due to the use of a box SOI in the reanalyses compared to a point observations in the SOI instrumental record, whereas the Niño 3.4 indices are from the same observational area in all of the datasets including the ERSST instrumental record.

ERA-20C has lower mean temperature values than ERSST, with the lowest 1979-2004 mean of the datasets, 299.60 K (Table 5.6). This is due to the extension of cooler temperature slightly further west than in ERSST as seen in Fig. 5.9. The 20CR has very strong similarity to the instrumental record with the same standard deviation and very similar mean (Table 5.6), while the two ERA-reanalyses have slightly lower means and slightly higher standard deviations. From Fig. 5.10, the 20CR reanalysis is slightly warmer than ERSST in the earliest half of the twentieth century (+0.32 K), whereas ERA-20C is slightly cooler (0.1 K) than ERSST in the latter decades of the twentieth century. This is reflected in a slightly cooler twentieth century mean in ERA-20C, 299.43 K, compared to 299.64 K in ERSSTv4, and a slightly warmer mean in 20CR. A two-sample t-test (Appendix B, Table B3) suggests that the reanalyses capture the mean temperature and its interannual variability in the Niño 3.4 region well over 1979-2004 (i.e. $H=0$) and the entire twentieth century, where the SOI was found to struggle to be as well represented.

Table 5.7 Number of El Niño and La Niña events defined using the 1.0 standard deviation threshold from the Niño 3.4 anomaly 1979-2004.

1979-2004				
	El Niño		La Niña	
	Events	Frequency (Years)	Events	Frequency (Years)
ERSSTV4	3	9	4	6.5
ERA-INT	3	9	4	6.5
ERA-20C	4	6.5	4	6.5
20CR	3	9	3	9
1900-1999				
ERSSTV4	18	5.5	16	6
ERA-20C	18	5.5	15	7
20CR	18	5.5	15	7

The frequency of El Niño and La Niña events in the instrumental record is best replicated by ERA-Interim, 1979-2004, (Table 5.7), however all datasets provided a realistic frequency of ENSO events, as expected from the similarities within the indices. ERA-20C suggests that El Niño events are more frequent than in the instrumental record, whereas 20CR underestimates the observed frequency of La Niña events. In Appendix B, Table B4, shows the years of the individual events from the various records over the twentieth century. Over the twentieth century the number of El Niño events defined from Niño 3.4 anomalies is the same from all three datasets, while La Niña is one less in the two reanalysis datasets than ERSSTv4. Therefore, overall the number of events is captured well in the reanalysis.

In the case of the SOI, 20CR does not match the instrumental record as well as ERA-20C in the earlier decades of the twentieth century. This is likely due to the fact that ERA-20C reanalysis assimilates both MSLP and surface wind observations, whereas 20CR only includes surface pressure observations. Therefore, when comparing climate model simulations to reanalysis in Chapter 6, ERA-20C is used as the

reanalysis over the twentieth century, whereas ERA-Interim is used when looking only at the modern satellite era.

5.4.5. Comparison of the DJF SOI and the DJF Niño 3.4 index

The DJF SOI from the reanalyses were then compared to the DJF Niño 3.4 index from these datasets. The instrumental record indicates strong agreement between these two commonly used indices of ENSO with a strong, significant negative correlation between the two indices over the entire twentieth century ($r=-0.82$) (Table 5.8). This is also the case when comparing the SOI to the Niño 3.4 index from the three reanalyses (Table 5.8). A small drop in the agreement of the records is found when comparing 1900-1999 with 1979-1999 in both ERA-20C and 20CR, however over the entire twentieth century the correlation remains strong, with $r=-0.85$ in ERA-20C and $r=-0.87$ in 20CR, 1900-1999. This suggests good coupling between the surface temperature and MSLP throughout the twentieth century, as expected from realistic reanalyses.

Table 5.8 The Pearson’s correlation coefficient between the DJF SOI and the Niño 3.4 index within the instrumental record and the three reanalyses over various periods covering the twentieth century.

	1979-2004	1900-1949	1950-1999	1900-1999
Instrumental SOI	-0.89	-0.75	-0.89	-0.82
ERA-INT	-0.92	---	---	---
ERA-20C	-0.89	-0.85	-0.88	-0.85
20CR	-0.90	-0.91	-0.91	-0.87

In order to investigate beyond these ENSO indices and gain a fuller understanding of the representation of ENSO within the reanalyses, composite anomaly plots for MSLP (Fig. 5.10; Fig. 5.11), surface temperature (Fig. 5.12; Fig. 5.13) and zonal wind (Fig. 5.14; Fig. 5.15) for ENSO events were created. Composite anomaly plots of El Niño and La Niña events defined by one standard deviation of the SOI and of the Niño 3.4 were compared. The 1979-2004 period was used to allow analysis of ERA-Interim along with the two historical reanalyses. However, composites using events over the entire twentieth century for ERA-20C and 20CR can be found in Appendix B for

comparison (Fig. B2-Fig. B7). As discussed above, the number of events defined using the SOI and the Niño 3.4 index differs and therefore the number of events informing the composite plots varies between dataset and index used (Table 5.4; Table 5.7).

The reanalyses capture the general spatial patterns of anomalies in MSLP, surface temperature and zonal wind during El Niño and La Niña events well however differences in the magnitudes of these anomalies and in their exact spatial distribution do exist. The strength of the MSLP anomalies in ERA-20C and 20CR are higher than for ERA-Interim for SOI-defined El Niño events, but lower than ERA-Interim in the Niño 3.4 defined El Niño events (Fig. 5.10; Fig. 5.11) For La Niña events, the 20CR anomalies in the Niño 3.4 defined events are stronger than in the ERA-reanalyses. Over the Darwin region, the MSLPA are positive during El Niño events and negative during La Niña events, while the opposite anomalies are found over Tahiti. This is the case for both the SOI- and Niño 3.4-defined events (Fig. 5.10; Fig. 5.11), and follows the typical spatial patterns associated with ENSO as explained in Chapter 1). The three reanalysis capture the patterns of MSLP well over these regions, thus supporting the strong agreement between the SOI in these three datasets, 1979-2004 (Fig. 5.7).

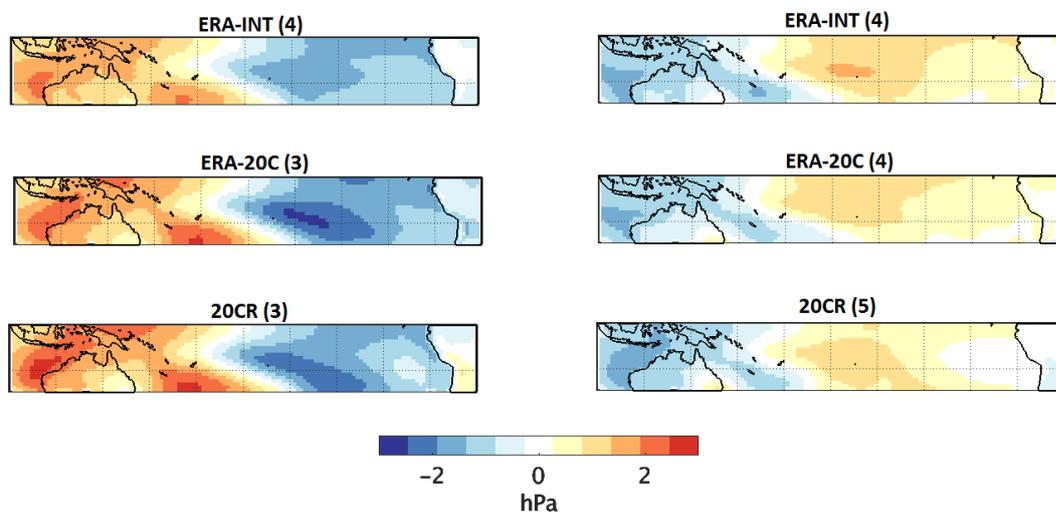


Fig. 5.10 DJF MSLP anomaly composites for El Niño events (left) and La Niña events (right) from reanalysis using the SOI one standard deviation threshold, 1979-2004. Numbers in brackets indicate number of events used to calculate each composite plot.

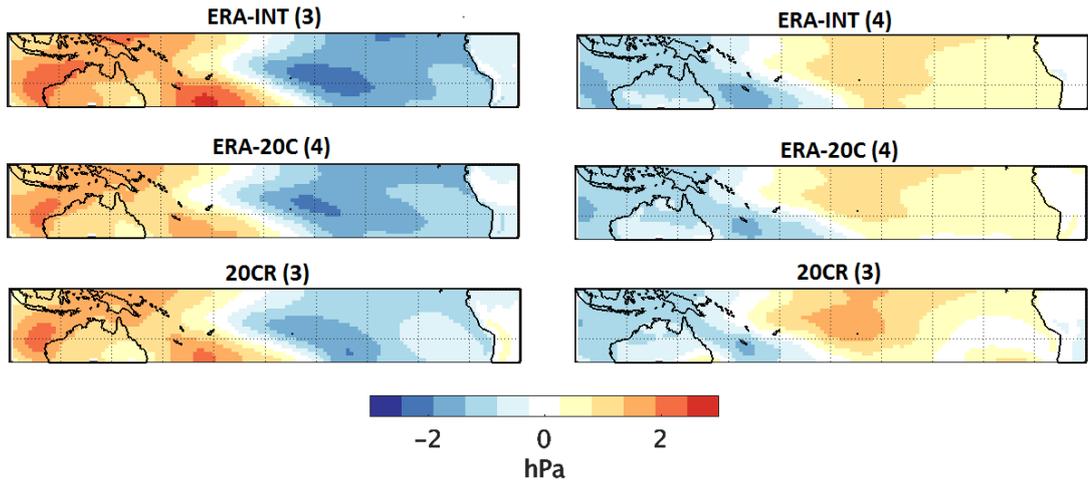


Fig. 5.11 DJF MSLP anomaly composites for El Niño (left) and La Niña events (right), from reanalysis using the Niño 3.4 one standard deviation threshold, 1979-2004.

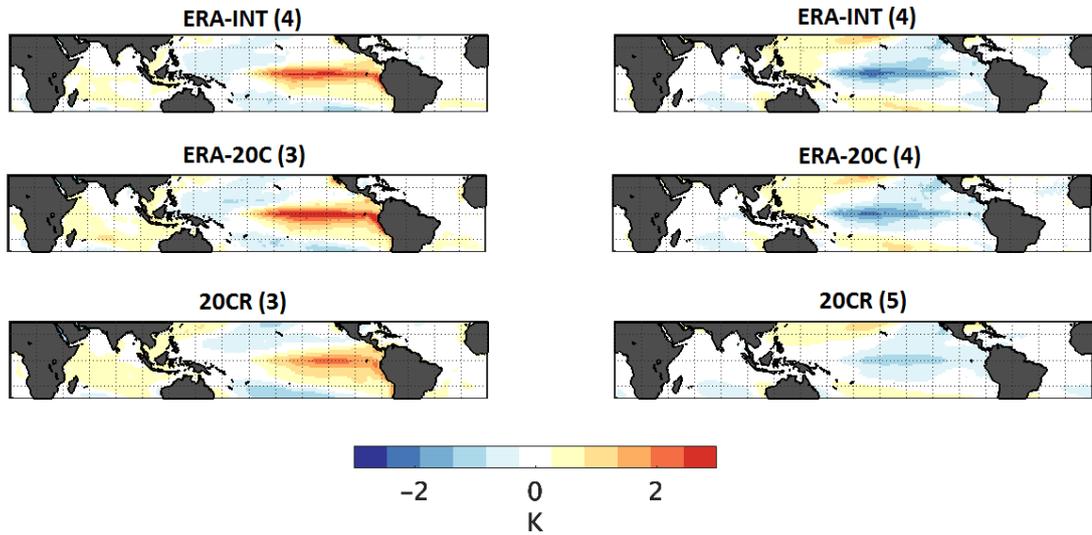


Fig. 5.12 Surface temperature anomaly composites for El Niño events (left) and La Niña events (right) from reanalysis using SOI one standard deviation, 1979-2004.

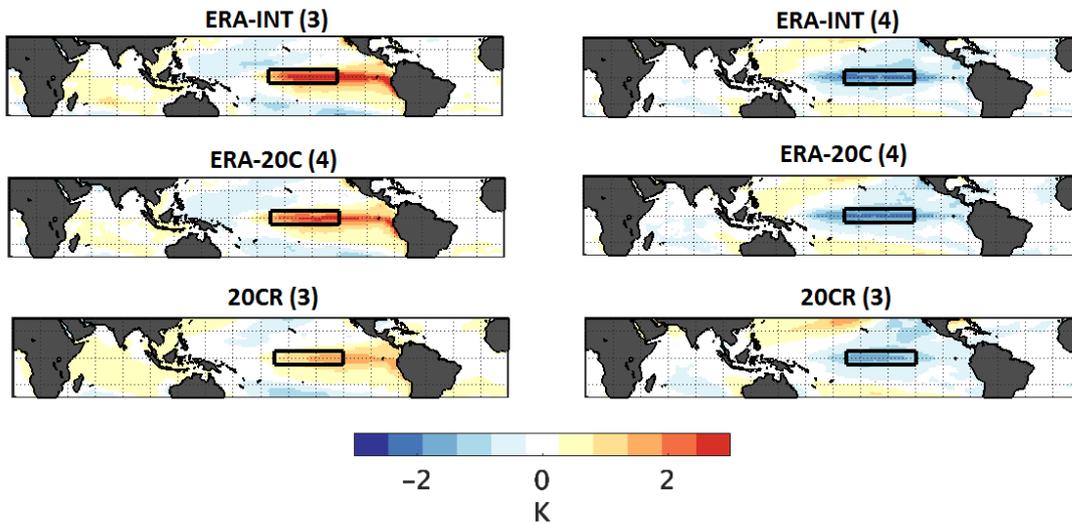


Fig. 5.13 Mean DJF surface temperature anomaly from reanalysis during years with Niño 3.4 one standard deviation defined El Niño (left) and La Niña events (right), 1979-2004. Black rectangle indicates the Niño 3.4 region.

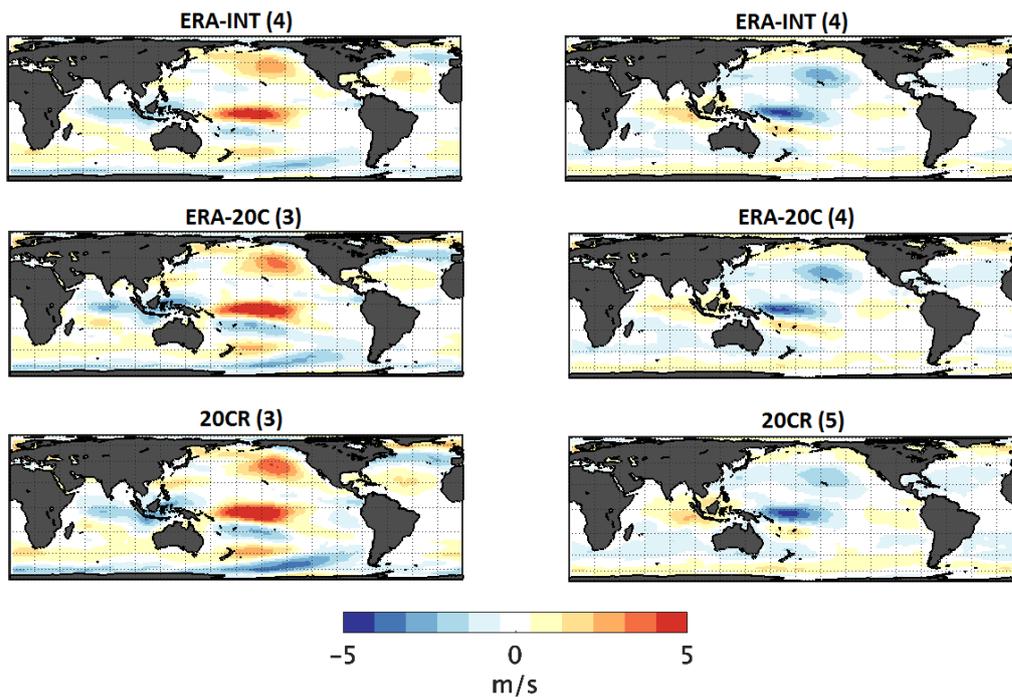


Fig. 5.14 Mean zonal wind anomaly (ms^{-1}) composites for El Niño events (left) and La Niña events (right) from reanalysis using the one standard deviation of the SOI, 1979-2004.

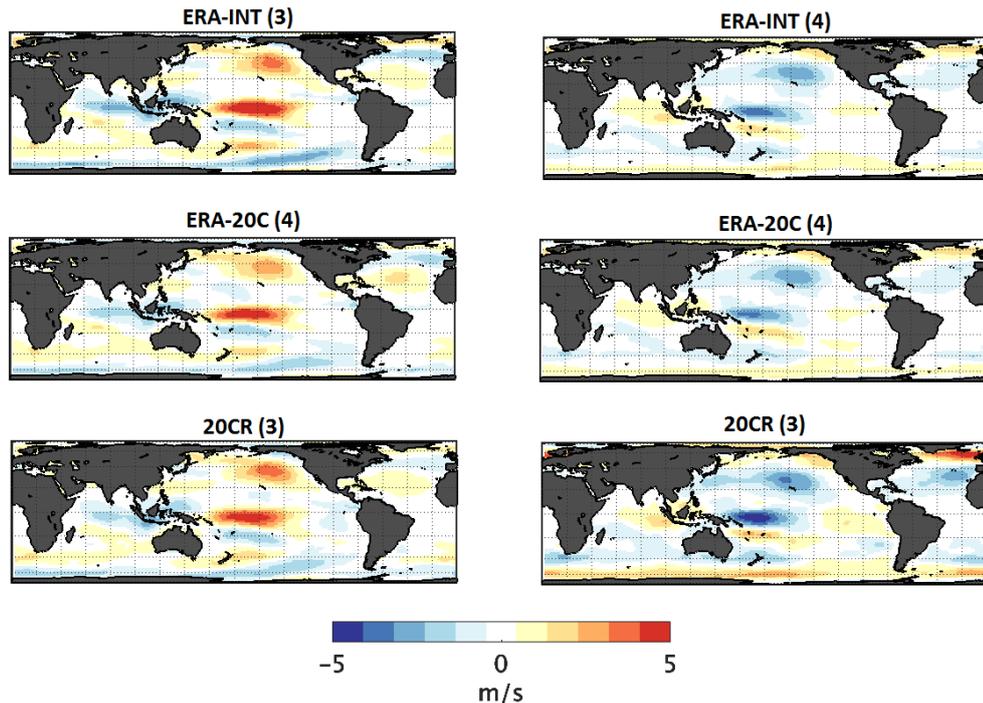


Fig. 5.15 Mean DJF zonal wind anomaly (ms^{-1}) composite during years with Niño 3.4 one standard deviation defined El Niño (left) and La Niña events (right), 1979-2004.

The magnitude of the surface temperature anomalies across the central and eastern Pacific are weaker within the 20CR composites compared to the ERA-Interim and ERA-20C composites, for both the SOI- and Niño 3.4-defined events (Fig. 5.12; Fig. 5.13). The location of the strongest zonal wind anomalies in the western Pacific are similar within the three reanalyses, but stronger for La Niña events in 20CR events defined by Niño 3.4 events (Fig. 5.14; Fig. 5.15). As ERA-Interim is seen as the superior reanalysis, due to its higher number of observations, increased spatial and temporal resolution, the spatial patterns within ERA-Interim are expected to be most realistic. Therefore the similarity of ERA-20C to ERA-Interim suggests that this may represent the spatial anomaly patterns associated with ENSO better than 20CR. This is especially the case when looking at surface temperature (Fig. 5.12 and Fig. 5.13). A good representation of spatial patterns of surface temperature and the magnitude of these anomalies is important consideration when looking into EP and CP El Niño events, discussed further in Chapter 6.

5.5 Summary

This chapter has built upon on Poli et al.'s (2016) comparison of ENSO indices within reanalysis by providing comparisons to the instrumental record. The conclusions of the inter-reanalysis comparisons are in agreement with Poli et al. (2016), with ERA-20C and 20CR having less agreement at the start of the twentieth century than later in the century. With the addition of the instrumental SOI it is found that ERA-20C captures the SOI best out of these two reanalyses during this period. This is useful because ERA-20C can then be selected as the best reanalysis to use in comparison to climate model simulations for this purpose in Chapter 6. However, it is interesting to note that 20CR captures Tahiti pressure better than ERA-20C. Overall ERA-20C is better than 20CR at capturing the SOI over the twentieth century, whereas there is no clear preference over the Niño 3.4 index captured from the two reanalysis. Thus for consistency between the Niño 3.4 index and SOI analysis, the use of ERA-20C as the preferential reanalysis in Chapter 6 is justified.

In addition, extending from Poli et al.'s (2016) index-based comparison, this chapter provided analysis of the frequency of ENSO events using the one standard deviation threshold and composite plots have been used to analyse the spatial patterns of MSLP, surface temperature and zonal winds during ENSO events in the historical reanalysis. ERA-20C composite plots perform better than 20-CR and therefore this provides additional justification for selection of ERA-20C for use in Chapter 6.

6. El Niño Southern Oscillation from historical climate model simulations: Model climatology, ENSO indices and ENSO diversity over the twentieth century

6.1 Introduction

6.1.1 Introduction to climate models

The use of climate models can help to develop our understanding of the climate of the past. The most comprehensive studies of past climate change include analysis of both climate model simulations and proxy-based reconstructions (Brohan et al. 2012; Jiang et al. 2015). During the instrumental period, proxy records and climate models can be compared to observations. However, when investigating beyond the limited instrumental record the use of proxies and models as a source of comparison to each other helps to provide the most robust understanding of past climate (Jones et al. 1998). In this chapter, historical climate model simulations and their climatological biases compared to historical reanalysis are investigated. In addition, two commonly used ENSO indices are explored: The Southern Oscillation Index and the Niño 3.4 index. Finally, ENSO diversity, in terms of the Eastern and Central Pacific El Niño events, over the twentieth century is investigated. Chapter 7 then provides a focus on zonal wind within the historical simulations, while Chapter 8 uses the climate model simulations to assess the stationarity in the ENSO-wind teleconnection and ENSO diversity over the Last Millennium.

The Coupled Model Intercomparison Project Phase 5 (CMIP5) was an extensive climate modelling project launched in 2008 with over 20 climate modelling groups worldwide performing a range of simulations using over 50 models (Taylor et al. 2012). A vast range of model outputs are available from this project. CMIP5 was used to inform the Intergovernmental Panel on Climate Change (IPCC) 5th Assessment Report and it built upon the success of previous CMIP projects by addressing major gaps in understanding of past and future climates (Taylor et al. 2012). Historical simulations cover the industrial period, 1850 to present, and are often referred to as twentieth century simulations (Taylor et al. 2012). Analysing this period enables an

evaluation of the ability of models to represent both natural and anthropogenic forcings over a period with good instrumental observational coverage.

In addition, CMIP5 was the first project of its kind to include paleo-simulations, with the aim of increasing our understanding of the past (Schmidt et al. 2014). Paleoclimate simulations carried out as part of the Paleoclimate Modelling Intercomparison Project Phase III (PMIP3), used in CMIP5, aim to assess the use of climate models used for future climate projections by using the same models to carry out simulations of past climates. The number of models used in the paleo-simulations is smaller than that of the full set of CMIP5 models. Therefore, although more historical simulations are available, this chapter only uses models which have both historical and paleo-simulations. Paleoclimate simulations have imposed boundary conditions which include solar variation, volcanic aerosols, well mixed greenhouse gases, land use and orbital parameters (Braconnot et al. 2011; Schmidt et al. 2011; Schmidt et al. 2012). The paleoclimate simulations in CMIP5 are separated into three periods: The Last Glacial Maximum (21ka), the Mid-Holocene (6ka) and the Last Millennium (850-1850 AD) (Schmidt et al. 2014). In the Mid-Holocene and the Last Glacial Maximum, the models all use the same boundary conditions prescribed by the CMIP5 paleoclimate experiments, but due to differences in the models' construction there are variations in the computed radiative perturbation from each model (Braconnot et al. 2011). For the Last Millennium (LM), the model groups used the same prescribed boundary conditions for orbital parameters, greenhouse gases, ozone and vegetation, however multiple reconstructions of solar irradiance and volcanic aerosols were available (Schmidt et al. 2011; Klein et al. 2016), with the number of reconstructions increasing as new datasets became available (Schmidt et al. 2012). Therefore, there are differences in the LM boundary conditions used between the models, discussed further in Section 6.2. However, Schmidt et al. (2011) describe this as an advantage due to the differing levels of uncertainty in the various solar and volcanic reconstructions.

The Last Millennium is a highly relevant period for paleo-simulations as it enables model and data comparison for a period in which the climate was in a similar mean state to the modern day, and a period in which the climate variability was dominated by slow variations in natural forcings such as solar irradiance and volcanic eruptions

(Braconnot et al. 2011). The LM period contains a good amount of proxy data which can help to assess past climate variations on annual and decadal timescales (Jiang et al. 2015). The Last Millennium simulations are investigated in Chapter 8, informed by analysis of historical climate simulations within this chapter and within Chapter 7.

6.1.2 ENSO in climate models

Climate models are mathematical representations of reality, however they are far from perfect and do not perform very well in the equatorial Pacific, the key ENSO region (Li et al. 2016). There is still much progress to be made in order for the models to fully capture modes of natural climate variability such as ENSO, and to increase confidence in the projections of their behaviour into the future. However, recent developments and new model comparison studies have shown that their ability to do so is improving (Christensen et al. 2013). Bellenger et al. (2014) compared the performance of the CMIP5 models to the CMIP3 models in terms of their ability to provide a realistic representation of ENSO. It was concluded that although improvements were evident, there was no large step increase in ENSO performance in the CMIP5 compared to CMIP3 (Bellenger et al. 2014). Some key issues remain in the representation of ENSO which relate to both the mean state of the Pacific Ocean, as well as ENSO-related feedbacks, especially those relating to convection and clouds (Guilyardi et al. 2009; Flato et al. 2013). These biases within the models lead to a number of errors in their representation of ENSO in terms of the amplitude, period, skewness and spatial patterns of ENSO teleconnections (Guilyardi et al. 2009).

In terms of sea surface temperatures (SST), CMIP5 simulations, in general, have biases in three areas: a cold bias over the equatorial Pacific cold tongue, a warm bias in the eastern oceans of the subtropical region, and a cold bias in the western portion of subtropical Pacific Ocean (Michael et al. 2013). Fig. 6.1 shows the observed SST climatology across the tropical Oceans and the locations of these three regions in which biases are commonly found. Fig. 6.2 shows the biases in SST across the equatorial Pacific from a multi-model mean of CMIP5 simulations (Li et al. 2016), with the cold tongue bias a clear feature along the equator. This bias limits the skill of predicting ENSO and ENSO-teleconnections within climate models (Ham and Kug

2012; Li et al. 2016). The warm bias in the eastern subtropical region can also be seen clearly in Fig. 6.2, and the large area of cold bias in the western Pacific.

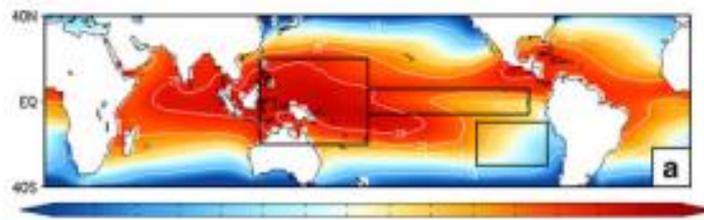


Fig. 6.1 Observed annual mean SST (Extended Reconstructed Sea Surface Temperature, ERSSTv3b), black boxes indicate regions of common SST bias within CMIP5 model simulations (Michael et al. 2013).

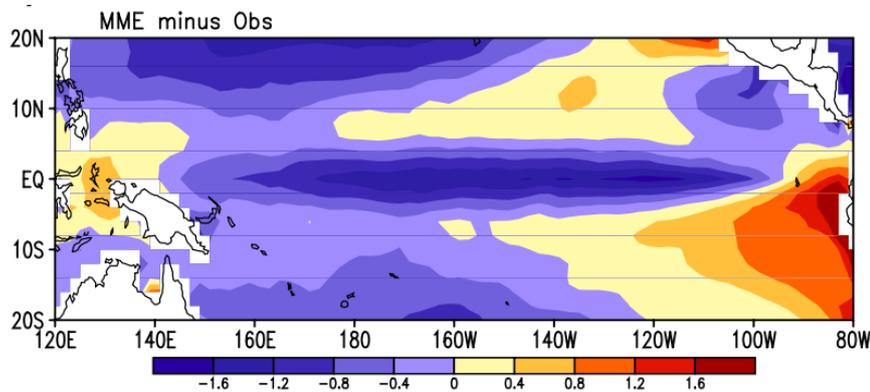


Fig. 6.2 Multi-model mean simulation of 24 CMIP5 CGCMs minus observations (Hadley Centre Sea Ice and SST dataset - HadISST) of annual mean SST over the tropical Pacific (Li et al. 2016).

The cold tongue bias is a commonly noted feature of climate models and it alters the representation of ENSO in two main ways. Firstly, a westward extension of the cold tongue reduces clouds and precipitation over the western Pacific warm pool which results in a weak SST-convective feedback. This leads to warming in the western Pacific which resembles La Niña-like pattern in SST (Li et al. 2016). Secondly, this excessive cold tongue results in surface easterly wind bias, forcing a thermocline that is too shallow over the equatorial eastern Pacific which thus promotes upwelling in the oceanic models. An easterly wind bias in the western half of the Pacific basin is common in Coupled Global Climate Models (Li and Xie 2014). Of the 24 CMIP5 models investigated by Li et al. (2016), those with the largest cold bias projected a weaker El Niño-like response, largely due to the weaker SST-convective feedback.

The cold tongue bias was reduced in CMIP5 models compared to CMIP3, however some biases such as central Pacific precipitation anomalies remain (Bellenger et al. 2013). Improvements in the representation of the mean state of the Pacific would lead to improvements in the representation of ENSO (Bellenger et al. 2013). Flato et al. (2013) state that there has been a 30% reduction in the cold tongue bias within climate model simulations since IPCC AR4. However, the CMIP3 and CMIP5 simulations still vary widely on their representation of 20th century changes in the tropical Pacific climate (Schmidt et al. 2014). Nevertheless, despite current biases, existing model simulations are useful for examination of spatial and temporal changes in ENSO characteristics and teleconnections (Lewis and LeGrande 2015), but it is important to be aware of these biases. Therefore, in Section 6.4, the climatological biases of the models are assessed prior to the evaluation of ENSO metrics.

6.1.3 The representation of El Niño diversity within climate models

Another issue with climate models is their representation of the different types of El Niño. As discussed in Chapter 1 and Chapter 4, El Niño events can be categorized into two general types, Eastern Pacific and Central Pacific (Wang et al. 2012). Most of the CMIP3 models were found to have systematic problems in simulating the distinctive patterns of SST and precipitation anomalies associated with the two types of El Niño (Ham and Kug 2012). A key reason for this can be linked back to the dry bias and the cold tongue bias found within a number of the models (Ham and Kug 2012). Therefore, in order for the two types of El Niño to be more successfully simulated these model biases within the climatology of the equatorial eastern Pacific need to be addressed (Ham and Kug 2012). The spatial patterns associated with the two types of El Niño events were improved in some CMIP5 models compared to CMIP3 (Kim and Yu 2012). However, many of the CMIP5 models, like CMIP3 models before them, were unable to realistically reproduce the differences between Central and Eastern Pacific El Niño events (Jourdain et al. 2013). Models tend to simulate a single type of El Niño rather than two types of El Niño (Ham and Kug 2012). Kim and Yu (2012) found that it is more difficult for models to reproduce observed intensity of EP events than the observed CP intensity, with only 50% of CMIP5 models able to realistically simulate both strong EP and CP events.

Eastern and Central Pacific El Niño events within models can be assessed by looking at the Niño 3 and Niño 4 index obtained from surface temperature in the historical model simulations. For consistency with analysis in Chapter 4, EP (CP) events are defined as years in which the Niño 3 (Niño 4) is greater than one standard deviation and greater than the normalised Niño 4 (Niño 3) (Kug et al. 2009; Ham and Kug 2012). The frequency of EP and CP events can also be analysed and the CP-EP ratio provides an indication of the number of CP events in a given period compared to the number of EP events (Xu et al. 2017). EP and CP events can be investigated further by looking at composite plots of key variables such as temperature and MSLP for the DJF season (Ham and Kug 2012). This provides an indication of the spatial patterns of ENSO and the teleconnections between variables.

6.1.4 Model evaluation: diagnostics and performance metrics

A main objective of this chapter is to evaluate models during the historical period, 1900-2004. Model evaluation is the comparison of model output with observations or reanalysis data, while inter-model comparison looks at evaluating models with other models. Model evaluation can be carried out qualitatively using ‘diagnostics’, such as spatial maps, time series analysis or frequency distributions (Knutti et al. 2010). Alternatively, quantitative evaluation can be carried out using ‘performance metrics’ (Knutti et al. 2010), which measure the agreement between a simulated and observed quantity. Performance metrics can be used to target a specific process such as ENSO. By compiling a number of metrics together a more comprehensive and objective overview of how well a model can represent ENSO can be formed. A number of previous studies have proposed a range of metrics for evaluating ENSO and ENSO-feedbacks within climate models (Guilyardi et al. 2009b; Bellenger et al. 2014).

Bellenger et al. (2014) evaluated the performance of CMIP3 and CMIP5 models using a number of performance metrics. An ‘ENSO SCORE’ was calculated from the average of four metrics that assessed the amplitude, structure, spectrum and seasonality of ENSO. Additional metrics relating to Bjerkness’ feedback and heat flux feedbacks were also calculated. In order to assess the Walker Circulation more specifically within the models, Sandeep et al. (2014) used the performance metric of the PWC (Pacific Walker Circulation) index, which as explained in Chapter 5 is a

measure of the vertical velocities at 500mb in the eastern and western equatorial Pacific which captures the strength of the Walker Circulation. Different studies use different metrics to assess climate models' ENSO performance. Here, metrics associated with the SOI and the Niño 3.4 indices are used. Focusing on the SOI instead of the Niño indices is less common within climate model approaches. This chapter therefore builds on previous work which tends to focus on either an oceanic or an atmospheric index, not both. Chapter 7 expands this further using metrics which address the correlations between ENSO and the zonal wind, addressing the ENSO-wind teleconnection in detail for the first time using these climate model simulations.

The aim of model evaluation is to assess the models in the historical period in order to select the best performing models in order to use simulations with these better models for the analysis over the Last Millennium. Previous ENSO studies have assessed models in the historical period and then continuing on into the Last Millennium have adopted a range of approaches and criteria for selecting and rejecting models. Lewis and LeGrande (2015) excluded one model (bcc-csm1-1) as its dominant ENSO periodicity was too short. Lewis and LeGrande (2015) and Hope et al. (2016) excluded the MIROC-ESM model from their LM analysis as simulations over the last 1000 years are known to display strong drift (Watanabe et al. 2011; Sen Gupta et al. 2013; Sueyoshi et al. 2013). However, Wittenberg (2009) notes that centennial changes in ENSO means that the observational period is not long enough to fully capture the natural range of ENSO variability. Therefore, even though Brown et al. (2016) find that over the historical period some models have ENSO characteristics closer to observations than others, they use all of the models over the Last Millennium due to the limited period available for assessment in the modern period. Therefore, no selection method is applied in Brown et al. (2016). Here, the selection criteria are used to select the top two models from this historical period for use in LM analysis in Chapter 8. This is justified as this approach uses the historical model simulations for their purpose of ranking models which is often overlooked in studies using LM simulations which use most or all of the climate models available regardless of their performance in the historical period. Focus is also given to interannual and decadal timescales in compliance with that available from the historical logbook-based ENSO reconstruction.

This chapter is organised as follows: *Section 6.2* introduces the new data source used in this chapter; the output from historical climate model simulations. *Section 6.3* describes the methodology used to assess the historical climate model simulations in terms of their performance relating to ENSO, and in comparison to the instrumental record and the reanalysis data. *Section 6.4* presents the results of these climate model evaluations focusing on MSLP, the SOI, the Niño 3.4 index and ENSO diversity. *Section 6.5* summarises the key findings from this chapter and how they will be used in the following chapters. This chapter addresses Aim 3, focussing on Objective 3B, as outlined in Chapter 1.

6.2. Data

In order to assess the climate model simulations available over the historical and LM periods a number of models were used. Only models with surface temperature, mean sea level pressure (MSLP) and zonal wind data available at monthly temporal resolution available over both the historical and LM periods were considered. While there is a multitude of climate models which run historical simulations, there is a more limited number of Last Millennium simulations available (Harrison et al. 2016). The surface temperature field is commonly used in model comparison for Niño 3.4 calculations (King et al. 2015; Lewis and LeGrande 2015) and skin surface temperature is used from the reanalysis products, as over the sea it is equivalent to SST. All datasets were converted for consistency from degrees Celsius to degrees Kelvin for temperature, hPa for pressure and ms^{-1} for wind.

Model output from a number of CMIP5/PMIP3 models was obtained via the Earth System Grid Federation (<https://esgf-node.llnl.gov/search/cmip5/>). The selection of models was limited by the availability of both historical and LM simulation outputs containing surface temperature (ts), zonal wind (uas) and Mean Sea Level Pressure (psl). Table 6.1 shows the key information for the five models, with this data available at the time of this analysis, which are used in this thesis: GISS-E2-R, HadCM3, IPSL-CM5A-LR, MIROC-ESM and MPI-ESM-P. These models were chosen due to the availability of all the required variables in both historical and Last Millennium simulations. For each model run the r1i1p1 ensemble member was used for consistency of comparison (Li and Xie 2014; Klein et al. 2016). R1 relates to a specific

set of initial conditions, i1 an initialization method and p1 the perturbed physics version.

Table 6.1 The five historical climate model simulations used in this thesis: Model name, institution, atmospheric resolution, time span and reference.

Model Name	Institution	Atmospheric resolution (lon*lat)	Time span	Reference
GISS-E2-R	NASA Goddard Institute for Space Studies	144*90	1851 - 2005	Schmidt et al. 2014
HadCM3	Met Office Hadley Centre	96*73	1859 - 2005	Collins et al. 2011
IPSL-CM5A-LR	Institut Pierre- Simon Laplace	96*96	1850 - 2005	Dufresne et al. 2013
MIROC-ESM	JAMSTEC*, AORI*, NIES*	128*64	1850 - 2005	Watanabe et al. 2011
MPI-ESM-P	Max Planck Institute for Meteorology	192*96	1850 - 2005	Jungclaus et al. 2013

*JAMSTEC – Japan Agency for Marine-Earth Science and Technology, AORI – Atmosphere and Ocean Research Institute (The University of Tokyo), NIES – National Institute for Environmental Studies.

As seen in Table 6.1, the original datasets are of varied spatial resolution and cover slightly different time periods. Therefore, the data were re-gridded to a common 1.5° x 1.5° spatial resolution and analysed over periods in which the reanalysis products and climate model outputs overlap (1979-2004 and 1900-1999). The solar and volcanic forcings used to set the boundary conditions for the LM simulations vary depending on the model used (Table 6.2). All but MIROC-ESM use the combination of Vieira and Solanki (2009) solar forcing reconstruction spliced to Wang et al. (2005), while a number of volcanic reconstructions are used. For all datasets, seasonal averages for boreal winter (DJF) were calculated, in line with the season of the logbook-based reconstructions and the peak of ENSO events.

Table 6.2 Forcings used as boundary conditions for Last Millennium simulations (adapted from Table 2 in Hope et al. 2016).

	Solar Forcing	Volcanic forcing
GISS-E2-R	Vieira and Solanki (2009) spliced to Wang et al (2005)	Crowley and Unterman (2013)
HadCM3	Steinhilber et al. (2009) spliced to Wang et al. (2005)	Crowley et al. (2008)
IPSL-CM5A-LR	Vieira and Solanki (2009) spliced to Wang et al (2005)	Ammann et al. (2007)
MIROC-ESM	Lean et al (2005)	Sato et al. (1993)
MPI-ESM-P	Vieira and Solanki (2009) spliced to Wang et al (2005)	Crowley and Unterman (2013)

6.3. Methodology

Firstly, the climatologies of the model data were assessed over the period 1979-2004. Mean DJF values of MSLP and surface temperature were plotted over the tropical Pacific and Indo-Pacific region. MSLP and surface temperature were analysed as these are two key variables used to define ENSO. In addition, zonal wind is assessed in Chapter 7 to investigate the representation of winds within climate models. Bias in the DJF climatology of these variables were calculated based on comparison to ERA-Interim, 1979-2004. Biases over 1900-1999 were compared to ERA-20C, however with added caution in interpretation due to the increased uncertainty in the historical reanalysis further back in time. Chapter 5 found strong agreement between ENSO indices from reanalysis datasets, and in their comparison to instrumental ENSO indices (Table 5.3; Table 5.5). Therefore, results here are not expected to be affected by the choice of reanalysis for comparison.

In order to assess ENSO, three main variables within the model simulations must be analysed within the relevant regions. Firstly, MSLP is the key atmospheric indicator of ENSO, with the difference in MSLP between Tahiti and Darwin used to calculate the Southern Oscillation Index. Secondly, surface temperature over the Niño 3.4 regions in the central equatorial Pacific will indicate which model best represents the

oceanic component of ENSO. TS/skin temperature is used here, as in previous studies to construct the Niño 3.4 index, and initial investigations found good similarity between surface temperature and SST justify use of this variable output. Finally, the zonal wind bias of the models needs to be assessed in order to identify which models capture zonal wind best. In Chapter 7, focus is given to zonal wind within the regions from which the logbook-based-ENSO reconstruction was carried out.

In addition to climatological biases, the models were assessed for their SOI and Niño 3.4 index, and the spatial composite anomalies from events defined using these two indices are compared. Climate models cannot be expected to replicate the observed timings of ENSO events, therefore analysis of ENSO within models focuses on the statistical nature of ENSO such as frequency and standard deviations (Hope et al. 2016). The Niño 3.4 SSTA is commonly used from models to infer ENSO variability. However, due to the biases within the Niño 3.4 region this index is not always the most suitable in capturing ENSO from models (Dai and Arkin 2017). Alternatively, the Southern Oscillation Index can be used to assess ENSO within models. Both are used here, adding to previous studies which have commonly only focused on one index, most often the SST-based ENSO indices. Inclusion of analysis of the SOI also allows a more comprehensive overview of the atmospheric circulation associated with ENSO, important for the focus on winds in ENSO-teleconnections in Chapter 7, looking at the stationarity of ENSO-wind teleconnections with regard to the logbook-based ENSO reconstruction. It also links this analysis more directly with that carried out using the SOI in Chapter 2.

Simple metrics such as mean and standard deviation provide a basic understanding of the differences in the indices obtained from the different datasets. Following on from this, the frequency of ENSO events was calculated using the one standard deviation threshold method as used with the ENSO reconstruction in Chapters 2 and 3 (Stahle et al. 1998). These were then compared to the observed frequency of events during this period. In addition, composite anomaly plots were made based on the mean SST and MSLP during those years classified as El Niño and La Niña. The spatial patterns of ENSO-related variables from the model simulations were then compared to those from ERA-Interim to investigate ENSO-teleconnections during extreme events, and to ERA-20C over the entire twentieth century, 1900-1999, in Appendix C.

Finally, ENSO diversity was assessed, using the NINO method as explained in Chapter 4. This focuses on spatial composite plots of EP and CP El Niño events defined from the individual model simulations, compared to ERA-20C, over the period 1900-1999. This analysis uses the entire twentieth century due to the limited number of EP and CP El Niño events obtained when using only the shorter period 1979-2004. The frequency of events and the ratio of CP to EP El Niño events is compared within the reanalysis and climate model simulations. The results from this chapter are then used to inform the analysis within Chapter 7 and 8, and provide detail which contributes to the overall discussion of the climate model analysis carried out in this thesis, discussed in Chapter 9.

6.4. Results

Firstly, the MSLP climatology and the SOI from the historical climate models are assessed, concluding with a ranking of the models based on these variables. This analysis is then repeated for the surface temperature and the Niño 3.4 index.

6.4.1 Mean Sea Level Pressure climatology

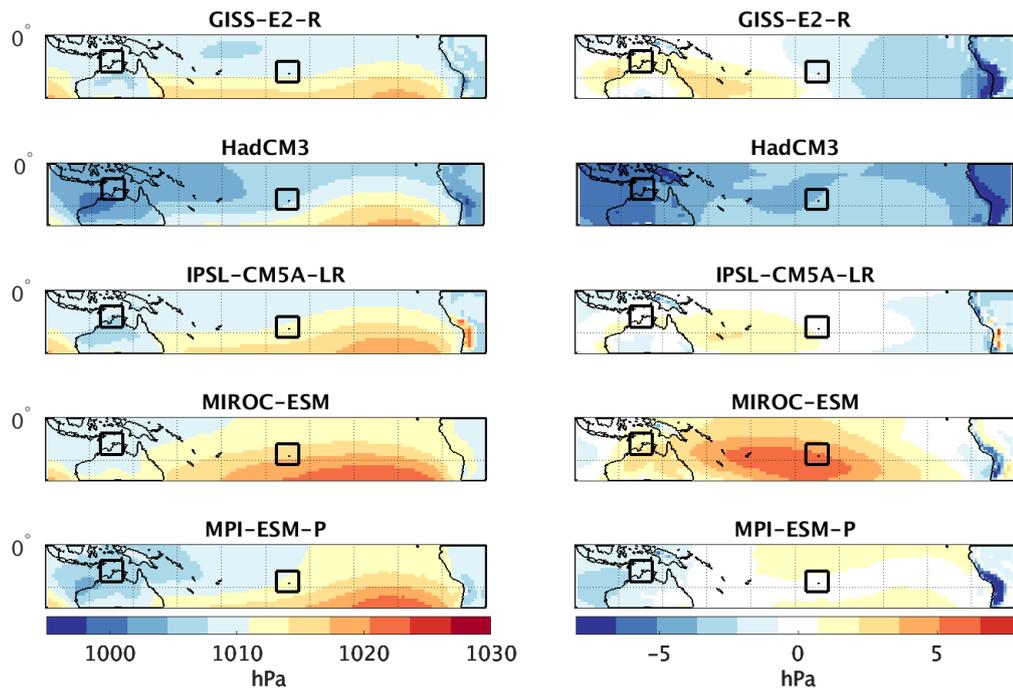


Fig. 6.3 DJF MSLP climatology (hPa) 1979-2004 from historical climate model simulations (left) and models' MSLP climatology minus ERA-Interim reanalysis (right). Black rectangles indicate the western and eastern regions used to define SOI from the reanalysis and model simulations.

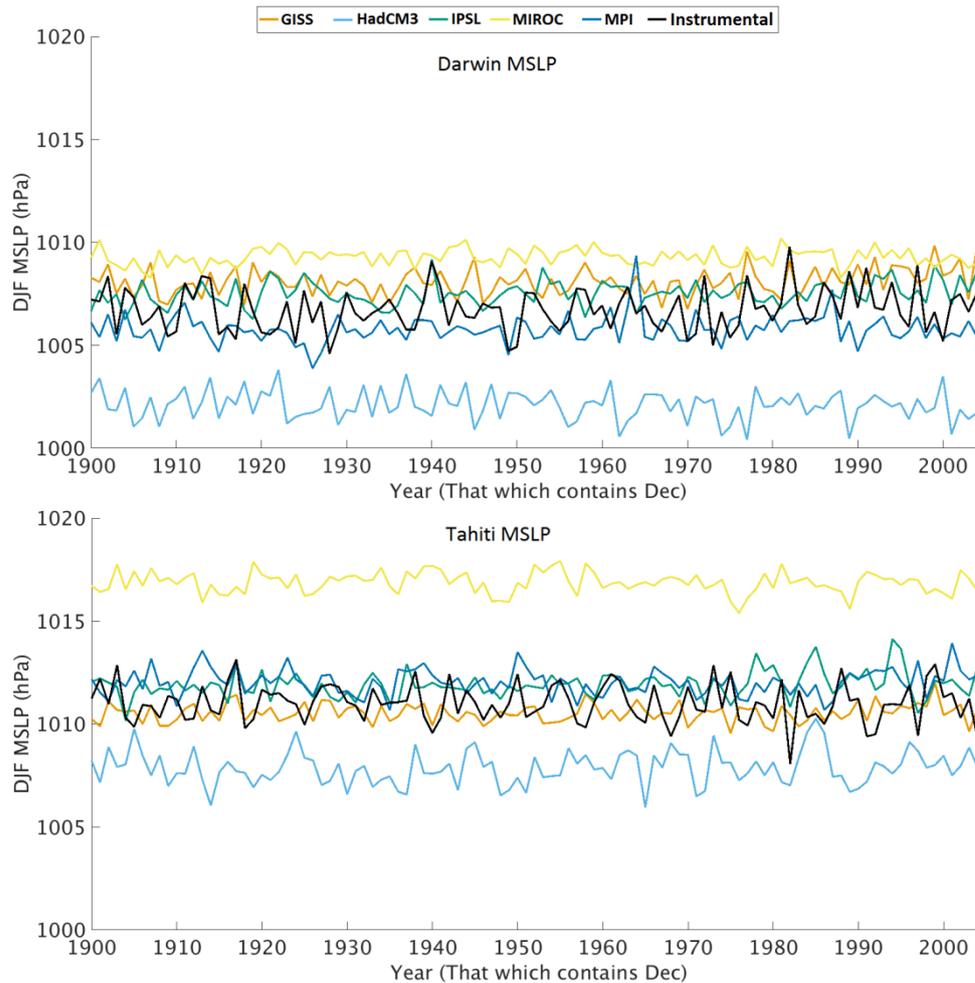


Fig. 6.4 DJF MSLP (hPa) in the Darwin (top) and Tahiti (bottom) regions, 1900-2004 from instrumental record and historical climate model simulations.

The spatial climatology of the DJF MSLP from the five historical climate model simulations and their bias with respect to ERA-Interim, 1979-2004, are shown in Fig. 6.3. HadCM3 has a negative MSLP bias across the entire equatorial Pacific, and especially over the Maritime Continent and western South America. MIROC-ESM has a positive bias over much of western and central Pacific. This can be seen to be the case over the entire twentieth century from Fig. 6.4, which shows the interannual variability in MSLP from the models compared to the instrumental record in Tahiti and Darwin. The Tahiti and Darwin MSLP values are taken from the models using the same method as explained for the reanalysis in Chapter 5. Looking at the Tahiti and Darwin regions, MPI-ESM-P and IPSL-CM5A-LR perform well over Tahiti, with little bias, whereas IPSL-CM5A-LR captures mean MSLP over Darwin best. This

suggests that IPSL-CM5A-LR represents MSLP climatology in the SOI regions closest to the reanalysis. The relatively strong biases in MSLP in HadCM3 and MIROC-ESM suggest that they may not represent the SOI regions as well as other model simulations.

Table 6.3 MSLP from Darwin and Tahiti box regions, and MSLP gradient between the two regions, 1979-2004 and 1900-1999 DJF seasonal average from instrumental record, ERA-reanalysis and historical climate model simulations.

	1979-2004			1900-1999		
	Darwin MSLP (hPa)	Tahiti MSLP (hPa)	MSLP gradient (hPa)	Darwin MSLP (hPa)	Tahiti MSLP (hPa)	MSLP gradient (hPa)
Instrumental	1007.1	1010.8	3.7	1006.7	1011.0	4.3
ERA-INT	1007.5	1011.6	4.1	---	---	---
ERA-20C	1006.8	1011.3	4.5	1007.0	1011.3	4.3
GISS-E2-R	1008.4	1010.6	2.2	1008.1	1010.5	2.4
HadCM3	1002.0	1008.1	6.1	1002.1	1007.9	5.8
IPSL-CM5A-LR	1007.7	1012.2	4.5	1007.5	1011.9	4.4
MIROC-ESM	1009.3	1016.8	7.5	1009.3	1016.9	7.6
MPI-ESM-P	1005.9	1012.1	6.2	1005.8	1012.0	6.2

The direction of the MSLP gradient between Tahiti and Darwin is replicated in all of the climate model simulations, with higher MSLP in Tahiti than Darwin, but with varying magnitudes (Table 6.3). Four of the historical simulations find a larger MSLP gradient than the ERA-Interim reanalysis and the instrumental record, ranging from 4.5 hPa (IPSL-CM5A-LR) to 7.5 hPa (MIROC-ESM). The MSLP gradient is closest to that found in ERA-Interim in IPSL-CM5A-LR, as suggested from the similarity in climatologies shown in Fig 6.3. This is also the case for the 1900-1999 analysis, with the difference between IPSL-CM5A-LR and both the instrumental and ERA-20C record only 0.1 hPa (Table 6.3). The larger pressure gradients in MIROC-ESM, HadCM3 and MPI-ESM-P suggests an over-exaggeration of the pressure differences in the tropical Pacific by the climate model simulations, GISS-E2-R has a reduced MSLP gradient compared to ERA-Interim, with the difference in DJF mean pressure

only 2.2 hPa between Darwin and Tahiti, suggesting that it might not capture the variability of SOI well.

The two-sample t-test is used to test the null hypothesis that the DJF MSLP from ERA-Interim and the climate model simulations from 1979-2004 are from datasets with equal means and equal but unknown variances in the Darwin region, and then in the Tahiti region. This tests if the regional MSLP from the climate models are statistically different from those found in the ERA-Interim reanalysis simulations. A t-test result of $H=0$, states that the null hypothesis cannot be rejected at the 95% confidence level, therefore suggesting similarity between the two datasets. Only IPSL-CM5A-LR is similar to MSLP from ERA-Interim in Darwin region (Table 6.4), whereas, the MSLP from all the models are statistically different (at 95% confidence. level) from ERA-Interim MSLP in the Tahiti region.

Table 6.4 Results from t-test comparing DJF MSLP from ERA-Interim of the datasets listed below, for Darwin and Tahiti, 1979-2004. Bold indicates where the null hypothesis can be rejected at the 95% confidence level.

	Darwin	Tahiti
GISS-E2-R	H=1 $p=6.67 \times 10^{-4}$	H=1 $p=2.2448 \times 10^{-5}$
HadCM3	H=1, $p = 1.56 \times 10^{-28}$	H=1, $p=1.932 \times 10^{-18}$
IPSL-CM5A-LR	H=0, $p=0.3469$	H=1 $p=0.0274$
MIROC-ESM	H=1 $p=4.2245 \times 10^{-11}$	H=1, $p = 3.8531 \times 10^{-29}$
MPI-ESM-P	H=1, $p = 8.80 \times 10^{-9}$	H=1, $p = 0.0414$

Therefore, this suggest that some models have difficulty in replicating the mean and variance of MSLP over Darwin and Tahiti, 1979-2004. IPSL-CM5A-LR captures the mean and variability of MSLP in Darwin and Tahiti the best out of the historical simulations, compared to ERA-Interim. Therefore it is expected to produce an SOI similar to that from ERA-Interim. However, the MSLP gradient is more important in terms of SOI representation. This was also found to be best represented in IPSL-CM5A-LR over 1979-2004 and also over the entire twentieth century (Table 6.3). This is therefore included as a metric in the ranking of SOI in Section 6.4.2.

6.4.2 The Southern Oscillation Index

The DJF SOI from the instrumental record and the reanalyses were found to be similar in Chapter 5 over the twentieth century. Disagreement between the records largest further back in time with statistically significant differences between the datasets between 1900-1950, and reduced strength of correlations between the SOI from the instrumental and reanalysis datasets in 1900-1949 compared to 1950-1999 (Table 5.3). Therefore, the instrumental record is used for comparison of statistics and number of events but the reanalysis is used when composite plotting is required, due to their complete spatial resolution. Composite plots of El Niño and La Niña from ERA-Interim are assessed as they provide the best possible representation of the spatial patterns of ENSO events, from a time with the highest number of observations and using the highest resolution temporal and spatial domains. Composite plots for the entire twentieth century can be found in Appendix C (Fig. C1; C2; C3; C4), and use ERA-20C as the reanalysis dataset to compare to the models. Fig. 6.5 shows the DJF SOI calculated from the historical climate model simulations using the methodology discussed in Chapter 5 (Section 5.3) and the instrumental DJF SOI over the twentieth century. Table 6.5 shows the Pearson's correlation coefficients between the SOI from the climate models and that from the instrumental record over various time periods of the last century.

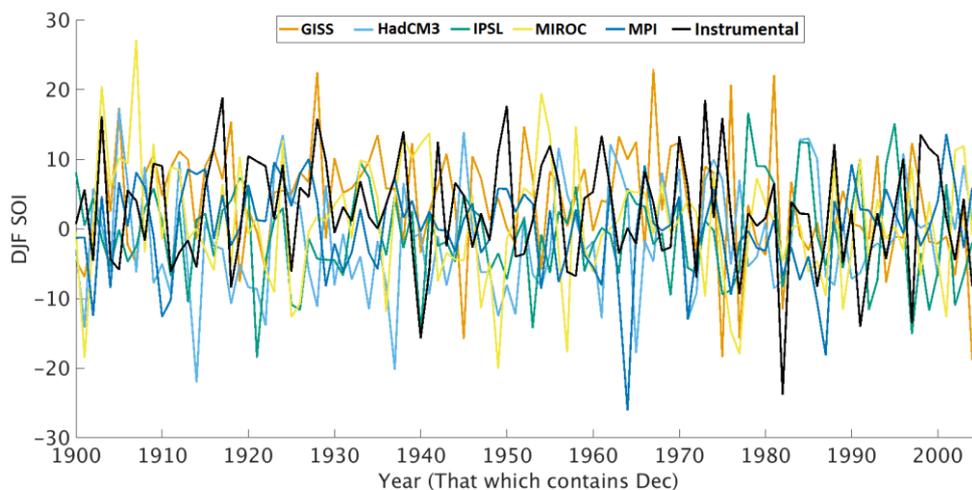


Fig. 6.5 The DJF SOI from historical climate model simulations and the DJF instrumental SOI 1900-2004 (calculated from a $1.5^\circ \times 1.5^\circ$ latitude-longitude grid), normalised over 1979-2004.

Table 6.5 Pearson’s correlation coefficient DJF SOI from the instrumental and climate model simulations over different analysis periods.

	1979-2004	1900-1949	1950-1999	1900-1999
GISS-E2-R	0.30	0.03	0.11	0.08
HadCM3	0.10	-0.03	0.04	-0.01
IPSL-CM5A-LR	0.15	0.06	0.05	0.05
MIROC-ESM	-0.02	-0.04	0.05	0.01
MPI-ESM-P	0.10	0.16	0.11	0.15

There is weak agreement between the SOI from the climate model simulations, with low correlation coefficients (ranging from $r=-0.02$ to $r=0.30$, 1979-2004 and from $r=-0.01$ to $r=0.15$ for 1900-1999) between the instrumental record and the five climate model SOI, none of which are significant at the 95% confidence level (Table 6.5). The 1982 El Niño event is captured in a number of indices, however there is less of an agreement throughout the rest of the time period. This is much weaker than the agreement found between reanalysis datasets (Fig. 5.7). Unlike the clear increase in correspondence found in the reanalysis when looking at the first and second half of the twentieth century (Fig. 5.7), there is no pattern of increased correspondence in the latter half of the twentieth century (Fig. 6.5). This is because models aren’t expected to replicate the exact years of observed events and significant correlations are a tough test for the models given this internal variability. Therefore, focus is more on the long term mean, variance and number of events, rather than the timing of events.

Table 6.6 Standard deviation of DJF SOI from the instrumental record, ERA-reanalysis and the historical climate model simulations 1979-2004 and 1900-1999.

	1979-2004	1900-1999
	SOI standard deviation	SOI standard deviation
Instrumental SOI	8.85	7.91
ERA-INT	8.84	--
ERA-20C	8.13	8.21
GISS-E2-R	7.84	7.87
HadCM3	6.62	8.08
IPSL-CM5A-LR	8.68	6.73
MIROC-ESM	6.99	8.60
MPI-ESM-P	6.44	6.31

The standard deviation is commonly used as an indicator of ENSO amplitude within models (Guilyardi et al. 2009b; Bellenger et al. 2014). The standard deviation of the DJF SOI, shown in Table 6.6, from the instrumental record, 1979-2004 is 8.85. All five of the climate model simulations produce an SOI with a standard deviation lower than that of the instrumental record over the period 1979-2004. IPSL-CM5A-LR has the standard deviation closest to that found in the instrumental record over this more recent period. The standard deviation is lower over the twentieth century in the instrumental record (7.91). However, over the twentieth century the standard deviation from some of the models is higher than the instrumental (IPSL-CM5A-LR; MIROC-ESM) and in other models it is lower (GISS-E2-R; HadCM3; MPI-EMS-P). A two-sample t-test (Appendix Table C1) shows that the null hypothesis that the datasets are statistically different cannot be rejected for 1979-2004. This is because the mean of the SOI from all datasets is zero as they have been normalised. However, statistically significant differences between the instrumental SOI and that from HadCM3, IPSL-CM5A-LR and MPI-ESM-P are found when comparing the indices over the entire twentieth century (Appendix Table C1).

Table 6.7 Number of El Niño and La Niña events defined by the instrumental, reanalysis and historical climate models using the threshold of one standard deviation from the DJF SOI to define events, 1900-1999.

	El Niño		La Niña		ENSO
	Events	Frequency (Years)	Events	Frequency (Years)	Frequency (Years)
Instrumental SOI	15	6.67	17	5.88	3.2
ERA-20C	15	6.67	13	7.69	2.8
GISS-E2-R	13	7.69	14	7.14	2.7
HadCM3	12	8.33	21	4.76	3.3
IPSL-CM5A-LR	13	7.69	14	7.14	2.7
MIROC-ESM	13	7.69	16	6.25	2.9
MPI-ESM-P	16	6.25	14	7.14	3.0

In terms of frequency of ENSO events, based on the one standard deviation threshold, some models identify less than the observed frequency from the instrumental record, while others identify more events more frequently (Table 6.7). ENSO frequency defines the frequency of an event which is either El Niño or La Niña. The frequency of ENSO events during the twentieth century within all of the climate model simulations is within the commonly defined 2-7-year frequency from the instrumental record (Allan et al. 1996), and close to the 3-year frequency suggested by the instrumental DJF SOI over this time period (Table. 6.7). This suggests that the frequency of events is well replicated from the models, with MPI-ESM-P capturing the closest to the number of El Niño events to the instrumental record during the years 1900-1999 and MIROC-ESM the closed to the observed number of La Niña events. However, the actual number and timings of events is not expected to be directly replicated within the models due to the strong role of internal variability in ENSO. Therefore, as the models all lie within the frequency range of the instrumental record then there are no obvious biases in the representation of the frequency of events over this period.

Using the events highlighted in Table 6.7, spatial composite anomaly plots show the MSLPA and SSTA associated with El Niño and La Niña events within ERA-interim

and the model simulations. Fig. 6.6 shows that during El Niño events there are low pressure anomalies over the central and eastern Pacific, and high pressure anomalies over western Pacific and Maritime Continent and Australia. The spatial pattern of MSLP anomalies varies between models, however a general pattern of positive anomalies in the west and negative anomalies in the east can be seen in the models for El Niño events. However, GISS-E2-R fails to show negative anomalies in the eastern Pacific during El Niño, and MIROC-ESM has very weak anomalies in this region. HadCM3, IPSL-CM5A-LR and MPI-ESM-P have the same MSLP anomaly pattern as ERA-Interim, within both the Darwin and Tahiti grid boxes for El Niño events. However, all the anomalies from the climate models are weaker than those in ERA-Interim. For La Niña events, MPI-ESM-P and HadCM3 capture the MSLP anomalies best. The HadCM3 magnitude is similar to that found in ERA-Interim in the central Pacific, whereas the magnitude of MSLPA over Darwin in La Niña is best captured in GISS-E2-R.

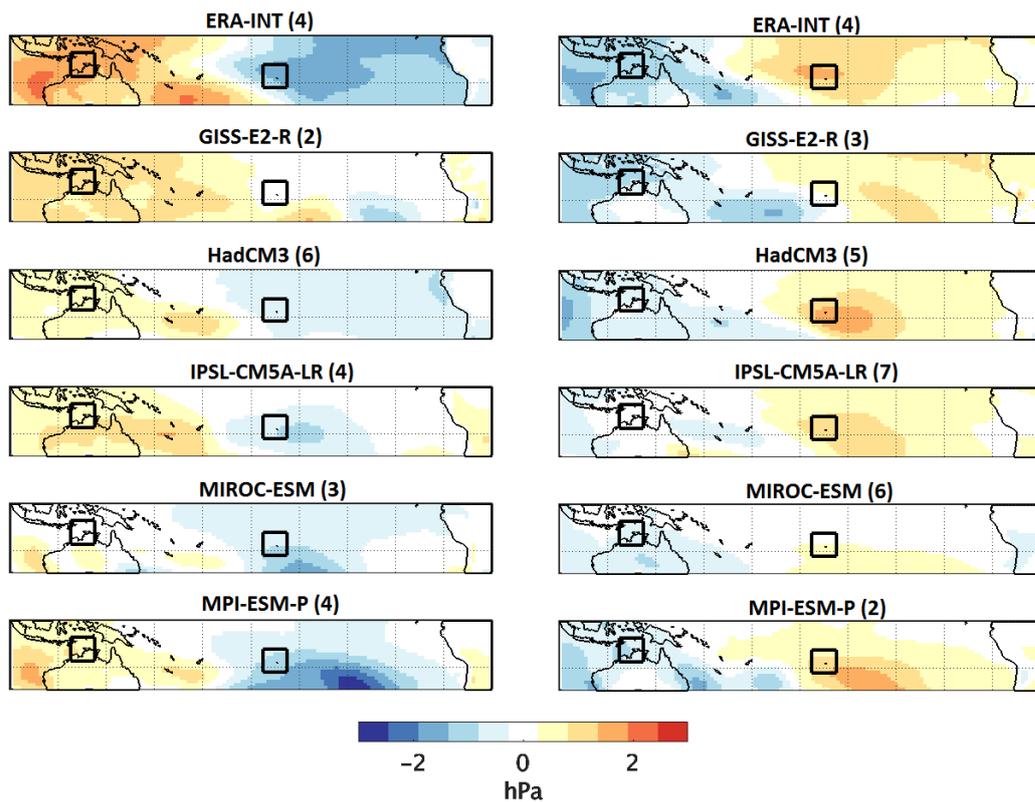


Fig. 6.6 Mean DJF MSLP composite anomaly from ERA-Interim reanalysis and climate model simulations during years with SOI one standard deviation defined El Niño (left) and La Niña events (right), 1979-2004. Anomaly calculated from 1979-2004 DJF mean. Numbers in brackets indicate number of events used to calculate each composite plot. Black rectangles indicate the western (Darwin) and eastern (Tahiti) regions used to define the SOI from the reanalysis and model simulations.

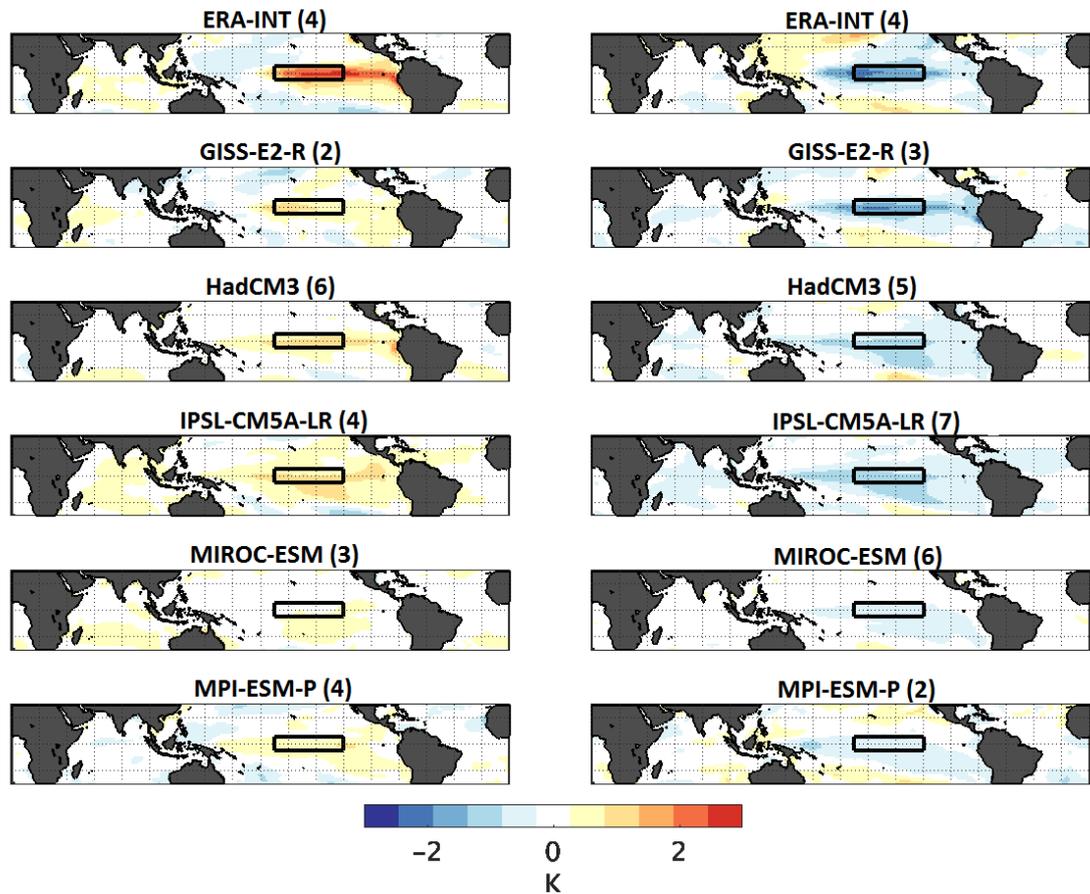


Fig. 6.7 Mean DJF surface temperature composite anomaly from ERA-Interim reanalysis and climate model simulations during years with SOI one standard deviation defined El Niño (left) and La Niña events (right), 1979-2004. Anomaly calculated from 1979-2004 DJF mean. Numbers in brackets indicate number of events used to calculate each composite plot. Black rectangle indicates Niño 3.4 region.

ERA-Interim surface temperature composites in Fig. 6.7 show that during El Niño events there are positive temperature anomalies in the central and eastern Pacific, focused in the Niño 3.4 region and extending eastwards from this region along the equator to South America. During La Niña events, negative surface temperature anomalies exist in a similar region, but extending further west than in El Niño, and not as strong in the far eastern equatorial Pacific. The spatial pattern of surface temperature anomalies varies between models but a general warming during El Niño and a cooling during La Niña is seen in all of the models over the central and eastern Pacific, but generally with a lower magnitude than the reanalysis. MIROC-ESM performs the worst over the Niño 3.4 region, and has generally weaker anomalies over

a smaller area than the reanalysis. This suggests that the relationship between pressure and surface temperature in MIROC-ESM is not as strong as other models for this region. All of the models have weaker surface temperature anomalies than ERA-Interim in both El Niño and La Niña events. GISS-E2-R captures the strength of the La Niña negative anomalies best, while IPSL-CM5A-LR, HadCM3, and GISS-E2-R captures El Niño surface temperature well in the Niño 3.4 region.

Composite plots for events over the twentieth century can be seen in Appendix Fig. C1 and Fig. C2 for MSLP and surface temperature respectively. They show good consistency with spatial patterns in Fig 6.6 and 6.7, however the strength of the anomalies is higher in the shorter period compared to the entire twentieth century, which could partly be a result of an average of a larger number of events over the longer period.

To bring together these different comparisons, the models were ranked one (best) to five (worst) on a number of metrics relating to how well they compare to the instrumental DJF SOI. The metrics include the MSLP gradient between Darwin and Tahiti, the standard deviation of the SOI and the number of El Niño and La Niña events. The models were ranked in comparison to the instrumental DJF SOI for each of the metrics and then the sum of the metrics ranked in order to give an overall rank for SOI parameters. The metrics sample across a range of characteristics spanning mean, variance and frequency, thus providing more than just a comparison of the means. From the SOI rankings 1979-2004 (Appendix Table C2), IPSL-CM5A-LR is ranked as the best of the five models over this period, as suggested at the end of Section 6.4.1 which explained that it captured the MSLP and MSLP gradient well. GISS-E2-R is ranked second despite its relatively low MSLP gradient, 1979-2004. The models were then ranked over the entire twentieth century compared to the instrumental record (Table 6.8).

Table 6.8 Ranking of historical model simulations for SOI-based metrics compared to the DJF SOI instrumental record (BoM), (1900-1999).

	MSLP gradient rank	SOI standard deviation rank	El Niño frequency rank	La Niña frequency rank	Overall ranking
GISS-E2-R	=3	1	=2	=2	1
HadCM3	2	2	5	5	5
IPSL-CM5A-LR	1	4	=2	=2	2
MIROC-ESM	5	3	=2	1	=3
MPI-ESM-P	=3	5	1	=2	=3

These two models also perform well when assessing performance of the models over the entire twentieth century, with GISS-E2-R first and IPSL-CM5A-LR second (Table 6.8). When comparing to ERA-20C, IPSL-CM5A-LR and GISS-E2-R rank joint first over the twentieth century (Appendix Table C3). Therefore, they both perform well with respect to the instrumental record but also to the extended reanalysis. As they both rank in the top two during both periods, they are used in Chapter 8 for SOI analysis over the Last Millennium. The surface temperature climatology and the Niño 3.4 were assessed within the models to identify which models perform better or worse when using temperature-based index of ENSO instead of the atmospheric SOI.

6.4.3 Surface temperature climatology

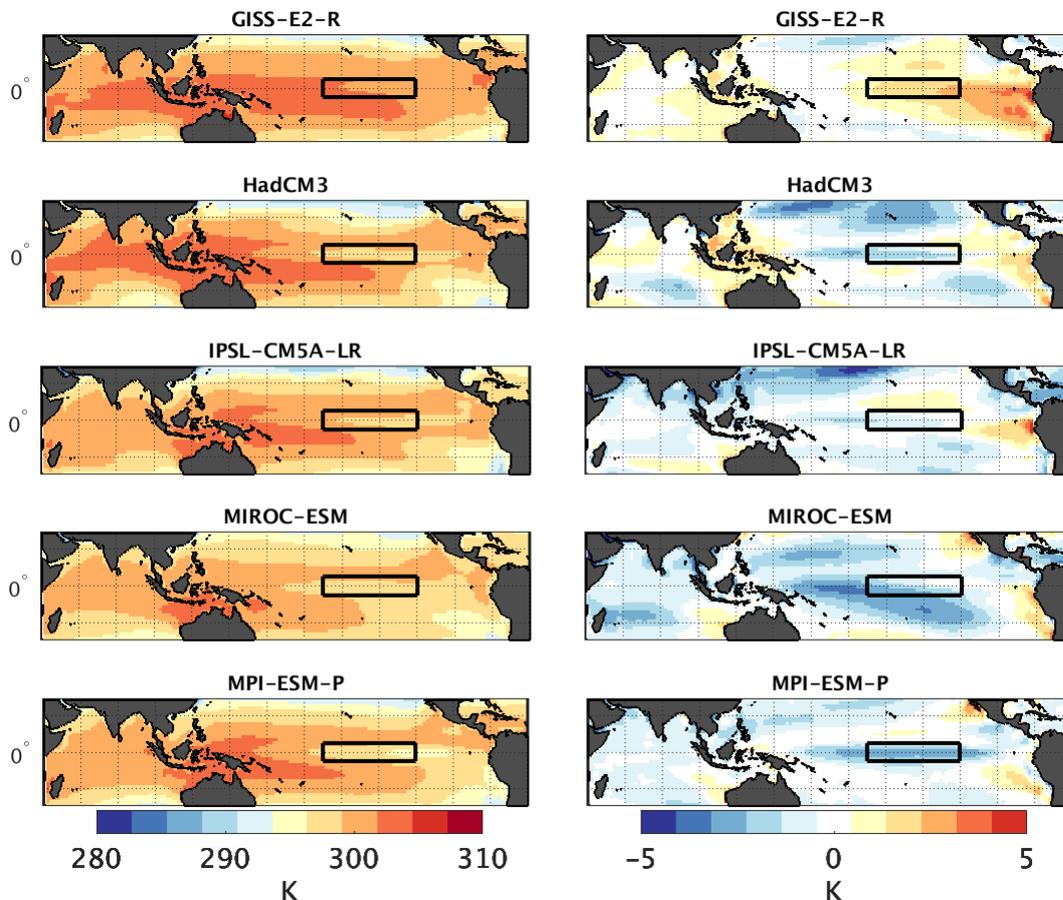


Fig. 6.8 DJF surface temperature climatology, 1979-2004 from historical climate model simulations (left) and Model climatology minus ERA-Interim reanalysis (right). Black rectangle indicates the Niño 3.4 region.

The models' surface temperature climatology and bias compared to ERA-Interim was then assessed in Fig. 6.8. GISS-E2-R has a warm bias over much of the central and eastern Pacific, failing to capture the cold equatorial tongue. Therefore, there are positive SSTAs in the Niño 3.4 region. In the other four models there are negative SSTA in the Niño 3.4 region. This is due to the cold tongue bias, commonly found in climate models, in which cold SSTs extended too far westwards along the tropical Pacific, as discussed in Section 6.1.2. The negative SSTAs are most widespread in MIROC-ESM, extending over a large portion of the southern tropical Pacific. HadCM3 has a cold bias across the central equatorial Pacific and SST anomalies which are too warm over the Maritime Continent (Turner et al. 2005). GISS-E2-R also simulates positive SSTA over the Maritime Continent. All of the models have the

warm bias in the far eastern Pacific identified as a common bias seen here in Fig. 6.2 (Michael et al. 2013), but this is most pronounced in GISS-E2-R, and is relatively strong in IPSL-CM5A-LR. However, the surface temperature bias in the Niño 3.4 region is lowest in IPSL-CM5A-LR, as this warm bias doesn't extend into this region in this model, therefore for the Niño 3.4 climatology IPSL-CM5A-LR agrees best with ERA-Interim. MPI-ESM-P is known to have a cold tongue that extends too far west (Jungclaus et al. 2013), and this is also seen here in Fig. 6.8.

6.4.4 The Niño 3.4 index

The Niño 3.4 index from the instrumental record and the reanalyses were found to be similar in Chapter 5 (Fig. 5.9) and therefore the instrumental record is used for the basic statistics but reanalysis used when composite plotting is required, due to their complete spatial resolution. The Niño 3.4 surface temperatures extracted from the models are compared to ERSST, 1900-2004 (Fig. 6.9). The positive bias in GISS-E2-R is clear in this region, and MPI-ESM-P shows the strongest negative bias. There is a lack of agreement between the Niño 3.4 indices calculated from the models when compared to each other, as well as when compared to the instrumental record, and no clear agreement on recent strong ENSO events such as 1982/3 or 1997/8. The Niño 3.4 temperature anomalies were calculated using the 1979-2004 base period, as was the case for the SOI, and the temperature anomaly indices from the models can be found in Appendix C (Fig. C5), and are used for index analysis and metrics. The Pearson's correlation coefficient comparing the models to ERSST highlight this poor agreement between the datasets over the twentieth century (Table 6.9). There is a strong negative correlation between ERSST and GISS-E2-R ($r = -0.89$) over 1950-1999 and moderately negative correlation between ERSST and MIROC-ESM over this period, suggesting some clear disagreement between the records. As discussed with respect to the SOI, it is beyond the scope of models to capture ENSO events in the same year as the instrumental record due to internal variability (Hope et al. 2016).

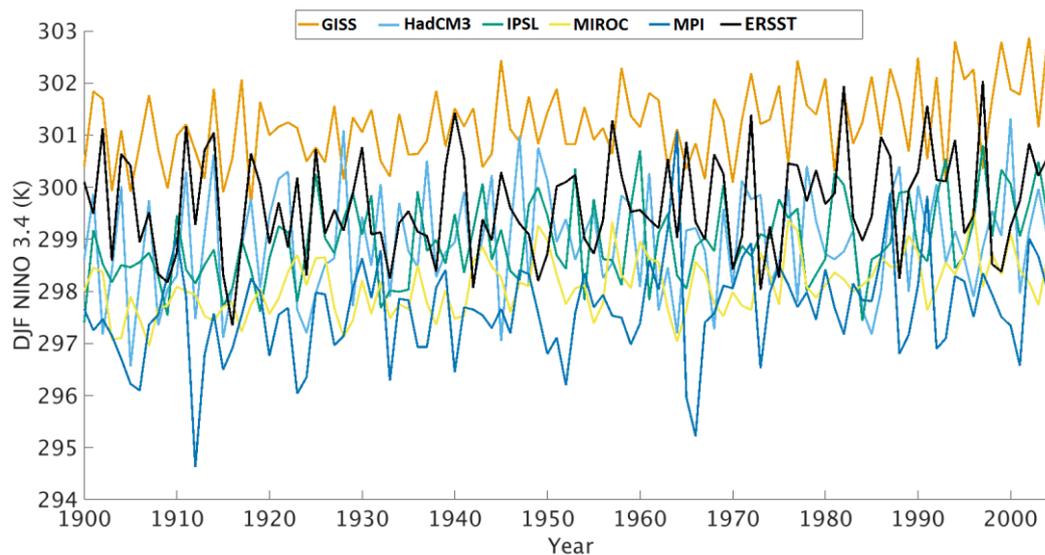


Fig. 6.9 The DJF Niño 3.4 from the historical climate model simulations and ERSST, 1900-2004 (Anomalies shown in Appendix Fig. C5).

Table 6.9 Pearson's correlation coefficient DJF Niño 3.4 from the ERSST and climate model simulations over different analysis periods, bold indicates significance the 95% confidence level.

	1979-2004	1900-1949	1950-1999	1900-1999
GISS-E2-R	-0.18	0.22	-0.84	0.10
HadCM3	-0.06	0.14	-0.03	0.07
IPSL-CM5A-LR	0.11	0.06	0.06	0.09
MIROC-ESM	-0.32	-0.04	-0.34	0.02
MPI-ESM-P	0.41	0.07	0.15	0.14

Table 6.10 Reanalysis and modelled DJF Niño 3.4 mean surface temperature and standard deviation, DJF 1979-2004 and 1900-1999.

	1979-2004		1900-1999	
	Niño 3.4 standard deviation (K)	Niño 3.4 mean (K)	Niño 3.4 standard deviation (K)	Niño 3.4 mean (K)
ERSSTv4	1.01	300.02	0.97	299.64
ERA-20C	1.13	299.60	0.98	299.44
GISS-E2-R	0.84	301.66	0.71	301.16
HadCM3	0.93	299.00	1.02	298.90
IPSL-CM5A-LR	0.83	299.35	0.78	298.91
MIROC-ESM	0.44	298.40	0.55	298.09
MPI-ESM-P	0.84	297.94	0.92	297.62

Over the twentieth century, all five of the climate model simulations produce a Niño 3.4 index with a standard deviation lower than that of ERSST, except HadCM3 which has the highest standard deviation of the datasets (Table 6.10) while MIROC-ESM has the lowest standard deviation of the climate models. In terms of the frequency of ENSO events over the period 1979-2004, MPI-ESM-P captures the same number of El Niño and La Niña events as ERSST while the other climate models identify a varied number of ENSO events (Table 6.11). However, when looking over the entire twentieth century MPI-ESM-P has lowest frequency of both El Niño and La Niña events, underestimating the number of observed events from ERSST. IPSL-CM5A-LR captures the frequency of both well.

There is a difference in the frequency of events defined using the Niño 3.4 index compared to the DJF SOI over the twentieth century when comparing both the instrumental records and the climate model simulations (Table 6.7; Table 6.11). All but MPI-ESM-P identify a higher number of El Niño events when using the Niño 3.4 index compared to the DJF SOI, and the overall ENSO frequencies suggest more ENSO events are found when using the Niño 3.4 index compared to the SOI, except for in MPI-ESM-P.

A two-sample t-test shows that the null hypothesis that the datasets are statistically different cannot be rejected at the 95% confidence level for all of the climate models, 1979-2004, and for all but HadCM3 1900-1999 (Appendix Table C4). This suggests that the differences in the mean and variance of Niño 3.4 from the climate models are not statistically significant, apart from in HadCM3, which was found to have the highest standard deviation of the climate models, suggesting the variance of Niño 3.4 surface temperature anomaly within this model simulation is over-represented compared to the observed record.

Table 6.11 Number of El Niño and La Niña events defined by the instrumental, reanalysis and historical climate models using the threshold of one standard deviation from the Niño 3.4 anomaly (1979-2004 base period) to define events, 1900-1999.

	El Niño		La Niña		ENSO
	Events	Frequency (Years)	Events	Frequency (Years)	Frequency (Years)
ERSSTV4	18	5.56	16	6.25	3.4
ERA-20	18	5.56	15	6.67	3.3
GISS-E2-R	17	5.88	17	5.88	3.4
HadCM3	19	5.26	15	6.67	3.4
IPSL-CM5A-LR	19	5.26	16	6.25	3.5
MIROC-ESM	16	6.25	16	6.25	3.2
MPI-ESM-P	11	9.09	12	8.33	2.3

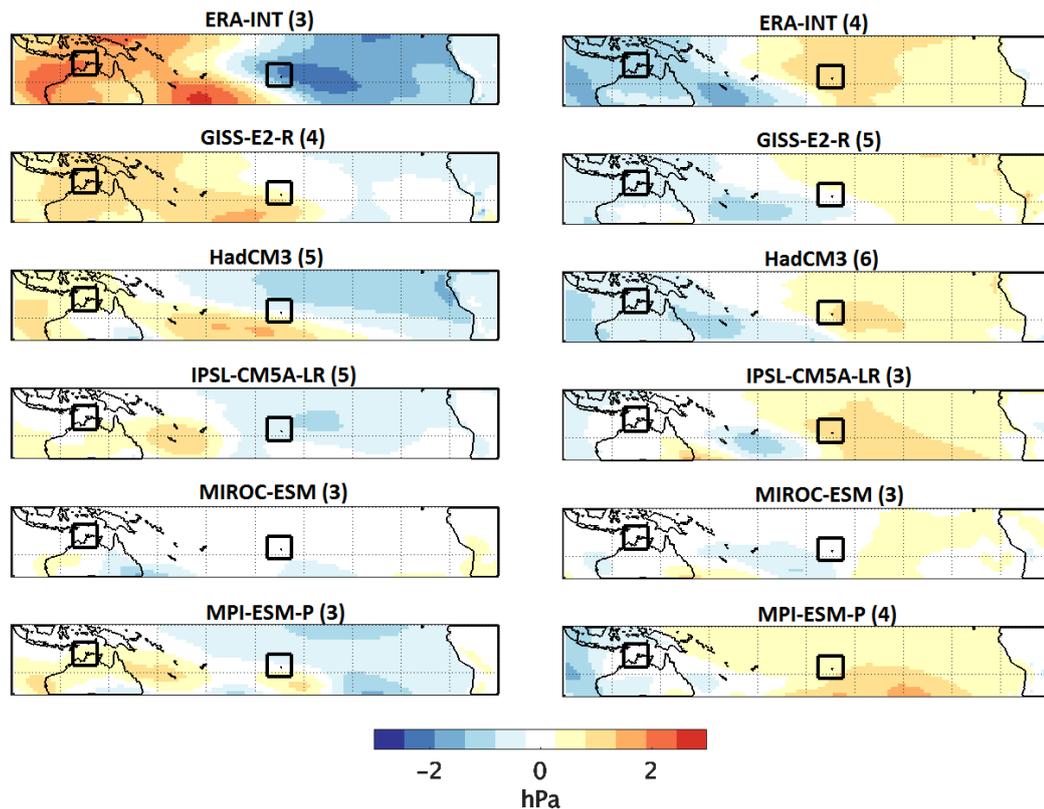


Fig. 6.10 Mean DJF MSLP composite anomaly from ERA-Interim reanalysis and climate model simulations during El Niño (left) and La Niña events (right) years defined by one standard deviation of the Niño 3.4 index, 1979-2004. Anomaly calculated from 1979-2004 DJF mean. Numbers in brackets indicate number of events used to calculate each composite plot. Black rectangles indicate the western (Darwin) and eastern (Tahiti) regions used to define the SOI from the reanalysis and model simulations.

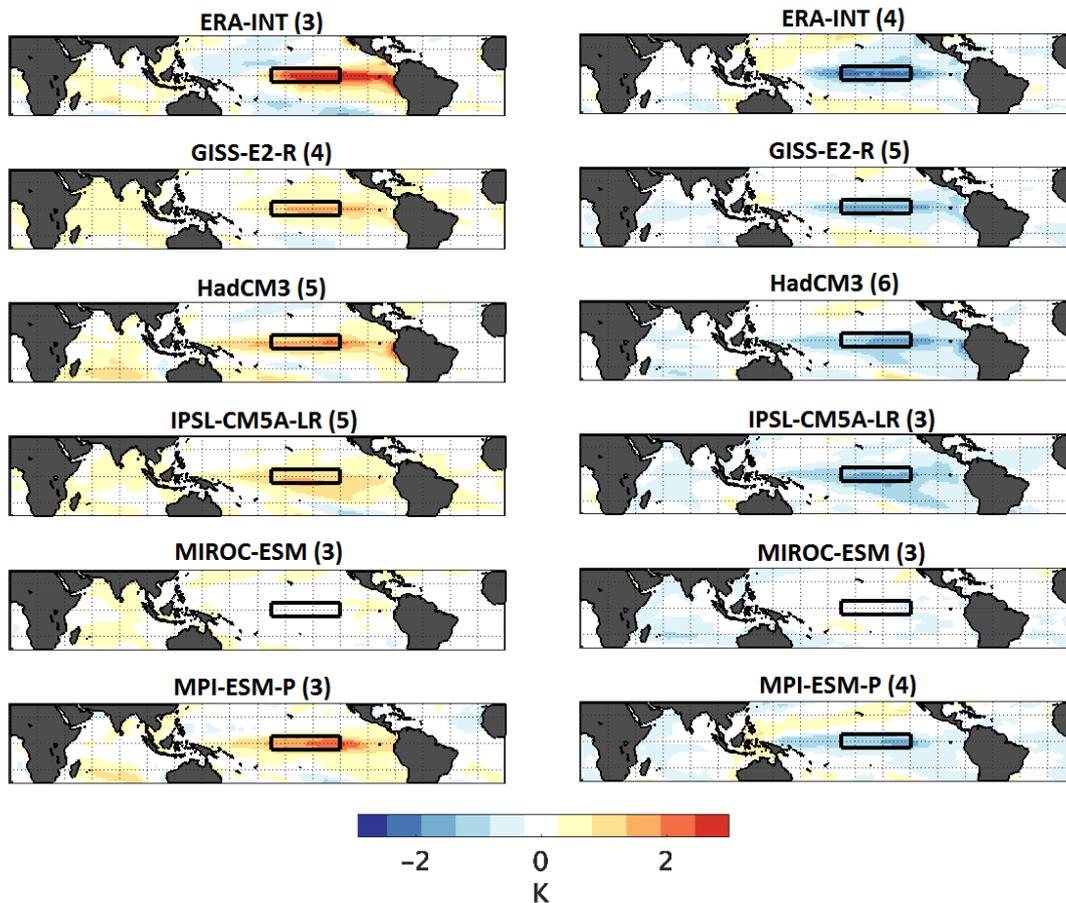


Fig. 6.11 Mean DJF surface temperature anomaly from ERA-Interim reanalysis and climate model simulations during El Niño (left) and La Niña events (right) years defined by one standard deviation of the Niño 3.4 index, 1979-2004. Anomaly calculated from 1979-2004 DJF mean. Numbers in brackets indicate number of events used to calculate each composite plot. Black rectangle indicates Niño 3.4 region.

The spatial composite calculated from the Niño 3.4-defined events (1979-2004) from ERA-Interim and the climate model simulations can be seen in Fig. 6.10 and 6.11. As the composites are calculated based on Niño 3.4 it is expected that anomalies in the Niño 3.4 region would be high, with positive anomalies during El Niño and negative anomalies during La Niña. As seen in the SOI surface temperature composites, the positive surface temperature anomaly spans the Niño 3.4 region and eastward to the coast of South America during El Niño, whereas the negative anomalies in La Niña are shifted further west into the central equatorial Pacific (Fig. 6.11). All the models except MIROC-ESM show similar anomalies within the Niño 3.4 region. MIROC-ESM anomalies are too weak. This model was also found to be the weakest in the SOI

surface temperature anomalies. During El Niño, the positive temperature anomalies HadCM3, IPSL-CM5A-LR and MPI-ESM-P extend too far west compared to ERA-Interim. The models were then ranked compared to the instrumental record, as carried out for the SOI in Table 6.8.

Table 6.12 Ranking of historical model simulations for Niño 3.4 metrics compared to ERRST (1900-1999).

	Niño 3.4 rank	Niño 3.4 standard deviation rank	El Niño frequency rank	La Niña frequency rank	Overall ranking
GISS-E2-R	3	4	=1	=3	3
HadCM3	2	=1	=1	=3	2
IPSL-CM5A-LR	1	3	=1	=1	1
MIROC-ESM	4	5	4	=1	4
MPI-ESM-P	5	=1	5	5	5

A similar ranking as for the SOI was undertaken. Overall, MPI-ESM-P performs the worst at capturing ENSO when using the Niño 3.4 index over the twentieth century (Table 6.12). Whereas, it ranked best over the later period (1979-2004 - Appendix Table C5) it struggles to capture the number of events in the earlier period. GISS-E2-R captures the spatial pattern of temperature anomalies well, but has a clear positive temperature bias in the Niño 3.4 region (Fig. 6.8). IPSL-CM5A-LR continues to rank well, as was seen with SOI analysis, and overall the ranking over the twentieth century compared to the ERRST (Table 6.12) suggests IPSL-CM5A-LR performs best at capturing the metrics of the Niño 3.4 index assessed, while HadCM3 ranks second and GISS-E2-R ranks third. HadCM3 ranked worst for the SOI, whereas IPSL-CM5A-LR and GISS-E2-R were the top two for the SOI (Table 6.8), therefore these two models rank highly for both indices.

6.4.5 ENSO diversity in the models over the twentieth century

Following on from the analysis of El Niño and La Niña within climate models, focus is now given to ENSO diversity. As discussed in Chapters 1 and 4, El Niño events can

have different characteristics and are commonly divided into Eastern Pacific (EP) and Central Pacific (CP) events. The defining difference between these two types of events is the location of SST anomalies in the equatorial Pacific, however there are also known differences in their regional teleconnections (see Chapter 4). The differences in MSLP and surface temperature for EP and CP El Niño events during boreal winter are assessed here within reanalysis data and the historical climate model simulations.

As in the above analysis, five historical climate model simulations are used here along with ERA-Interim and ERA-20C reanalysis data. Additionally, a monthly index of Niño 3 (5°N-5°S 150°W-120°W) and Niño 4 (5°N-5°S 160°E-150°W) from NOAA is used as the ‘instrumental’ record, spanning 1870 to present. These indices were calculated from HadISST and are available to download from National Oceanic and Atmospheric Administration (http://www.esrl.noaa.gov/psd/gcos_wgsp/Timeseries/Nino3/;http://www.esrl.noaa.gov/psd/gcos_wgsp/Timeseries/Nino4/). The HadISST ‘instrumental record’ defines three EP Niño events and only one CP El Niño event during 1979-2004. Therefore, due to this limited number of events, ENSO diversity and composite plots focused the entire twentieth century, 1900-1999 with the use of ERA-20C. Comparing the models over the entire twentieth century provides a higher number of EP and CP events to be analysed and used in the composites.

Yeh et al. (2009) compared the frequency of EP and CP events over the historical period (1854-2007) and found an increase in the frequency of CP events in the latter decades of this period. They used the Niño 3 and the Niño 4 SST index from ERSST, and in addition compared results from HadISST (1870-2007) and Kaplan SST (1856-2007) and found similar conclusions from the various records. A decrease in correlation between Niño 3 and Niño 4 in recent decades was found suggesting increased independence of the two indices (Yeh et al. 2009). They found an occurrence ratio of CP to EP El Niño to be 0.2 from the observations, much smaller than from the models they investigated, which had a mean ratio of 0.97, thus suggesting that CGCMs tend to simulate CP El Niño events more frequently than they are observed. This is likely due to the common cold bias over the cold tongue region which means El Niño warm anomalies are shifted further west in the models. Yeh et al. (2009) also found

only subtle differences between raw and detrended data from Niño 3 and Niño 4 and present results from both.

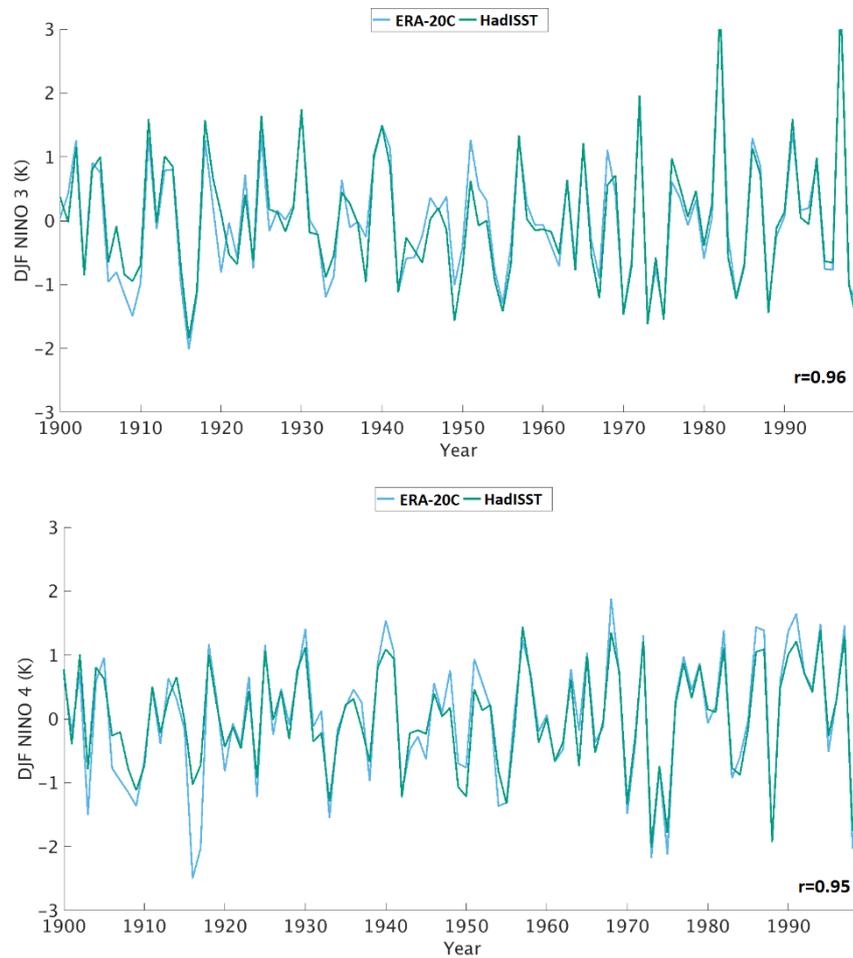


Fig. 6.12 The DJF Niño 3 (top) and Niño 4 (bottom) from HadISST and ERA-20C reanalysis, 1900-1999 ($1.5^{\circ} \times 1.5^{\circ}$ grid), Pearson's correlation (r) between the indices from the two datasets is shown.

There is a close correspondence between the two indices and clear interannual variability of the Niño 3 and Niño 4 indices (Fig. 6.12). Interannual variability is a common feature of ENSO indices as seen from the instrumental SOI (Fig. 6.5) and the Niño 3.4 index (Fig. 6.9). Data are normalised over the 1900-1999 base period, and ERA-20C used as the reanalysis for comparison to HadISST. The Pearson's correlation coefficient between the Niño 3 index from HadISST and ERA-20C, 1900-1999, is 0.96, and is 0.95 for the Niño 4 index. Some years are more extreme in one index than in others (Fig 6.12). The 1982/83 and 1997/98 El Niño events are represented by the two biggest peaks in the Niño 3 index, however the magnitude of

the Niño 4 anomaly in these same years is not as large. Both of these events are well known EP events (Wang et al. 2012), thus explaining why the temperature anomaly is largest in the Niño 3 index, which covers a region further east than the Niño 4 index, as seen in Fig. 1.6. Positive anomalies are generally higher in the Niño 3 index whereas the negative anomalies are generally of higher magnitude in the Niño 4 index. This links with findings above, in which negative SSTA during La Niña events are further west than the positive SSTA during El Niño events.

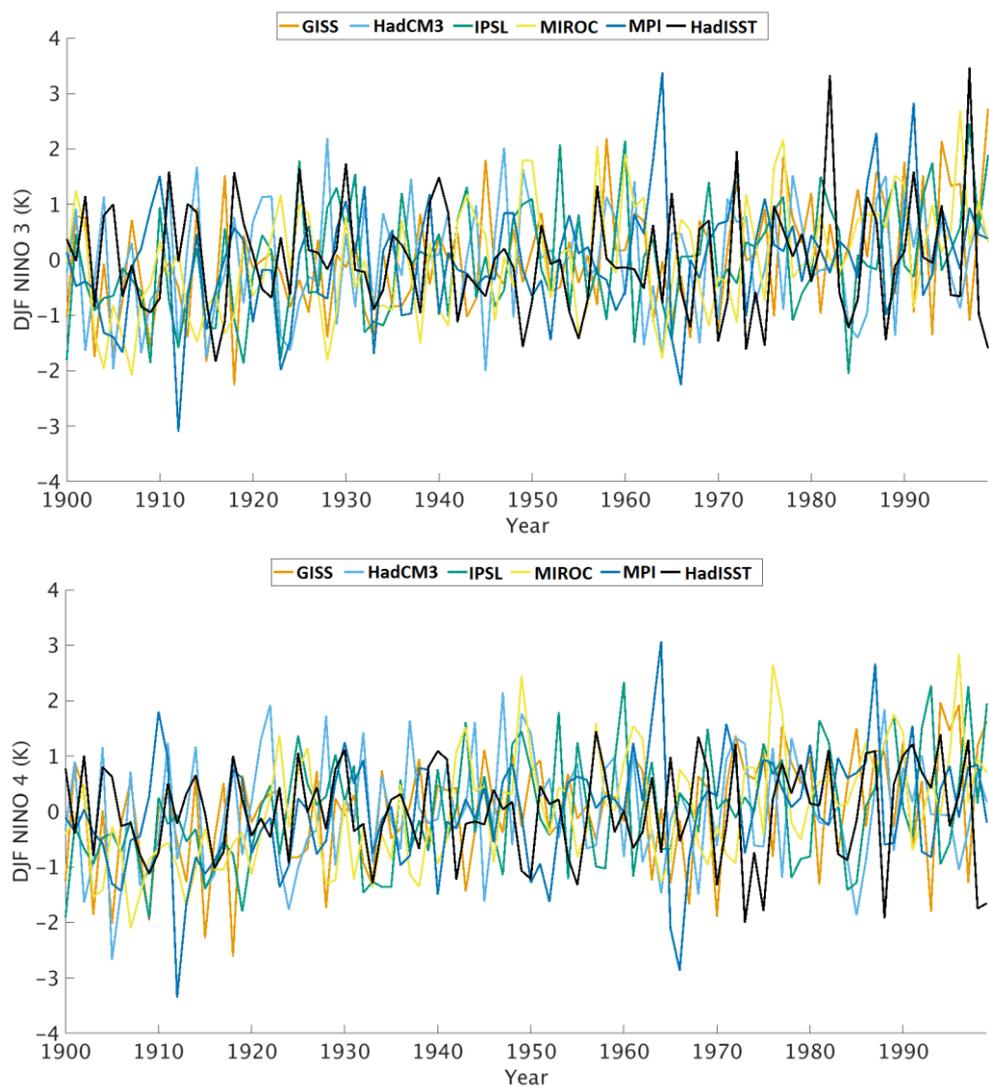


Fig. 6.13 The DJF Niño 3 index (top) and Niño 4 index (bottom) from HadISST and climate model historical simulations, 1900 -1999 (1.5° x 1.5° grid).

The Niño 3 and Niño 4 indices were then extracted from the five climate model simulations (Fig. 6.13). Interannual variability is again a key feature of the indices.

The Niño 3 mean temperature is higher than Niño 4 in all of the datasets, as expected due to the Pacific mean state as seen in Fig 6.8. The standard deviation in Niño 3 region is smaller in all of the model simulations compared to the reanalysis. This is also generally the case for the Niño 4 region, however HadCM3 and IPSL-CM5A-LR have standard deviations similar to the reanalysis. In the Niño 4 region there is a larger difference in the standard deviation in the ERA-20C and HadISST datasets, than there is in the Niño 3 region.

From these indices, the number of EP and CP El Niño events, and a CP-EP ratio, were found (Table 6.13) based on the one standard deviation threshold. Xu et al. (2017) used a number of historical runs (1950-99) of the CMIP5 models and calculated the CP-EP El Niño ratio. There is no overlap between the models used by Xu et al. (2017) and the models used here, therefore the CP-EP ratios cannot be directly compared on a model-by-model basis. However, the CP-EP ratio from the 17 CMIP5 models used by Xu et al. (2017) ranged from 0.78 to 1.8. This covers a similar range to that found using the models within this chapter, in which the CP-EP ratio ranges from 0.64 to 1.75 for the entire twentieth century (1900-1999). HadCM3, IPSL-CM5A-LR and MIROC-ESM all have a higher CP-EP ratio than the HadISST instrumental record, suggesting more CP events than observed. This suggests that within the models strong CP events are identified more than strong EP events. This supports the findings of Kim and Yu (2013) who suggest that CMIP5 models find it more difficult to reproduce the magnitude of strong EP events compared to strong CP events. The one standard deviation threshold is at the higher end of suggested thresholds for defining ENSO events, and therefore focuses on medium to strong events. One reason for models overestimating the number of CP events is the cold SST bias in the eastern Pacific within the models (Fig. 6.8) which results in the warm SSTA shifted further west across the equatorial Pacific into the Niño 4 region. This was noted by Yeh et al. (2009). GISS-E2-R has a lower CP-EP ratio than observed but similar to that from ERA-20C; Fig. 6.8 showed that this model did not have the common cold tongue bias. However, MPI-ESM-P also has a lower CP-EP ratio than observed, but did contain the cold tongue bias. Therefore, there must be additional factors influencing the CP-EP ratio within the climate models.

Table 6.13 Number of EP and CP El Niño events from reanalysis and historical climate models, 1900-1999.

	EP Niño events	CP El Niño events	CP-EP El Niño ratio (CP ÷ EP)
HadISST	12	10	0.83
ERA-20C	12	7	0.58
GISS-E2-R	11	7	0.64
HadCM3	10	13	1.3
IPSL-CM5A-LR	8	14	1.75
MIROC-ESM	8	14	1.75
MPI-ESM-P	8	6	0.75

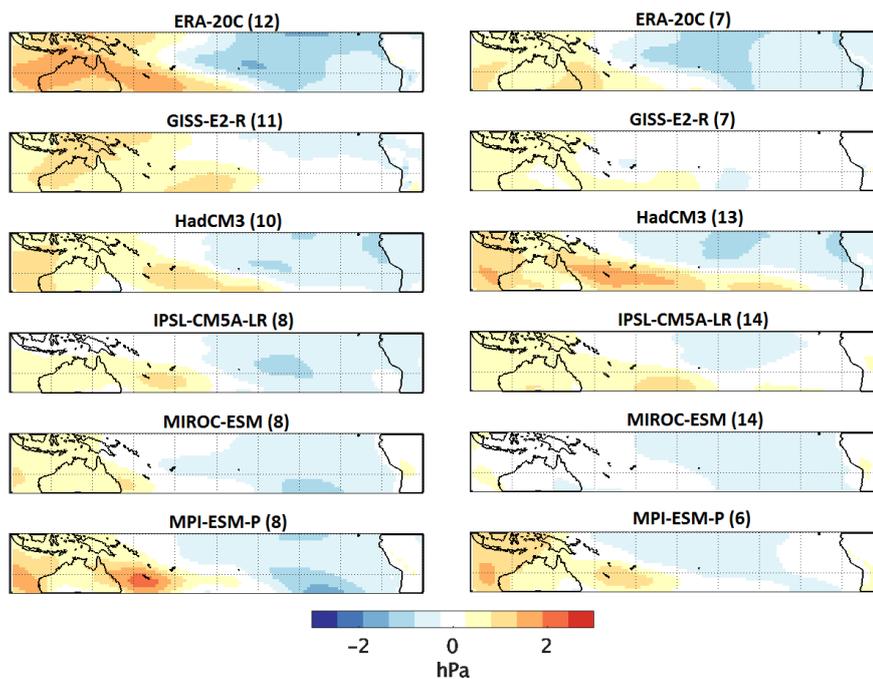


Fig. 6.14 Mean DJF MSLP anomaly from ERA-20C reanalysis and climate model simulations during EP (left) and CP El Niño events (right), 1900-1999. Anomalies based on 1900-1999 DJF mean. Numbers in brackets indicate number of events used to calculate each composite plot.

Composite anomaly plots show the spatial pattern of MSLP (Fig. 6.14) and surface temperature (Fig. 6.15) during EP and CP events from the reanalysis and historical

climate model simulations, 1900-1999. In ERA-20C reanalysis during EP El Niño, negative MSLP anomalies are found in the central and eastern Pacific, and positive anomalies in the western Pacific and Maritime Continent (Fig. 6.14). The MSLP anomalies are of smaller magnitude during CP events, and the centre of the negative MSLP anomalies are focused further westward into the central Pacific as expected. The EP MSLP anomalies are weaker in the climate model simulation composites compared to ERA-20C. Generally weaker anomalies within the models compared to the reanalysis were also found when looking at El Niño and La Niña events using the SOI and Niño 3.4 index (Fig. 6.6 and Fig.6.10)

During EP events, all the models identify the general spatial pattern of positive MSLPA in the west and negative MSLPA in the central and eastern Pacific. GISS-E2-R captures the positive MSLP anomalies over the Maritime Continent and into the western Pacific well during EP events, but fails to capture the negative anomalies over the central and eastern Pacific. The MSLP anomalies during CP events are generally weaker than those of EP events in the climate model simulations. Again, GISS-E2-R fails to capture the negative MSLPA in the central Pacific. MIROC-ESM does not show positive anomalies during CP events in the western Pacific, Australia and the Maritime Continent. HadCM3 has an eastward extension of positive MSLP south of the equator to a larger extent than that seen in ERA-20C. MPI-ESM-P and IPSL-CM5A-LR capture the CP MSLP patterns best, but with reduced magnitude compared to ERA-20C in the central Pacific.

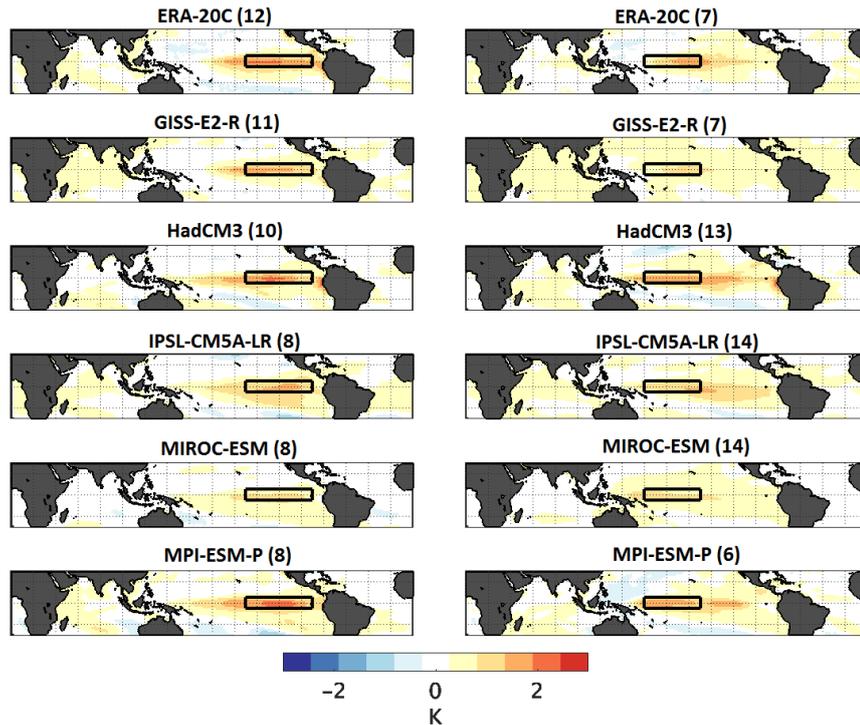


Fig. 6.15 Mean DJF surface temperature anomaly from ERA-Interim and climate model simulations during EP (left) and CP El Niño events (right), 1900-1999. Anomalies based on 1900-1999 DJF mean. Black rectangle indicates Niño 3 (left) and Niño 4 (right) regions. Numbers in brackets indicate number of events used to calculate each composite plot.

The surface temperature composite anomalies from ERA-20C are stronger during EP events compared to CP events and extend further east than in the CP events as expected (Fig. 6.15). This is the case for all of the model simulations. The anomalies in the model simulations are generally weaker than those from the reanalysis. HadCM3 and MPI-ESM-P temperature anomalies extend further east in the CP composites compared to the reanalysis. Again MIROC-ESM has the weakest anomalies across the region for both EP and CP events. Difficulties in the representation of surface temperature patterns across the Pacific could be related to the cold tongue bias found within a number of climate models. GISS-E2-R ranks top for ENSO diversity (Table 6.14), capturing the number of events well, although slightly lower than the number observed. IPSL-CM5A-LR ranks second for ENSO diversity.

Table 6.14 Ranking of historical models compared to Niño 3 and Niño 4 metrics from HadISST (1900-1999).

	Niño 3	Niño 4	EP	CP	Overall
	standard	standard	events	events	Ranking
	deviation	deviation	rank	rank	
	rank	rank			
GISS-E2-R	4	1	1	=1	1
HadCM3	3	5	2	=1	=3
IPSL-CM5A-LR	2	2	=3	=3	2
MIROC-ESM	5	3	=3	=3	5
MPI-ESM-P	1	4	=3	=3	=3

6.5 Summary of historical simulation results

Overall the models have been found to perform with a varying degree of correspondence to the reanalysis with respect to climatology and for ENSO events. MIROC-ESM is found to be the weakest at representing MSLP and surface temperature patterns associated with both typical ENSO (El Niño and La Niña) and with ENSO diversity (EP and CP El Niño events). A strong positive MSLP bias exists over much of the equatorial Pacific (Fig. 6.3) resulting in an over-exaggeration of the MSLP gradient between Tahiti and Darwin (Table 6.3). It ranks third in the SOI comparison (Table 6.8). The surface temperature climatology shows a negative bias over the western and central equatorial Pacific (Fig. 6.8), this results in a mean Niño 3.4 index 1.55 K lower than the instrumental record, 1900-1999 (Table 6.10). The standard deviation of Niño 3.4 over this region is 0.55 K, which is almost half of that of the instrumental record (0.97 K). Therefore, MIROC-ESM has weaker variability than observed record over this period. When looking at EP and CP events, composite plots show MIROC-ESM has the weakest temperature anomalies of the reanalysis data (Fig. 6.15). The ranking methods applied to the historical climate models ranked MIROC-ESM last in terms of ENSO diversity (Table 6.14). This model was ranked 4th for the Niño 3.4 index and 3rd for the SOI, suggesting an improvement in performance when using the atmospheric index over surface temperature-based

indices. This suggests that MIROC-ESM has biases in its MSLP and surface temperature climatology, resulting in issues in the representation of ENSO and ENSO diversity within its historical simulation. In addition, previous studies have excluded MIROC-ESM from their analysis over the LM as its simulations over the last 1000 years are known to display strong drift (Lewis and LeGrande 2015; Brown et al. 2016).

HadCM3 was found to have a negative MSLP bias compared to ERA-Interim over the entire study region, with the strongest bias found over the land masses (Fig. 6.3). The MSLP composites showed that it captured the general spatial anomaly patterns relatively well for El Niño and La Niña defined by the SOI (Fig 6.6), and to a slightly less extent when defining events using the Niño 3.4 index (Fig. 6.10). The surface temperature anomalies extend too far west during El Niño and La Niña events over the twentieth century, spanning the entire equatorial Pacific (Fig 6.7; Fig 6.11). HadCM3 was found to rank worst out of the models when comparing to the instrumental SOI (Table 6.8), however it is second best when comparing to Niño 3.4 from ERRST over the entire twentieth century (Table. 6.12). In terms of ENSO diversity, HadCM3 was ranked third (Table 6.14). Therefore, it captures ENSO amplitude and frequency better when using the temperature-based index compared to the SOI in the twentieth century.

MPI-ESM-P captured the climatological MSLP over Tahiti and Darwin well and produces an SOI which ranks joint third over the twentieth century (Fig. 6.3; Table 6.8). MPI-ESM-P ranked top when comparing models to ERSST Niño 3.4 metrics over the shorter time period 1979-2004 (Appendix Table C5), however when this was extended over the entire twentieth century it ranked worst (Table 6.12). This was due to identifying the lowest frequency of El Niño and La Niña events over this period, a frequency which was furthest from that found using the ERRST record. This shows that differences in the results of some of the models can change strongly when comparing over the two different periods. In terms of ENSO diversity, MPI-ESM-P ranked third. Therefore, it ranks mid to bottom of the models for the SOI, the Niño 3.4 and ENSO diversity metrics.

Within GISS-E2-R, the MSLP bias is such that there is a reduced MSLP gradient over Tahiti to Darwin, with positive bias in the west and negative bias in the central and

eastern regions (Fig. 6.3). The MSLP gradient in GISS-E2-R is the smallest of the models and reanalysis (Table 6.3). Despite this, it produces an SOI which has a standard deviation which is most similar to the instrumental record, and identifies the frequency of ENSO events relatively similarly to that identified from the instrumental DJF SOI over 1900-1999 (Table 6.8). As a result, GISS-E2-R ranks highest of all the models for the SOI metrics over the twentieth century (Table 6.8). GISS-E2-R was also found to have bias in its surface temperature climatology compared to ERA-Interim (Fig. 6.8), being the only one of the models to have a positive bias over a large extent of the eastern equatorial Pacific. However, it still captures the surface temperature anomalies relating to El Niño and La Niña events well (Fig. 6.11). It ranks third over the twentieth century for Niño 3.4 metrics (Table 6.12). GISS-E2-R captures the surface temperature patterns associated with EP El Niño events and La Niña well, but with a lower magnitude than ERA-20. It ranks top of the models for ENSO diversity capturing the number of EP and CP El Niño events closest to that from the HadISST-based record of ENSO diversity (Table 6.14). Therefore, GISS-E2-R performs well over multiple rankings of the models.

IPSL-CM5A-LR has the lowest MSLP bias compared to ERA-Interim, capturing MSLP closest to climatology in Tahiti and Darwin (Fig. 6.3), thus having a MSLP gradient which is closest to the observed gradient in these regions (Table 6.3). IPSL-CM5A-LR ranks second in the SOI rankings, capturing the frequency of events well but struggling to capture the standard deviation of the SOI as well as some of the other models over the twentieth century (Table 6.8). It also has lowest surface temperature bias in the Niño 3.4 region (Fig. 6.8) and ranks top in the Niño 3.4 ranking over the entire twentieth century when compared to observations (Table 6.12). IPSL-CM5A-LR has MSLPA more similar to ERA-Interim when using the SOI to define events (Fig. 6.6), compared to the Niño 3.4 (Fig. 6.10), where it fails to capture the strength and extent of anomalies over northern Australia. It captures the surface temperature climatology well across the majority of the equatorial Pacific, especially in the Niño 3.4 region as seen in Fig. 6.8, however it does have a warm temperature anomaly in the far eastern Pacific, a common issue in CMIP models as explained in Section 6.1.2. This was also noted as a bias within the IPSL-CM5A-LR simulations by Dufresne et al. (2013). However, as this is restricted to coastal areas, it does not influence the Niño

3.4 index, thus in part explaining the high performance of this model for the Niño 3.4 analysis. The regions of positive surface temperature anomalies in IPSL-CM5A-LR are larger in extent than those seen in ERA-20C for both EP and CP events. Dufresne et al. (2013) assessed some of the basic ENSO features and climatology within IPSL-CM5A-LR, and found that it produces an ENSO with a spectral peak at 3 to 3.5 years, followed by a secondary peak over 4 years. This supports the frequency findings for this model in Table 6.11 using the Niño 3.4 index, a higher frequency is found when using the SOI to define events. IPSL-CM5A-LR ranks second to GISS-E2-R in terms of ENSO diversity over the twentieth century, therefore ranking highly for both temperature-based and atmospheric-based metrics of ENSO.

Overall, it can be concluded that there are a number of biases in the mean climatology of the equatorial Pacific within the climate model simulations analysed. HadCM3 has a negative MSLP bias over the entire region, while GISS-E2-R has a positive surface temperature bias over the eastern Pacific. There are also differences in the mean and variations in the SOI and the Niño 3.4 indices extracted from the climate model simulations. Using a ranking system which focused on metrics for the SOI and the Niño 3.4 index, it is found that IPSL-CM5A-LR and GISS-E2-R rank within the top three of the five models for both of these indices, and the top two for the SOI (Table 6.8). They are also ranked the top two models for representing ENSO diversity (Table 6.14). As focus is given to the SOI in the following chapters, the results from this chapter suggested that these are the best two models to use in terms of ENSO representation. HadCM3, which performed better for temperature-based indices rather than the SOI, could be considered for further analysis using the Niño indices in the future. Due to the varied performances of different environmental attributes found within the models and the variations in the rankings when analysing different time periods, the use of both IPSL-CM5A-LR and GISS-E2-R for analysis of the Last Millennium would be the most justified from the results of this chapter. Chapter 7 analyses the zonal wind from the five historical climate models in more detail to explore if these models still rank highly, and if so should be used for analysis over the Last Millennium in Chapter 8.

Throughout this analysis, models were re-gridded to a common spatial resolution ($1.5^{\circ} \times 1.5^{\circ}$), however as noted in Table 6.1 the original spatial resolution of the models

were not all the same. IPSL-CM5A-LR has the lowest spatial resolution, but still performed well compared to some of the higher resolution models. On the other hand, GISS-E2-R has one of the highest spatial resolutions and it also performs well. Therefore, the original spatial resolution of the data does not appear to have been a driving factor in the performance of the models. This supports similar findings from Toh et al. (2017) who found that the horizontal resolution is not a good indicator of model performance of precipitation and winds over the Maritime Continent, with regional biases in the models more related to biases in the Hadley circulation and global monsoons. As discussed in Section 6.1.4, there are a number of methods and metrics which can be used for the ranking of climate models. Using different metrics based around temperature and atmospheric measures of ENSO, different rankings were obtained from the five historical climate model simulations. This highlights the necessity of selecting the most appropriate ranking criteria to apply to a given part of the research.

Chapter 7 analyses the ENSO-wind teleconnection within reanalysis and climate models and is directed towards informing logbook-based reconstructions. This focuses on the historical model simulations. Chapter 8 concludes the analysis of climate model output by investigating the ENSO diversity and ENSO-wind teleconnections of selected Last Millennium simulations. The results from these chapters which focus on climate models are summarised within the chapters but also discussed in more detail in Chapter 9.

7. The representation of the mean zonal wind and the stationarity of ENSO-wind teleconnection within historical reanalysis and climate model simulations

7.1 Introduction

This chapter focuses on the representation of zonal wind within a number of reanalysis data products and in the historical climate models simulations of the five selected CMIP5/PMIP climate models introduced in Chapter 6. It assesses the ability of the models to capture mean zonal wind climatology and then explores the teleconnections between ENSO and zonal wind during the boreal winter season. The stability of these teleconnections over the twentieth century is assessed within the reanalysis and the climate model simulations. The stability of ENSO-teleconnection is an underlying assumption of many paleoclimate studies including the logbook-SOI reconstruction presented in Chapter 2. However, stationarity in relationships between the variables over time is not always the case (Gallant et al. 2013). Climate models which span longer than observations and reanalysis provide a useful framework for assessing stationarity, despite their biases (Batehup et al. 2015).

The logbook-based ENSO reconstructions presented in Chapter 2, are based on the assumption of stationarity in the ENSO-wind teleconnection in the study regions from which the logbook data are extracted. Chapter 2 assessed the strength of the ENSO-wind teleconnection between 1979-2013, using ERA-Interim reanalysis data, re-gridded to a coarse resolution ($7.5^\circ \times 8^\circ$). ERA-Interim was used over this modern period as it covers a time span from which observations from satellites are an input into reanalysis and thus informed by the largest number of observations (Poli et al. 2016). However, in order to assess longer term stability in the relationship between zonal wind and ENSO longer datasets are needed. Chapter 5 showed a good agreement between ENSO indices extracted from ERA-Interim and historical reanalyses, ERA-20C and 20CR, which span the entire twentieth century. Therefore, ERA-20C and the 20th Century reanalysis are used here as the best informant of stability in teleconnections. In addition, historical simulations from five climate models are used

to assess the strength and stability of these teleconnections within the climate models over the twentieth century. Chapter 8 builds upon the results presented in this chapter, using paleo-simulations over the Last Millennium from the best performing models to investigate ENSO-zonal wind teleconnections over a longer time period than possible from reanalysis data and from the historical simulations.

7.1.1 Zonal winds within historical reanalysis

Firstly, the winds within the historical reanalysis are investigated. The number of observations of wind speed and direction which are assimilated into the reanalysis varies spatially and temporally within ERA-20C reanalysis over the twentieth century. Fig. 7.1 shows the number of wind observations across the globe in the years 1900, 1920, 1940, 1960, 1980 and 2000 assimilated into ERA-20C, showing an increase in coverage over the twentieth century. Poli et al. (2016) noted that there are regions and time periods of sparse observational coverage but there is an increase in spatial coverage over time. There is a clear increase in observations in the latter half of the twentieth century. The Pacific in the early part of the twentieth century is notable for its few observations (Fig. 7.1). This was an issue also found within the historical logbook period (Chapter 2). There are also differences in the type of wind report, with the majority of observations obtained from ships, although from the mid-1970s buoy measurements are also assimilated, as well as input from fixed ocean platforms from mid-1970s to mid-1990s (Poli et al. 2013). The 20CR reanalysis only assimilated surface pressure, not surface winds, and therefore is expected to be less reliable than ERA-20C.

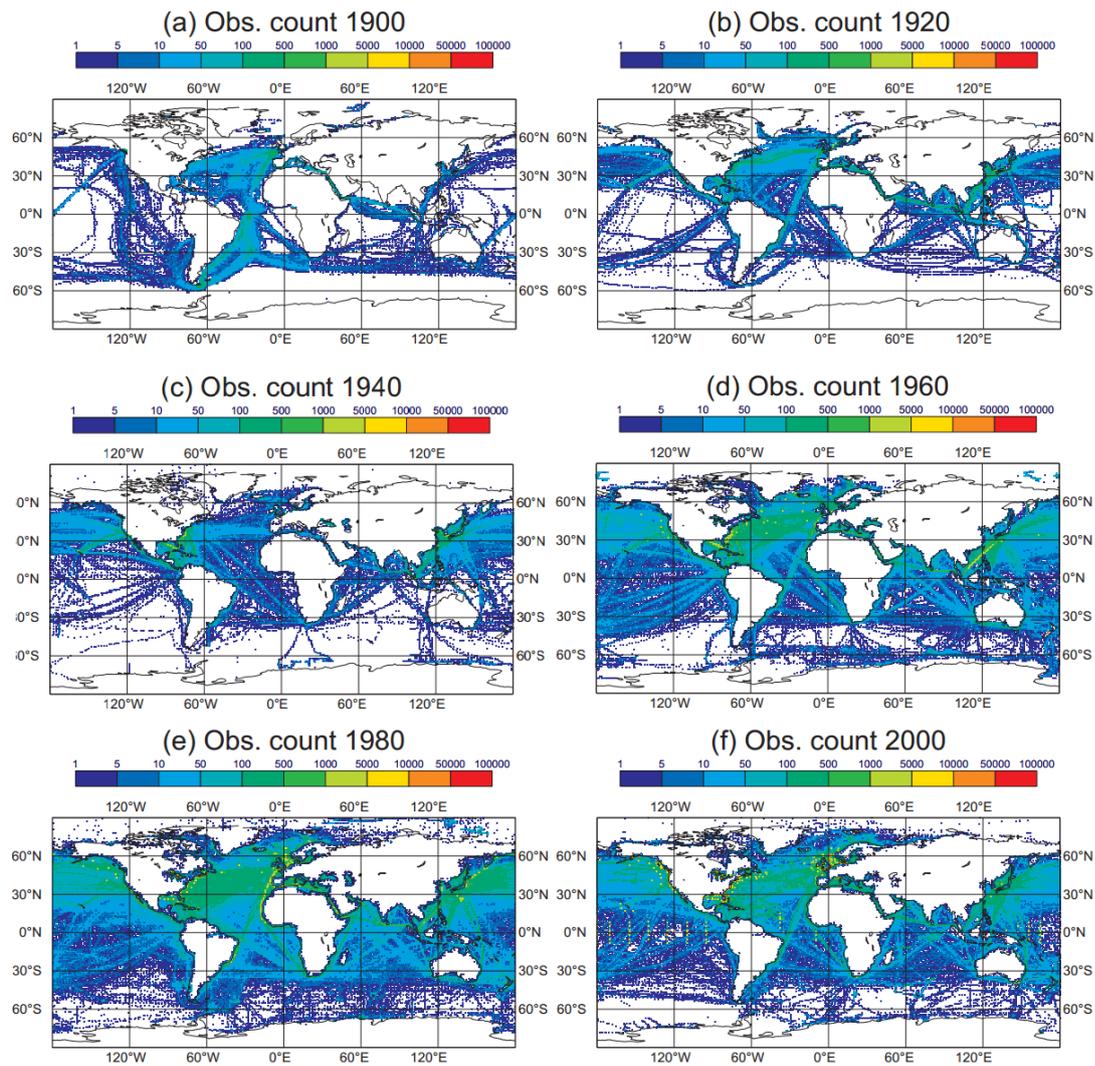


Fig. 7.1 Surface wind vector observations for selected years (1° latitude x 1° longitude grid) from ICOADS 2.5.1 used in ERA-20C. Fig. from Poli et al. (2013).

Winds, and their associated zonal component, are associated with regional and global scale wind, pressure and circulation patterns. The ability of the reanalysis and climate model simulations to simulate the large scale circulation patterns is therefore important for their ability to correctly simulate regional-scale winds. Fig. 7.2 is a general schematic showing mean global surface wind patterns typical in January (Barry and Chorley 2010). In relation to the location of the predictor grid boxes used in the logbook-based reconstruction (locations of which can be seen in Fig. A2), the grid boxes off the coast of South Africa (grid box 2-4) lie around the boundary between the South East Trades and southern hemisphere westerlies winds, grid boxes 5 and 6 lie within the South Easterly Trade region, while 7, 8 and 9 extend into the region of

equatorial westerlies. The spatial patterns of zonal wind climatology are analysed within the historical reanalysis in Section 7.4.1.

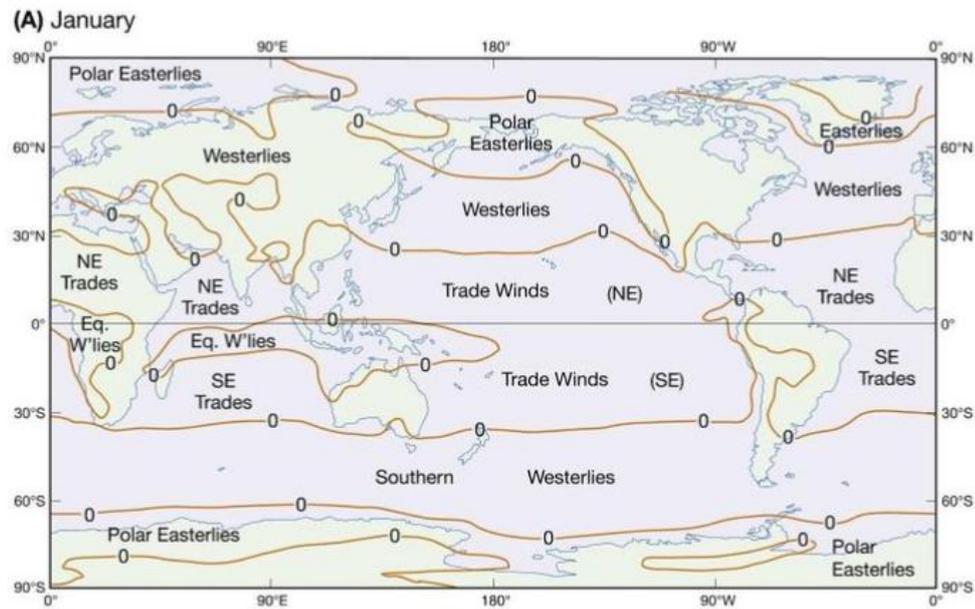


Fig. 7.2 Global wind zones at 1000mb in January, with the zero line representing the boundary between westerly and easterly zonal winds. Based on NCEP/NCAR Reanalysis 1970-1999, Fig. from Barry and Chorley (2010).

Few studies directly focus on zonal winds as their key parameter in climate model investigations, and those which do are largely regionally focused on phenomena such as monsoons, and thus are combined with a focus on precipitation (Jourdain et al. 2013; Brown et al. 2013; Sperber et al. 2013). Toh et al. (2017) analysed the DJF precipitation and wind over the Maritime Continent and its surrounding seas and oceans, a domain which included the predictor grid box 7-9 and part of grid box 6 from this study. They used a multi-model mean (MMM) of the climatological wind bias of 28 CMIP5 models in the Atmospheric Model Intercomparison Project, AMIP5, 1979-2008, and found a weak easterly wind bias in the region in which the predictor grid boxes of the logbook-based reconstruction are located. Looking at the models individually, the easterly bias is stronger than the MMM in GISS-E2-R and IPSL-CM5A-LR, the two models which overlap with the models used here.

7.1.2 Historical and paleo-winds in proxies and climate models

ENSO teleconnections have been extensively researched due to their global impact (Diaz et al. 2001). In terms of historical and paleo- ENSO-teleconnections, temperature and precipitation are often focused upon (Coates et al. 2013; Brown et al. 2016). The ENSO-wind teleconnection is less directly investigated, but is sometimes included within the scope of paleo-monsoon studies (Sperber et al. 2013). Proxy data are more commonly representative of temperature and precipitation than wind, especially in regions where monsoon features are not a characteristic of the regional climate. Paleo-proxy studies do not provide any direct record of winds however a number of processes can be evaluated which provide indications of the behaviour of paleo-winds. These include indications of moisture, dust, surface ocean temperature and fronts, ocean productivity and ocean circulation (Kohfeld et al. 2013). Former wind directions can be reconstructed based on variation in wind-blown sediments such as loess and coversands (Lowe and Walker 2013). Identification of dust within ice and marine cores can also provide indications of wind direction when tracked back to their source (Lowe and Walker 2013). The role of wind direction and velocity has also been investigated using dendrochronology, focusing on the eccentricity of tree ring growth as an indication of prevailing wind direction (Hamilton 2005). However, the emphasis on the winds themselves in this chapter provides a unique focus to ENSO-teleconnections. The wind observations from ships' logbooks are more reliable and more abundantly available than pressure and temperature, therefore analysis of wind patterns over the period 1750-1850 can help to inform the reconstructions presented in Chapter 2 and logbook-based reconstructions in the future.

7.1.3 Strength and stationarity of historical ENSO teleconnections

Climate models have been used to assess the stationarity of ENSO teleconnections (Coats et al. 2013; Gallant et al. 2013; Batehup et al. 2015; Brown et al. 2016). Stationarity is a key assumption of proxy-based ENSO reconstructions, which refers to a static relationship between variables over time, for example the strength of the correlation between the Niño 3.4 SST index and surface temperature in a given region (Batehup et al. 2015). Variables including temperature, precipitation, wind and pressure are all relevant for the assessment of the stationarity of ENSO-climate teleconnections. The instrumental period is not long enough to fully assess the

stationarity of ENSO teleconnections over time, therefore climate models can be used to investigate these statistical relationships on longer time scales. In order to assess the full range of ENSO variability, a period of five centuries has been suggested for use from previous modelling studies (Wittenberg 2009), thus making the LM model simulations highly suited for this purpose.

ENSO is a globally important climatic phenomenon due to the wide-reaching extent of its teleconnections. Therefore, previous work has assessed the stability of the teleconnections in various different regions, concluding that the stability of teleconnections can be regionally dependent (Lewis and LeGrande 2015). Focus here is given to previous studies which, like this one, have used climate models which have both historical and LM simulations, with the aim of using the evaluation of historical simulations to inform the interpretation of the LM results. The LM conclusions from these studies are then discussed further in Chapter 8.

Lewis and LeGrande (2015) assessed the correlation between Niño 3.4 region temperatures and precipitation in the west, central and eastern Pacific over 1979-2005. It was found that the correlation coefficients between the Niño 3.4 index and precipitation were generally of the same sign within the 20CR and the MMM of seven CMIP5 simulations, suggesting a broad agreement between reanalysis and models. However, as seen from Fig. 7.3 the strength of the correlation was weaker in the MMM compared to the 20CR (Lewis and LeGrande 2015). The use of MMM often results in damped characteristics due to the spread of the individual model results. The composites from individual models show better correspondence to the strength of 20CR composite anomalies, which would also likely to be the case for the strength of the correlations within some of the individual models, which is not shown within their paper.

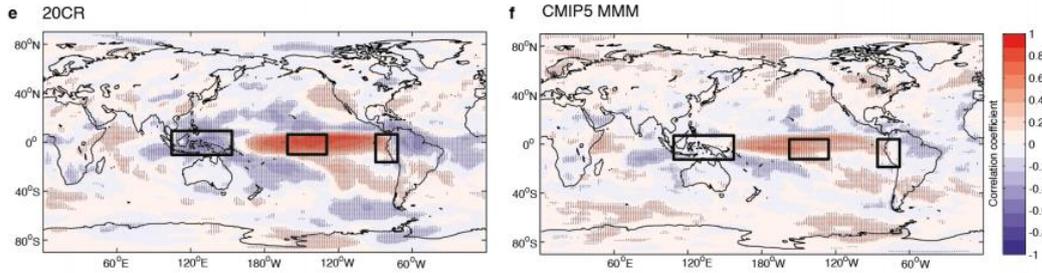


Fig. 7.3 Spatial correlation coefficients between the Niño 3.4 index and 20CR precipitation (left) and CMIP5 7 models MMM (right), model years 1976-2005 (SONDJF). Rectangular boxes indicate East, Central and West Pacific regions (Lewis and LeGrande 2015).

Weaker correlation coefficients when comparing strength of ENSO-teleconnections from observations and models over the historical period were also found by Brown et al. (2016). They analysed the spatial correlation patterns between the Niño 3.4 index and Sept-Feb rainfall over Australia, 1900-1999 (Fig. 7.4). Variations were found in the spatial patterns and in the magnitude of the correlations over the region, comparing six historical climate model simulations from CMIP5/PMIP3 (1900-1999) to observations (1979-2008). In general, they find that the spatial pattern of precipitation anomalies over Australia are broadly well represented. However, a number of known model biases in this region affect the ENSO teleconnection anomaly patterns within the individual models (Brown et al. 2016). These include a South Pacific Convergence Zone that is too zonal (Brown et al. 2013) and the placement of the boundary between negative and positive precipitation anomalies in the equatorial western Pacific too far to the west. Therefore, the ability of models to capture ENSO and ENSO-precipitation teleconnections in Australia varied between individual models.

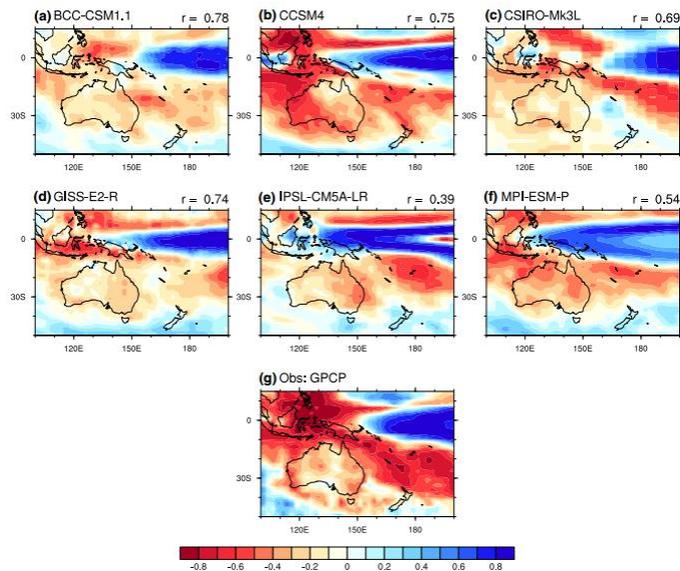


Fig. 7.4 Correlation between Sept-Feb rainfall and the Niño 3.4 index for six CIMP5 simulations, 1900-2000, (a-f), and observations, 1979-2008 (g). All linearly detrended before correlations were calculated. r = spatial pattern correlation coefficient between each model and the observed correlation patterns in the domain (Brown et al. 2016).

ENSO teleconnections over North America were evaluated by Coats et al. (2013) using CMIP5/PMIP3 model simulations. They divided the Last Millennium into 56-year windows and calculated the correlation between the Niño 3.4 and the 200mb geopotential height field over North America, a signal of the North American teleconnection pattern. It was found that the strength and the character of the correlation between North America teleconnection and ENSO was highly variable within and between models (Coats et al. 2013). This suggests that improvements are needed in order for the models to converge to a more realistic representation of ENSO and its teleconnections.

Gallant et al. (2013) examined the stationarity of the relationship between Australian climate and ENSO using running correlations from instrumental, proxy and climate model data sources and found a number of non-stationary climate relationships. However, they suggest that methods which rely on this underlying assumptions can still be useful, as long as these uncertainties are fully acknowledged (Gallant et al. 2013). As the stability of teleconnections can be regionally dependent, as found by Lewis and LeGrande (2015) over the Pacific, it is important to consider using multiple regions within ENSO reconstructions. This was considered and taken into account for

the logbook-based reconstructions and is recently more often identified within ENSO reconstruction literature (Braganza et al. 2009; Capotondi et al. 2015).

This chapter is organised as follows: *Section 7.2* briefly summarises the historical climate models used here, as introduced in Chapter 6, and explains any additional methods adopted within this chapter. *Section 7.3* presents the results of this chapter which focuses on the performance of the reanalysis and historical climate models in capturing the ENSO-zonal wind relationship over the twentieth century. *Section 7.4* summarises the results from this chapter and the historical climate model analysis as a whole, and how they are used to inform the analysis within Chapter 8. The work within this chapter addresses Aim 3, specifically Objective 3C, as outlined in Chapter 1.

7.2 Data and methods

The datasets used here were also utilised in the previous chapter and include instrumental records of the SOI, the Niño 3.4 index, the three reanalysis datasets (ERA-Interim, ERA-20C and 20CR) and the five historical climate model simulations (GISS-E2-R, HadCM3, IPSL-CM5A-LR, MIROC-ESM and MPI-ESM-P). Initially, the data were used at a common $1.5^{\circ} \times 1.5^{\circ}$ resolution, and then were re-gridded to match that of the resolution of data used in the logbook-based reconstructions, $7.5^{\circ} \times 8^{\circ}$ latitude-longitude resolution. Correlations between the DJF zonal wind and the DJF SOI, and the Niño 3.4 index, were analysed in Chapter 2 for the logbook-based reconstruction using ERA-Interim data, 1979-2013. The SOI was selected as the main index as it is the atmospheric component of ENSO and therefore is the ENSO index focused upon here. The same analysis was carried out using the Niño 3.4 index instead of the SOI to analyse the difference in using a different ENSO index on the concluding results. The full results from this analysis can also be found in the Appendix D and are only briefly summarised within this chapter.

Firstly, the mean DJF zonal wind in the logbook regions were assessed in the various datasets, to investigate how well winds are represented with climate model simulations. The mean and standard deviation from the reanalysis and models were compared. A two-sampled t-test was carried out to assess if the mean zonal winds are statistically different within the reanalysis and models in the locations of the logbook observations.

This is then followed by analysis of Pearson's correlation coefficient between reanalysis- and model-derived SOI, and reanalysis and model mean zonal wind. This was carried out at both high resolution ($1.5^\circ \times 1.5^\circ$) and logbook resolution ($7.5^\circ \times 8^\circ$) over the period 1979-2004. This was then repeated over the longer time period, 1900-1999. All time series were linearly detrended before correlations were calculated. Running correlations were calculated using a 30-year window. Spatial climatology and composite anomaly plots were also created as a tool for diagnostic evaluation to assess patterns across the regions. The models were then ranked in order to select two models that capture the ENSO-wind teleconnections best over the twentieth century. The metrics used for ranking of the models were specific to the logbook-based reconstruction presented in Chapter 2, and therefore, although providing some indication of the larger scale representation of ENSO-winds, they are more focused on aiding further research within logbook-based reconstruction and use a method in which future studies can assess the stationarity of teleconnections within their chosen region.

7.3 Results

7.3.1 Mean DJF zonal wind in logbook predictor grid boxes within reanalysis

The mean DJF zonal wind climatology from the three reanalyses over their period of overlap, 1979-2004, is shown in Fig. 7.5. The coarser resolution predictor grid boxes used in the logbook-ENSO reconstruction are highlighted. In terms of the general circulation all three of the reanalysis products show easterly winds over much of the tropical Pacific, namely, the surface wind component of the Walker Circulation (as explained in Chapter 1). In the Indian Ocean, there are negative zonal wind anomalies broadly centred around 20°S in the central Indian Ocean, spanning across grid boxes 5 and 6. This is associated with typical cyclonic conditions to the just to the south of this region, which promote south-easterly winds in this region, and westerly winds further south. Weak mean winds are found in the region around South Africa as they lie on the border of southern hemisphere westerlies and the southeast trade winds, in agreement with the general pattern illustrated by Fig. 7.2 (Barry and Chorley 2010). Equatorial westerlies, as suggested by Fig. 7.2, can be seen in the three reanalysis over the equatorial Indian Ocean and the southern Maritime Continent. These equatorial

westerlies are stronger across the Indian Ocean in the ERA-Reanalysis than in 20CR. The maximum strength of the winds is higher in the 20CR in the region just to the north of the equator in both the central Pacific and western Atlantic Ocean. The differences in the mean zonal wind between ERA-20C and 20CR could be in part due to the assimilation of wind observations within ERA-20C, whereas winds are not assimilated in the 20CR. Therefore, it is expected that ERA-20C would better represent zonal wind climatologies than 20CR.

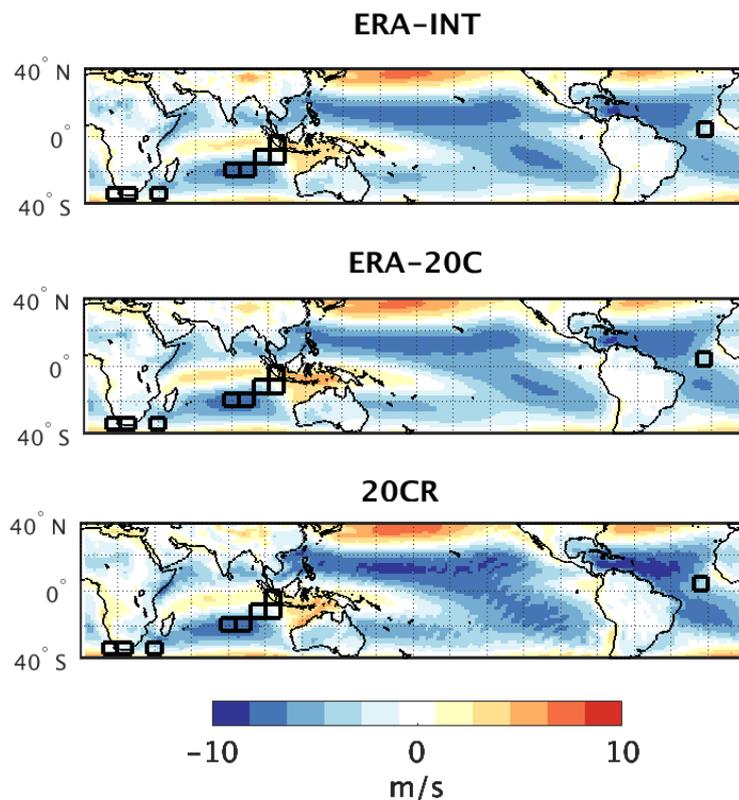


Fig. 7.5 Mean DJF zonal wind (ms^{-1}) 1979-2004 from ERA-Interim, ERA-20C and 20CR reanalysis at higher ($1.5^\circ \times 1.5^\circ$) resolution. Black boxes indicate grid boxes used in logbook-based reconstruction, with grid box 1 in the Atlantic Ocean, and grid boxes 2-4 from west to east around South Africa and west to east 5-9 in the Indian Ocean.

In order to carry out the logbook-ENSO reconstruction the zonal wind from the reanalyses was re-gridded to $7.5^\circ \times 8^\circ$ latitude-longitude resolution. The climatological mean values at this coarser resolution for each predictor grid box from the three reanalyses are shown in Fig. 7.6. As seen in Fig. 7.5, grid box 1 is in the equatorial

Atlantic Ocean, grid boxes 2, 3 and 4 are located to the South of South Africa, boxes 5 and 6 are in the Indian Ocean centred around 20°S, and grid boxes 7, 8 and 9 are also in the Indian Ocean but closer towards the Maritime Continent, with grid box 9 over western Indonesia. The exact locations of these grid boxes can be found in Appendix A (Table A1; Fig. A2). The strongest zonal winds are found in grid boxes 5 and 6. The mean values from ERA-20C are higher than those from ERA-Interim for all of the nine logbook grid boxes (Fig. 7.6), while the mean values from 20CR are higher than ERA-Interim in all but grid box 2 and 7. This discrepancy in wind strength between the century-length reanalyses and ERA-Interim has not been reported previously, but it must be noted that the results are largely specific to the logbook-based reconstruction regions. Wider scale comparisons of zonal winds from reanalysis are suggested as an avenue for further work.

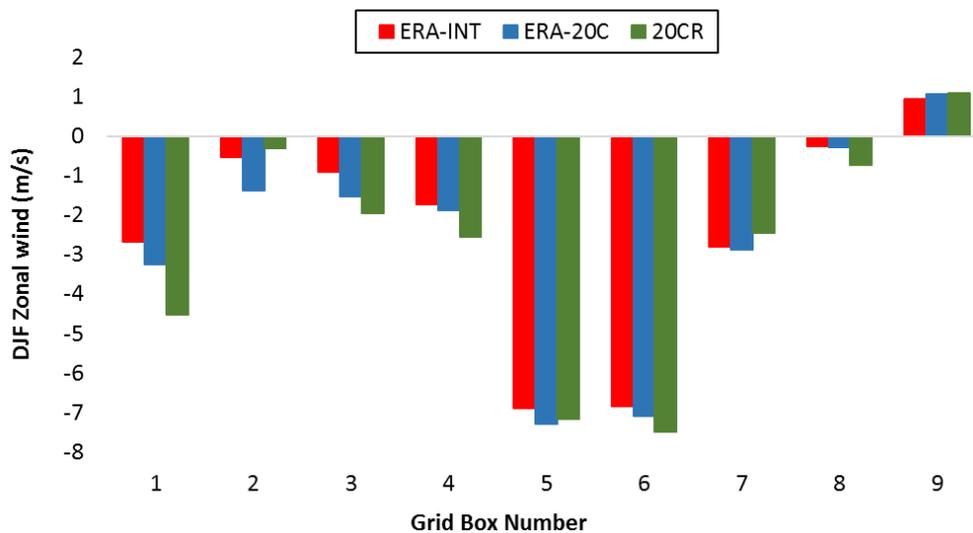


Fig. 7.6 DJF zonal wind (ms^{-1}) from logbook grid boxes, ERA-Interim, ERA-20C and 20CR, 1979-2004 ($7.5^\circ \times 8^\circ$ resolution).

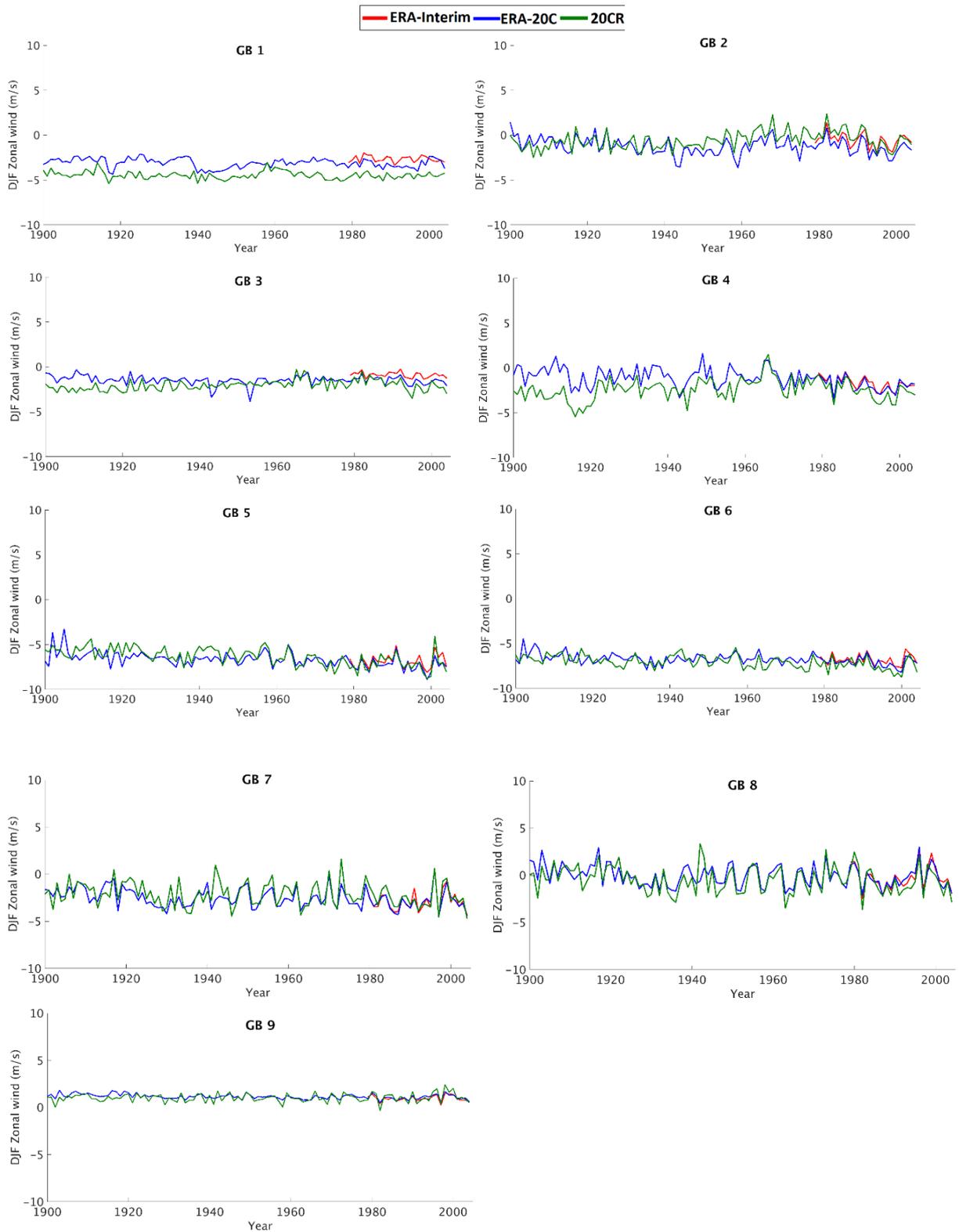


Fig. 7.7 DJF zonal wind (ms^{-1}) for the nine logbook grid boxes from ERA-Interim (1979-2004), ERA-20C and 20CR reanalysis, 1900-2004.

In order to explore the coherence between the reanalysis datasets further the interannual variability in the DJF zonal winds from each predictor grid box 1900-2004 is analysed (Fig. 7.7). This, along with correlation analysis and t-tests comparing the zonal winds from the various datasets, provide additional analysis which assesses both the mean but also the variance of the datasets. The Pearson's correlation coefficient between the mean DJF zonal wind is calculated for each grid box, comparing the values from each of the reanalysis datasets for each predictor grid box. Firstly, the two historical reanalyses were compared to ERA-Interim (1979-2004) and good agreement between the datasets was found, with positive significant (95% confidence level) correlations found in each of the predictor grid boxes (Table 7.1).

Table 7.1 Pearson's correlation coefficient between mean DJF zonal wind in re-gridded logbook grid boxes, from ERA-20C and 20CR when correlated with ERA-Interim over 1979-2004.

	Predictor grid box								
	1	2	3	4	5	6	7	8	9
ERA-20C	0.41	0.94	0.87	0.94	0.88	0.89	0.95	0.95	0.92
20CR	0.55	0.88	0.76	0.91	0.80	0.80	0.90	0.88	0.90

When comparing to ERA-20C, the weakest Pearson's correlation coefficient between the ERA-20C and ERA-Interim is found in grid box 1 ($r=0.41$), suggesting the least agreement in this Atlantic grid box compared to the grid boxes from the other regions. The remaining eight grid boxes have strong positive correlations between the datasets ($r=0.87$ to $r=0.95$), suggesting better agreement in the pattern of interannual-variability of the two reanalysis datasets. A two-sampled t-test was carried out to test the null hypothesis that the DJF zonal wind from ERA-Interim and ERA-20C have equal means and equal but unknown variances. The results from the t-test indicate that for grid boxes 4, 6, 7, 8 and 9 the difference between ERA-Interim and ERA-20C values within each grid box is not statistically significant at the 95% confidence level (Appendix Table D1). However, this is not the case for grid boxes 1, 2, 3 and 5, despite the good temporal agreement for the latter three grid boxes. Therefore, the DJF zonal wind from the different ERA- reanalysis datasets vary in some regions suggesting differences in the reanalysis datasets.

When comparing the mean DJF zonal wind from 20CR with ERA-20C reanalysis, similar but weaker results are generally found ($r=0.76$ to $r=0.91$ in boxes 2-9) (Table 7.1; Fig. 7.7). The 20CR has lower correlations with ERA-Interim than it was found for ERA-20C in all but grid box 1 ($r=0.55$) (Table 7.1). This suggests that there is better agreement at the grid box scale between the two ERA-Reanalysis products compared to the 20CR. A two sampled t-test found that the mean DJF zonal wind is statistically different in four of the nine predictor grid boxes when comparing both ERA-20C and 20CR to ERA-Interim over this shorter climatological period, 1979-2004 (Appendix Table D1).

Table 7.2 Pearson’s correlation coefficient DJF zonal wind from ERA-20C and 20CR.

	Predictor grid box								
	1	2	3	4	5	6	7	8	9
1900-1949	0.49	0.38	0.19	0.50	0.24	0.17	0.62	0.64	0.56
1950-1999	0.26	0.81	0.57	0.82	0.83	0.75	0.88	0.89	0.87

The two historical reanalyses were then compared over the entire twentieth century. Fig. 7.7 shows the variability within the zonal wind in the predictor grid boxes over this period. Pearson’s correlation coefficient comparing the data over 1900-1949, and 1950-1999, in Table 7.2, show that there is a reduced agreement between the datasets in all of the grid boxes, except grid box 1, when looking at the earlier half of the twentieth century compared to the latter half. A two-sampled t-test was carried out comparing the datasets from the two reanalysis for each of the predictor grid boxes suggesting that there is a significant difference between the zonal wind at this grid box scale resolution. Overall, due to the stronger correlations and the closer match to ERA-Interim within the climatology plots (Fig. 7.5) it is found that ERA-20C compares better to ERA-Interim than 20CR does. Therefore, henceforth historical model simulations are compared to ERA-20C over this period.

7.3.2 Mean DJF zonal wind in logbook grid boxes within historical climate model simulations

Following on from the analysis of the mean DJF zonal wind within the reanalysis, this is now investigated for the five historical climate model simulations to assess their

ability to capture the general global wind patterns as well as the mean wind values within the predictor grid box locations. The mean DJF zonal wind climate model bias is shown in Fig. 7.8, calculated relative to the ERA-Interim climatology, 1979-2004. Spatial plots are calculated over this period and compared to ERA-Interim due to the higher number of observations assimilated into ERA-Interim compared to the historical reanalyses.

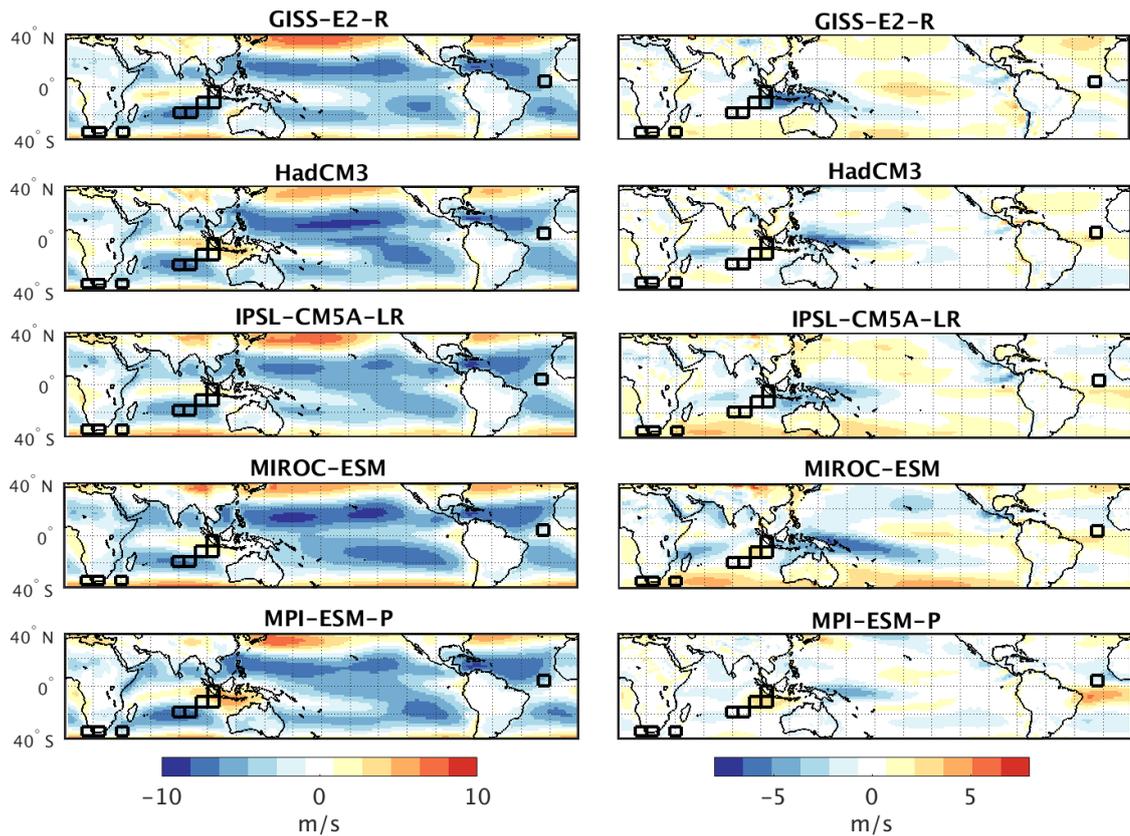


Fig. 7.8 DJF zonal wind (ms^{-1}) climatology, 1979/80-2004/05 from historical climate model simulations (left) and model zonal wind bias (Model climatology minus ERA-Interim reanalysis) (right).

The main global patterns of zonal wind circulation from the reanalysis and climate models are similar, with all five models capturing the easterly zonal wind across the equatorial Pacific, the south-easterly trade winds in the southern Indian ocean and the equatorial westerlies in the equatorial Pacific Ocean (Fig. 7.5; Fig. 7.8). However, the strength of the zonal winds and the exact locations of these regions is not entirely consistent with the reanalysis, as seen from the model zonal wind bias plots (Fig. 7.8). A number of models (GISS-E2-R; IPSL-CM5A-LR; MIROC-ESM) have a positive

zonal wind anomaly in the south east trade region in the south west Indian Ocean, suggesting difficulty in capturing the strength of these trade winds. GISS-E2-R has a negative wind bias over the eastern Indian Ocean region containing grid box 7, 8 and 9, whilst HadCM3, MIROC-ESM and MPI-ESM-P have a weak positive zonal wind bias covering some of this region (Fig. 7.8). This suggests the models have difficulty capturing the equatorial easterlies in this region, with positive anomalies suggesting that the models have these winds stronger than the reanalysis, and negative anomalies suggesting they are weaker than reanalysis. IPSL-CM5A-LR best corresponds to ERA-Interim zonal wind in grid boxes 7, 8 and 9, with weak negative anomalies. Toh et al. (2017), using a MMM and the individual GISS-E2-R and IPSL-CM5A-LR, also found a negative wind bias in this region, in agreement with the results for these two models in Fig. 7.8.

The zonal wind bias plots also show good correspondence between MPI-ESM-P and ERA-Interim in the region around South Africa and into the southern Indian Ocean, corresponding with good agreement in grid boxes 2, 3, 4 and 6. HadCM3 zonal wind is also close to ERA-Interim in this region, whereas the remaining three models GISS-E2-R, IPSL-CM5A-LR and MIROC-ESM have a positive zonal wind bias in this region, which suggests that the southern hemisphere westerlies are shifted further north in these models compared to the reanalysis. Combining the zonal wind bias plots with the results from the t-test (Appendix Table D2) and time series comparison over 1979-2004 it can be concluded that IPSL-CM5A-LR performs best in the eastern Indian Ocean and western Maritime Continent (i.e. grid boxes 7, 8 and 9), whereas HadCM3 and MPI-ESM-P generally perform best in the South African and southern Indian Ocean region (i.e. grid boxes 2-6).

The mean DJF zonal wind from the five climate model simulations were extracted for the logbook grid boxes, and plotted along with ERA-Interim 1979-2004 (Appendix Fig. D1) and ERA-20C over the entire twentieth century (Fig. 7.9, Fig. 7.10). In cases where the mean DJF zonal wind in ERA-Interim and ERA-20C are close to zero such as grid boxes 2, 3, 4 and 8, the sign of the mean zonal wind value within the models does not always match that of the reanalysis. Grid box 8 has a large divergence in mean values, with an ERA-20C mean of -0.04 ms^{-1} and climate model means ranging from -3.86 ms^{-1} (GISS-E2-R) to 2.73 ms^{-1} (MPI-ESM-P). The closest correspondence

between the reanalysis and models is found in grid boxes 5 and 6, where strong easterly winds are present (i.e. Indian Ocean Trade wind regions). Disagreement is found in grid boxes 7 and 8 which links with issues previously highlighted in the literature with Maritime Continent climate such as a weak ENSO-monsoon connection and a lack of consensus from the models on seasonal precipitation in this region (e.g. Jourdain et al. 2013; Toh et al. 2017).

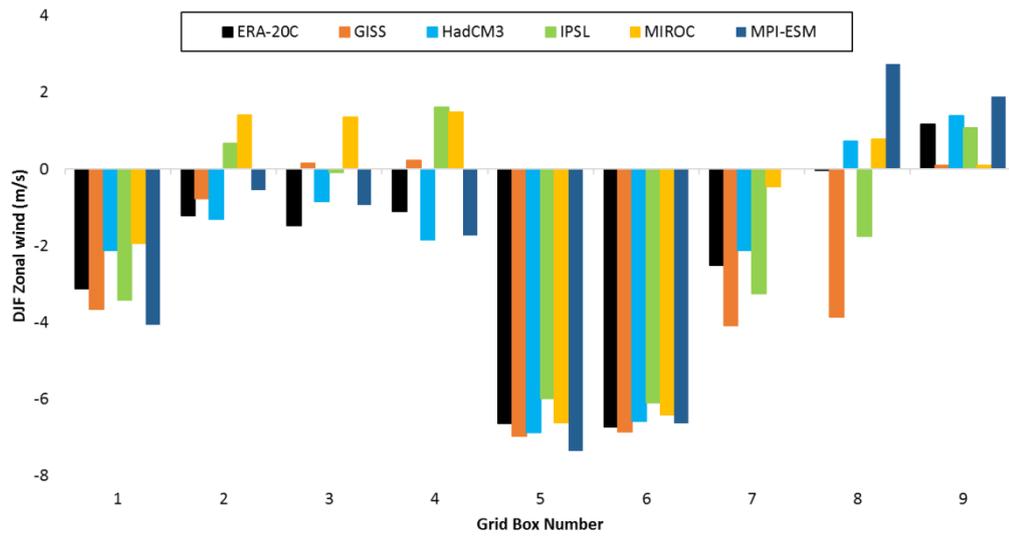


Fig. 7.9 DJF Mean zonal wind (ms^{-1}) per grid box (1-9), 1900-1999 from ERA-20C and climate model simulations at $7.5^\circ \times 8^\circ$ latitude-longitude.

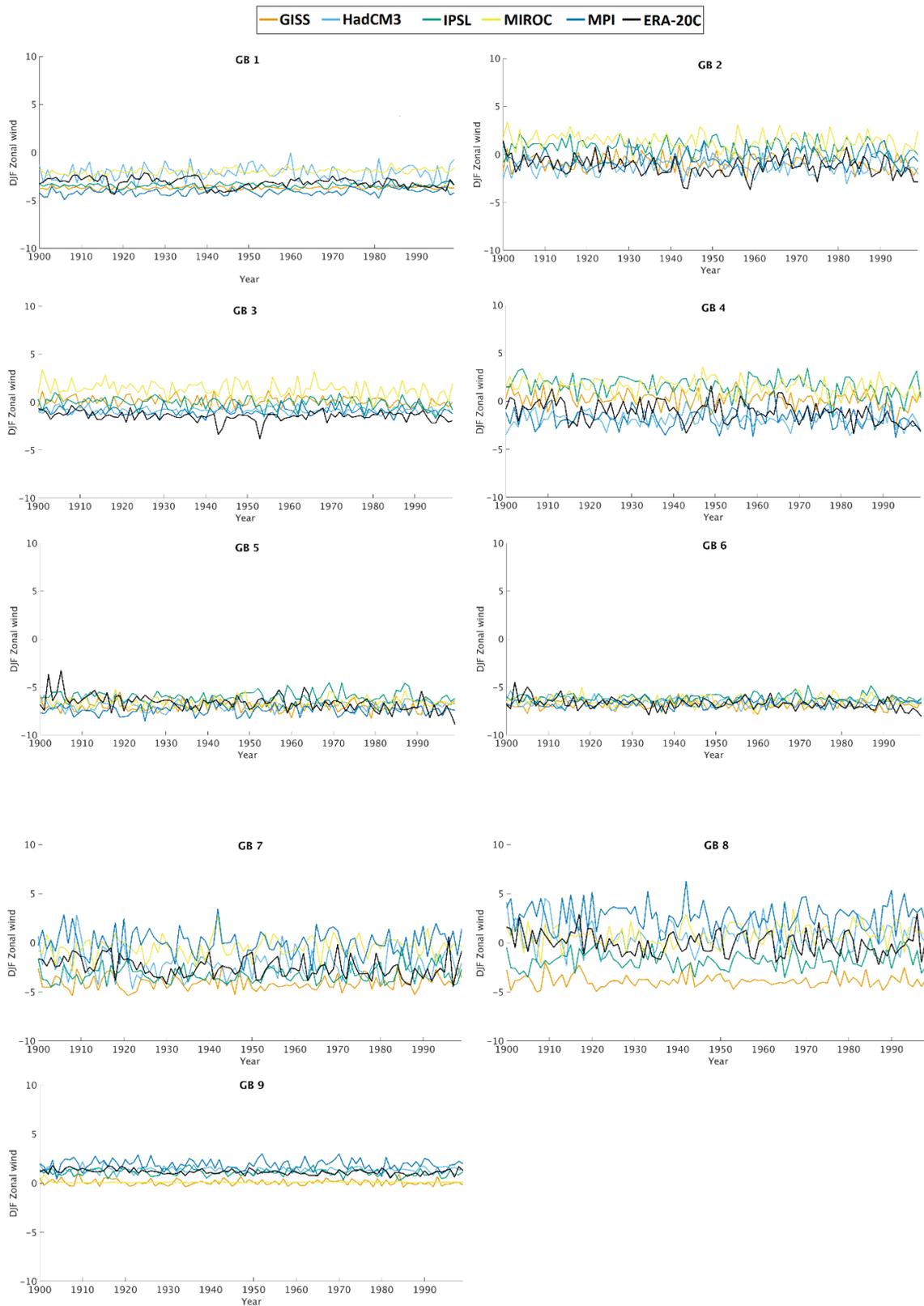


Fig. 7.10 The DJF mean zonal wind (ms^{-1}) in logbook predictor grid boxes, 1-9 ($7.5^\circ \times 8^\circ$) from ERA-20C and climate model simulations, 1900-1999.

There is variability in the spread of the zonal means from the models in the predictor grid boxes when analysing the wind over the entire twentieth century (Fig 7.10). The spread of DJF mean zonal winds in the different grid boxes varies between the models. In grid boxes 1,3,5,6, and 9 there is much smaller model spread than in grid boxes 2,4 7 and 8. This suggests that in the later grid boxes the models fail to capture the mean zonal wind as well.

A two-sample t-test was carried out to assess if the time series of zonal mean winds from the climate model simulations were statistically different from those from ERA-20C dataset (see Appendix Table D3). The two-sample t-tests comparing the five model simulations to the ERA-20C reanalysis found that in each model only one of the grid boxes had a similar mean and variance to that found in the reanalysis. The location of the grid box with similarity to ERA-20C varied between the five models, with no region showing better agreement. The disagreement between the zonal wind between reanalysis and the climate model simulations could in part be a result of the resolution used, with climate models being better at capturing large-scale circulation, but not as good at a grid box level.

Another reason for this general lack of agreement in the zonal wind mean and variance is due to the decreasing quality of the reanalysis data further back in time, as discussed in Chapter 6. It is also challenging for climate models to replicate mean zonal wind at specific grid box scale, especially in regions of model bias. Therefore, the analysis is continued by looking into the correlation between zonal wind and the SOI during the boreal winter season, motivated by the use of this teleconnection within the logbook-based reconstruction.

7.3.3 Correlations between DJF zonal wind and the DJF SOI, 1979-2004

7.3.3.1 Reanalysis comparison

In order to further assess the relationship between the DJF zonal wind and ENSO, the Pearson's correlation coefficient between zonal wind and the SOI, and the Niño 3.4 index were calculated and plotted. The DJF SOI were extracted from the reanalysis and model outputs (1.5° x 1.5° resolution data), and is focused on here as this ENSO measure was used in the logbook-based reconstruction. As explained in Chapter 2, this

measure has good agreement with the Niño 3.4 index, but is thought to be a better representation of the atmospheric component of ENSO, and thus the winds. Fig. 7.11 shows the spatial correlation between the DJF zonal wind and DJF SOI from ERA-Interim and ERA-20C reanalysis data. Time series were detrended over the entire period (1979-2004) prior to calculation of correlations. This is shown at the higher resolution, with the coarser logbook resolution for ERA-Interim seen in Chapter 2 (Fig. 2.5). The spatial patterns of ENSO-wind teleconnections are shown to be similar within both datasets. Within the nine logbook grid boxes, the strength and sign of the correlation coefficients are similar in both reanalyses over this period, with negative correlation in grid boxes 2-5 and positive correlations in 1 and 7-9. Positive correlations in grid boxes 7-9, and in a larger region covering much of the Maritime Continent and the northern Indian Ocean suggest that during El Niño events there is a reversal in the climatological equatorial westerlies as seen in Fig. 7.5, and zonal wind weakens or is reversed. This is due to the divergence of surface winds as the Walker Circulation in the Pacific is reversed and descending air masses prevail over the western Pacific and the Maritime Continent, as depicted in Fig. 1.2. The descent of air in the western Pacific during El Niño (negative SOI events), results in a negative correlation between SOI and zonal wind over the central equatorial Pacific, with a reduction in the strength of the typical easterly winds in this region, and in some cases westerly winds prevailing. In the region around South Africa which contains grid box 2-4, negative correlations are seen due to westerly winds during El Niño events, and more easterly zonal wind in La Niña events.

The correlation analysis was also carried out using the same datasets, but re-gridded to the $7.5^\circ \times 8^\circ$ latitude-longitude resolution of the logbook data analysis. The correlation coefficients for each logbook grid box at this coarser resolution is shown in Fig. 7.12, with grid box one showing the largest difference in correlation strength between the ERA-reanalysis datasets. This can be seen by the weaker and smaller region of positive correlations in the central Atlantic in ERA-20C, compared to ERA-Interim.

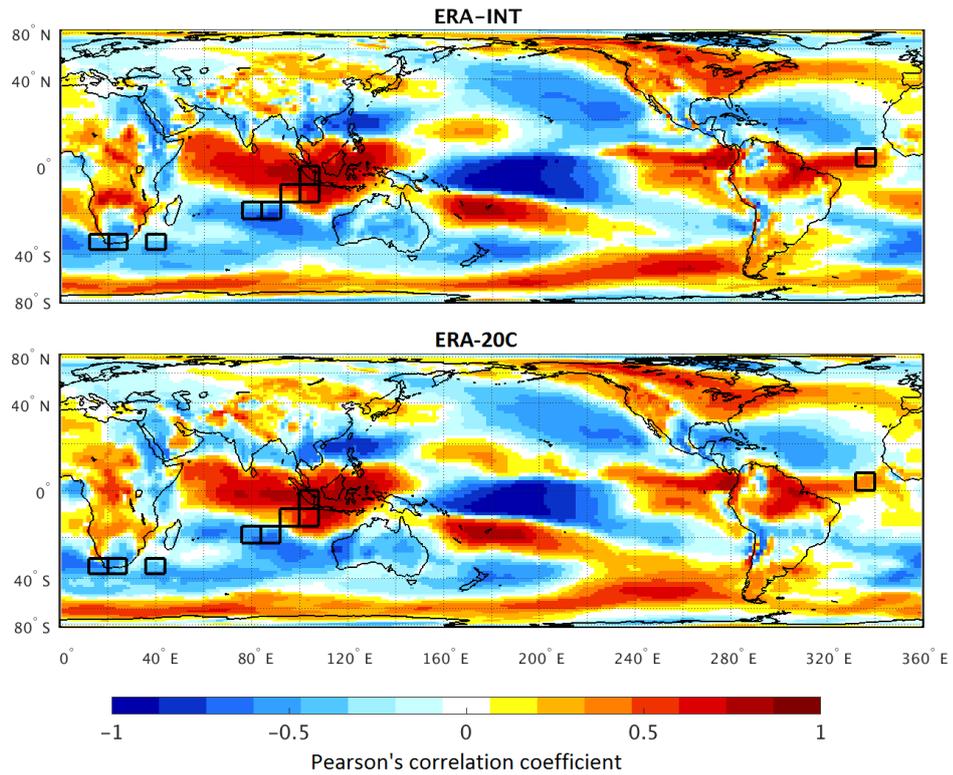


Fig. 7.11 Pearson's correlation coefficients between mean DJF zonal wind and DJF SOI from ERA-Interim (top) and ERA-20C (below), 1979-2004. Data detrended prior to calculations of correlations. Black rectangles indicate logbook grid boxes. All values shown.

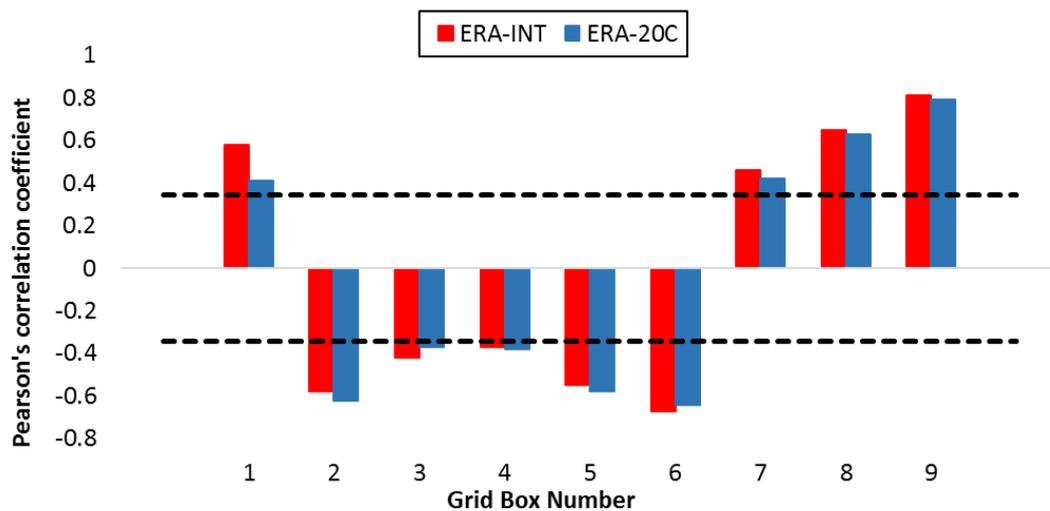


Fig. 7.12 Pearson's correlation coefficient between DJF Zonal wind per grid box (1-9) and DJF SOI from reanalysis. Dashed lines indicates 90% confidence level. Time series detrended prior to calculations of correlations.

7.3.3.2 Climate model historical simulation comparison

Correlation analysis was also then carried out using the five climate model simulations over the same climatological period (See Appendix Fig. D2; Fig. D3). The spatial pattern of correlation varies more clearly when comparing the various climate model outputs (Fig. D3) compared to comparisons of the reanalysis datasets (Fig. 7.12). Within the logbook grid boxes, there is less consistency in the strength and the sign of the correlations in the climate models compared to the reanalysis data. All of the models except MIROC-ESM have weaker correlation coefficients than ERA-Interim in all nine of the logbook grid boxes. GISS-E2-R has very weak, non-significant (at the 90% confidence level) correlations in all but grid box 9. MPI-ESM-P only have a significant correlation in two of the nine grid boxes. The 90% confidence level was used here, in-line with that used in the initial correlation analysis which informed the logbook-based reconstruction (Chapter 2). MIROC-ESM has the highest number of grid boxes with significant correlations (at the 90% confidence level), 2, 3, 4, 5 and 8, however the strength of the correlation in grid boxes 3 and 4 is stronger than that found in the same grid boxes in ERA-Interim. HadCM3 has significant correlation between zonal wind and the SOI in only three of the nine grid boxes, whereas IPSL-CM5A-LR has significant correlation in four of the grid boxes.

As well as the number of grid boxes with correlations significant at 90% confidence level, the number of grid boxes in which the sign of the correlation is the same as that in ERA-Interim can be used as an indication of ENSO-wind teleconnection representation within the historical model simulations (Table 7.3; Table 7.4). The sign of the correlation is opposite to ERA-Interim in three of the logbook grid boxes in HadCM3 and in two of those in MIROC-ESM. The remaining three models only have the opposite sign correlation to ERA-Interim in one of the nine grid boxes. Grid box 1 in the Atlantic has the opposite sign in three of the five models compared to ERA-Interim, which is more than any of the other eight grid boxes. This suggests that the ENSO-wind teleconnection in the Atlantic is less well represented within the five model simulations overall compared to the other regions.

The results from the analysis of the SOI-wind correlations within the logbook grid boxes suggests that MIROC-ESM captures the SOI-zonal wind teleconnection best compared to ERA-Interim over this period, despite its wind bias. However, large

uncertainties exist when evaluating models over short periods such as that covered by ERA-Interim, due to inter-decadal and inter-centennial variability in ENSO behaviour (Wittenberg 2009; Lewis and LeGrande 2015). Therefore, although MIROC-ESM performs best in this measure over this period, this may be in part due to the time slice analysed. In addition, it is a challenge for climate models to represent characteristics at a grid box level. In general, climate models can capture the spatial patterns associated with large scale circulations but not always in the locations at a higher spatial resolution of an individual grid box.

Correlation coefficients between zonal wind and the DJF SOI within the models are generally found to be weaker than those in the reanalysis. This was also found in previous ENSO-teleconnection studies (Lewis and LeGrande 2015; Brown et al. 2016). The models used here were then ranked using four metrics (Table 7.3) to assess their similarity to reanalysis. The ability of the models to capture mean zonal wind, the SOI and the correlation between the two variables is assessed, with models ranked from 1-5, with one being closest to the reanalysis, and thus assumed to be most realistic. The models were then given an overall ranking based on the average of their ranking for the four metrics assessed. Over the shorter period, 1979-2004, the models were compared to ERA-20C and it was found that IPSL-CM5A-LR performs best, followed by GISS-E2-R and MPI-ESM-P jointly (Table 7.4). This analysis and ranking method was then carried out over the entire twentieth century, 1900-1999.

Table 7.3 Metrics used to rank the models against each other when compared to ERA-20C.

Metric 1	Results from grid-box t-test of mean
Metric 2	SOI rank from Chapter 6
Metric 3	Rank of the number of predictor boxes with significant correlations (90% confidence level) between SOI and zonal wind
Metric 4	Rank of the number of predictor boxes with the same sign correlation as that from ERA-20C

Table 7.4 SOI-wind correlations ranks compared to ERA-20C, 1979-2004.

	Metric 1	Metric 2	Metric 3	Metric 4	Overall rank
GISS-E2-R	=3	2	5	=1	=2
HadCM3	=1	4	3	5	=4
IPSL-CM5A-LR	=3	1	2	=1	1
MIROC-ESM	5	3	1	4	=4
MPI-ESM-P	=1	5	4	=1	=2

7.3.4 Correlation between zonal wind and the SOI over the twentieth century

The focus of this section is to look at a 100-year historical period in order to provide consistency with the periods that will be analysed over the Last Millennium, Therefore, henceforth the 1900-1999 period is focused upon in this chapter. As previously stated, the inter-decadal and inter-centennial scale variations in ENSO cannot be fully revealed when evaluating models over shorter time scales (Lewis and LeGrande 2015). Thus, suggesting that this longer time period may provide more robust results from the climate models. The rankings from this period are used to select the climate models for use over the LM in Chapter 8.

The Pearson's correlation coefficients are generally weaker when analysing the period 1900-1999 compared to 1979-2004, and this will be partly due to the increase in sample size. However, the critical value for significance at the 90% confidence level is lower as a result of this. Weaker correlation values can be seen in the spatial correlation plot using the SOI and DJF zonal winds from ERA-20C (Fig. 7.13), compared to Fig. 7.11 (1979-2004). Nevertheless, the correlations within the logbook grid boxes are of the same sign as that from the later period, 1979-2004. It should be noted that in the earlier part of the twentieth century the number of observations input into ERA-20C is drastically reduced compared that that of the more recent period, and therefore there is increased uncertainty in the dataset. However, it is still a useful indicator of the behaviour of ENSO and ENSO-teleconnections and the ENSO indices extracted are close that that of the instrumental record as seen in Chapter 5 (Poli et al. 2016).

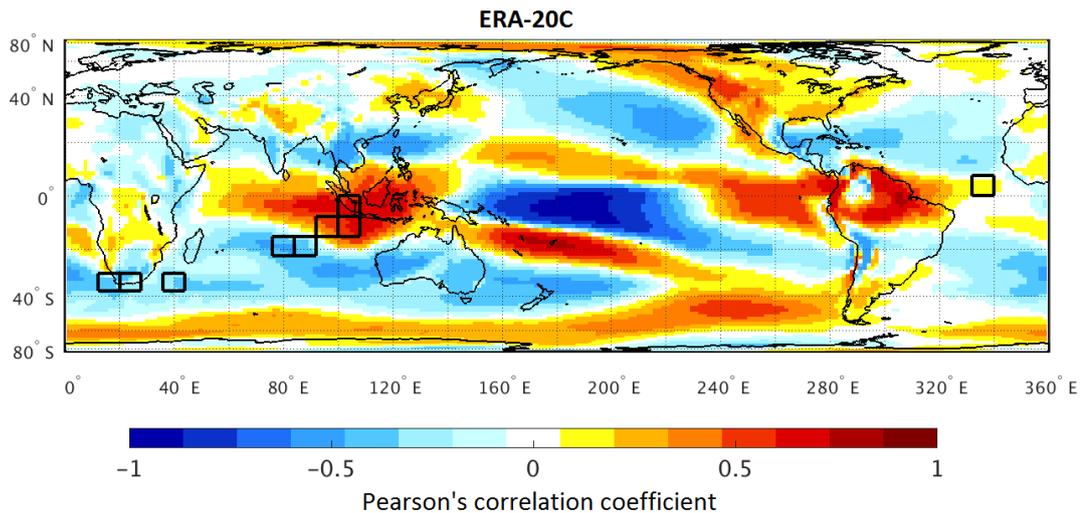


Fig. 7.13 Pearson's correlation coefficients between mean DJF zonal wind and DJF SOI from ERA-20C 1900-1999. Data detrended prior to calculations of correlations. Black rectangles indicate logbook regions. All values shown.

The Pearson's correlation coefficient between the DJF SOI and the DJF zonal wind from the five historical climate models over the period 1900 to 1999 can be seen in Fig. 7.14. Again the correlations from the climate models are weaker than those from the reanalysis. The correlation coefficients between the DJF SOI and the zonal wind in the nine predictor grid boxes is shown in Fig. 7.15, for ERA20C and the five climate model simulations. The correlation coefficients from the climate models are generally weaker than those from ERA-20C, supporting the findings of Fig. 7.13 and Fig. 7.14.

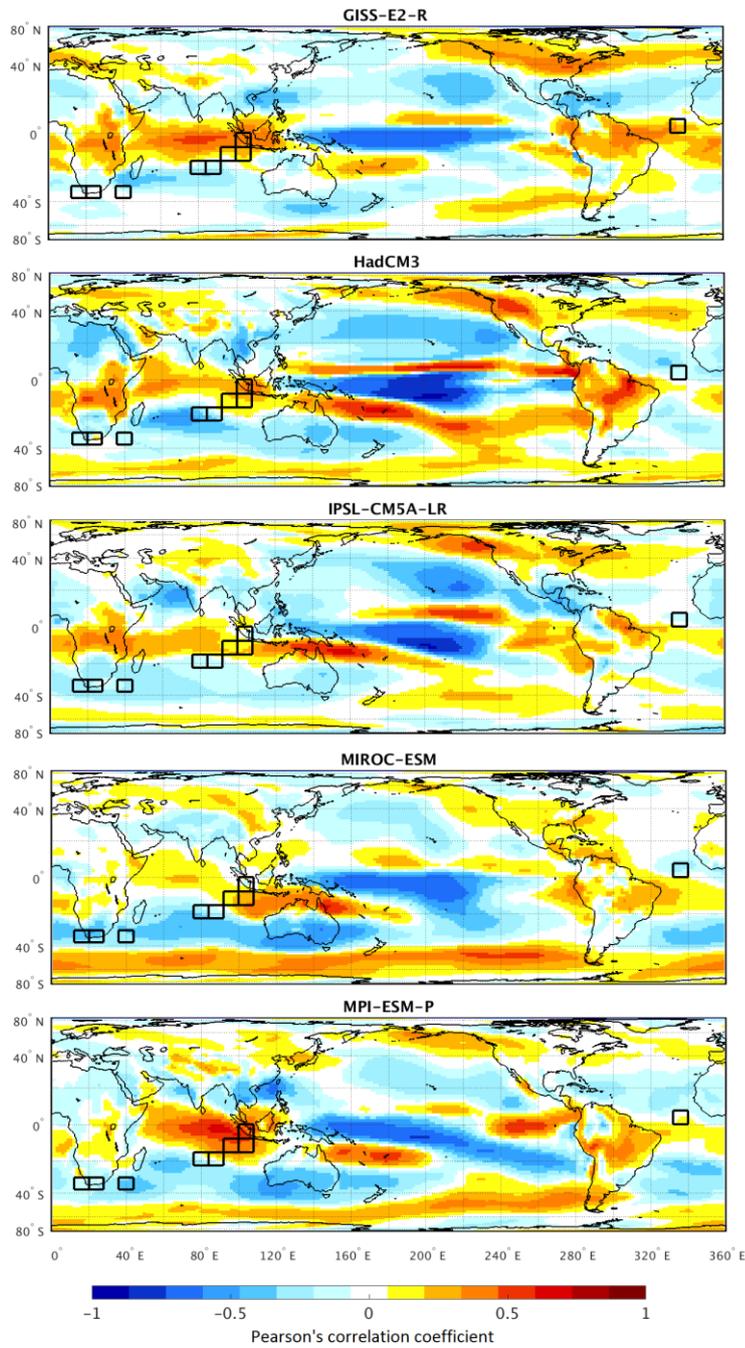


Fig. 7.14 Pearson's correlation coefficient between DJF zonal wind and DJF SOI from climate model simulations, 1900-1999. All time series detrended prior to calculation of correlations. Black rectangles indicate logbook grid boxes.

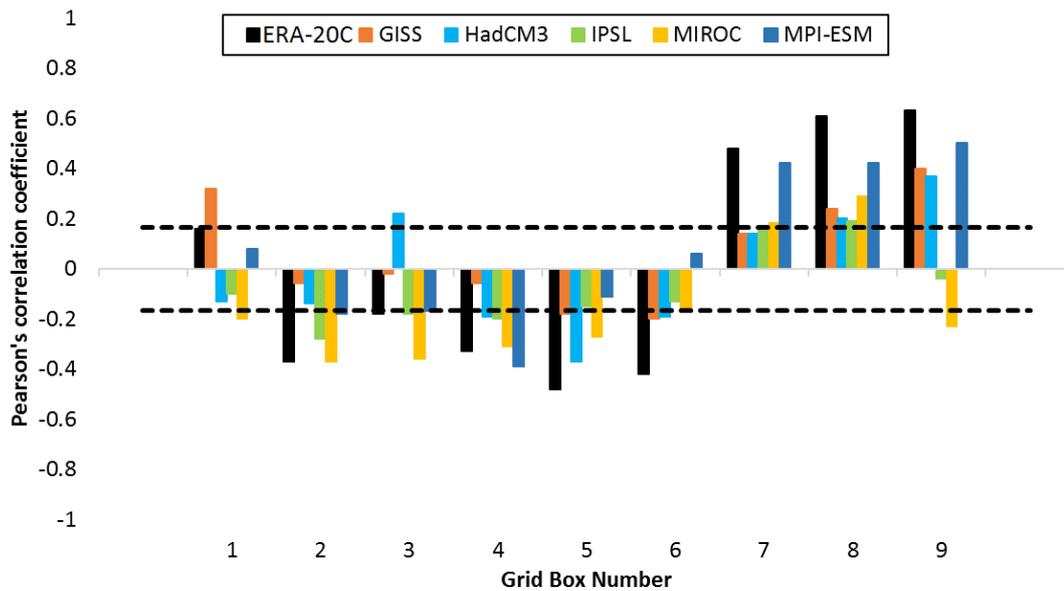


Fig. 7.15 Pearson's correlation coefficient between DJF Zonal wind per grid box (1-9) and DJF SOI from reanalysis and climate model simulations. Dashed lines indicates the 95% confidence value. Time series detrended prior to calculations of correlations. The results from this longer period (Table 7.5) agree with that from the shorter period (Table 7.4), in that the SOI-zonal wind correlations are significant in the highest number of logbook grid boxes in MIROC-ESM. Over this longer period, MPI-ESM-P is best at representing the teleconnection within grid boxes 7, 8 and 9, performing better than in the 1979-2004 period in this region. GISS-E2-R is the only model which has the sign of the correlation between zonal wind and the SOI the same as that from ERA-20C over the twentieth century.

Table 7.5 Number of logbook grid boxes with significant correlation (95% confidence level) of correct sign, and the number of grid boxes with the same sign correlation as ERA-20C, 1900-1999, between DJF zonal wind and the SOI.

	Number of sig correlation grid boxes	Number of grid boxes with same sign correlation as reanalysis
GISS-E2-R	5	9
HadCM3	6	7
IPSL-CM5A-LR	5	7
MIROC-ESM	8	7
MPI-ESM-P	6	8

Table 7.6 SOI-wind correlations ranks compared to ERA-20C, 1900-1999.

	Metric 1	Metric 2	Metric 3	Metric 4	Overall rank
GISS-E2-R	=1	=1	4	1	1
HadCM3	=1	4	=2	=3	5
IPSL-CM5A-LR	=1	=1	4	=3	3
MIROC-ESM	=1	5	1	=3	4
MPI-ESM-P	=1	3	=2	2	2

The models were then ranked for their performance over the entire twentieth century (Table 7.6). GISS-E2-R is ranked highest when comparing wind-SOI metrics over the twentieth century, followed by MPI-ESM-P and IPSL-CM5A-LR. Their strength and spatial pattern of correlation over this historical period are compared further in Chapter 8, when comparing to correlations over periods during the Last Millennium. The same analysis was carried out using the Niño 3.4 index instead of the SOI in order to explore the implications of using a different ENSO index on the overall results. The analysis can be found in the Appendix D (Fig. D4-D10; Table D5-D7). Some differences are found in the ranking of the models over the twentieth century when comparing the SOI–wind correlations to those from the Niño 3.4 index. However, GISS-E2-R ranks

first again and IPSL-CM5A-LR ranks third. MPI-ESM-P is ranked second by the SOI, whereas HadCM3 is ranked second by the Niño 3.4 index analysis, with MPI-ESM-P ranking joint third. However, as seen in Table 7.6, HadCM3 ranks worst when using the SOI.

GISS-E2-R and IPSL-CM5A-LR are ranked in the top three over both periods and were found to capture the SOI well in Chapter 6. MPI-ESM-P also performs well at capturing the zonal wind and its correlation with the SOI over the twentieth century within the predictor grid boxes. However, as GISS-E2-R and IPSL-CM5A-LR rank within the top three of both the SOI and the Niño 3.4 index rankings from Chapter 6, as well as ranking highly here, then they are selected to be analysed further and used over the LM simulations in Chapter 8, and to assess the stability of the SOI-wind teleconnection in Section 7.4.5.

7.3.5 Stability of the SOI-wind teleconnections over the twentieth century

The stability of the teleconnection between the SOI and the zonal wind within the predictor key boxes is assessed over the twentieth century within reanalysis and historical climate model simulations. The statistical relationship between the SOI and zonal wind is the basis for the logbook-based reconstructions presented in Chapter 2. The assumption of stationarity is a key to paleoclimate reconstructions, however in reality there is no guarantee that the relationship from which a reconstruction is built is stable over longer time periods than the initially assessed (Gallant et al. 2013). Therefore, here analysis of the initial correlation which focused on 1979-2014 within Chapter 2, is expanded over the twentieth century using a range of datasets.

Running correlations, using a 30-year window, were calculated over the twentieth century between the SOI and the DJF zonal wind in the nine predictor grid boxes. This was carried out initially using ERA-20C reanalysis (Fig. 7.16) and 20CR (Appendix Fig. D12). As noted in Section 7.1, there is larger uncertainty in the earlier half of the twentieth century within the reanalysis due to a reduced number and reduced quality of observations assimilated. Therefore, the conclusions drawn from the historical reanalyses are limited by this. Grid boxes 1, 2, and 3 show a gradual increase in the strength of the correlation over the twentieth century with all three having correlation values close to zero at the start of the analysis period. Grid boxes 4, 5 and 6 all show

a similar pattern in their running correlations, with a decrease in strength of their correlation coefficient centred around 1940's. These grid boxes are located in the Indian Ocean around 20°S (5 and 6) and the far South West of the Indian Ocean (GB 4). GB 8 and 9 also have a similar pattern of variation in correlation strength with time as each other. In some grid boxes the correlation between the two variables is not always significant throughout this period and in some cases changes sign for short periods. Grid boxes 8 and 9 show the most consistency significant teleconnection, suggesting higher stability within the reanalysis in this regions to the west of Sumatra and Java. The running-correlations were also calculated using 20CR, as an alternative reanalysis over the twentieth century, and similar conclusions are drawn using this reanalysis compared to ERA-20C (Appendix Fig. D11).

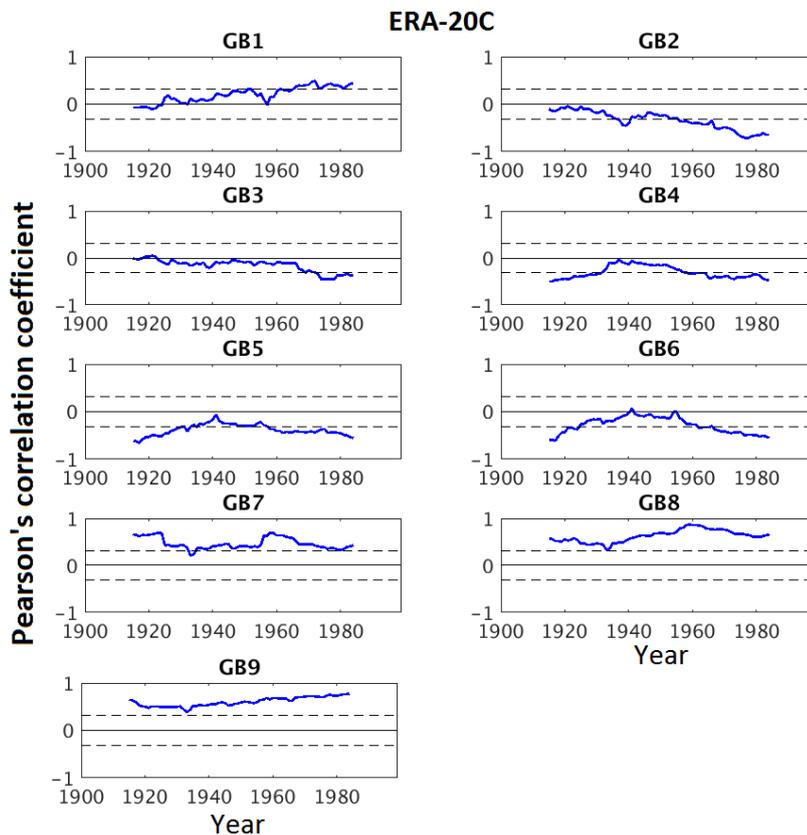


Fig. 7.16 30-year sliding/running correlation of DJF zonal wind and DJF SOI 1900-1999, ERA-20C. All time series were detrended before calculating running correlations. The year shown is the 15th year of the window, dashed line indicates significance at the 90% confidence level.

The stability of the ENSO teleconnections is regionally dependent, as found in previous research (Lewis and LeGrande 2015). A reduced strength in ENSO-precipitation teleconnection centred around 1940 was also found by Brown et al. (2016) in a regional analysis of Eastern-Australia. They found this reduction when correlating the Niño 3.4 index with precipitation observations from the Australian Water Availability Project over the extended boreal winter seasons (SONDJF), 1900-1999. Similar periods of teleconnection break down were also found in some of the historical climate models they analysed. They concluded that the reduction in the observed correlation was the result of a period with fewer ENSO events and weaker ENSO variance (Brown et al. 2016). In addition, the quality and quantity of the observational data also reduces further back in time as seen with reduced agreement between reanalysis datasets in the earlier part of the twentieth century (Chapter 5). This could also cause a reduction in the strength of correlation between observational and reanalysis datasets.

The running correlations were then carried out using the two selected historical climate model simulations (Fig. 7.17) and all five historical simulation simulations can be seen in Appendix D, Fig. D12. The correlations are generally weaker than those found in ERA-20C and the models suggest that the correlations between zonal wind and the SOI are not stable over this time period with the sign of the correlations within a given grid box changing over the period, and often not significant at the 95% confidence level. As found in section 7.4.4, it is a challenge for the models to capture the zonal wind at the predictor grid box scale, therefore it is also challenging for them to capture the ENSO-zonal wind teleconnection at this smaller spatial scale.

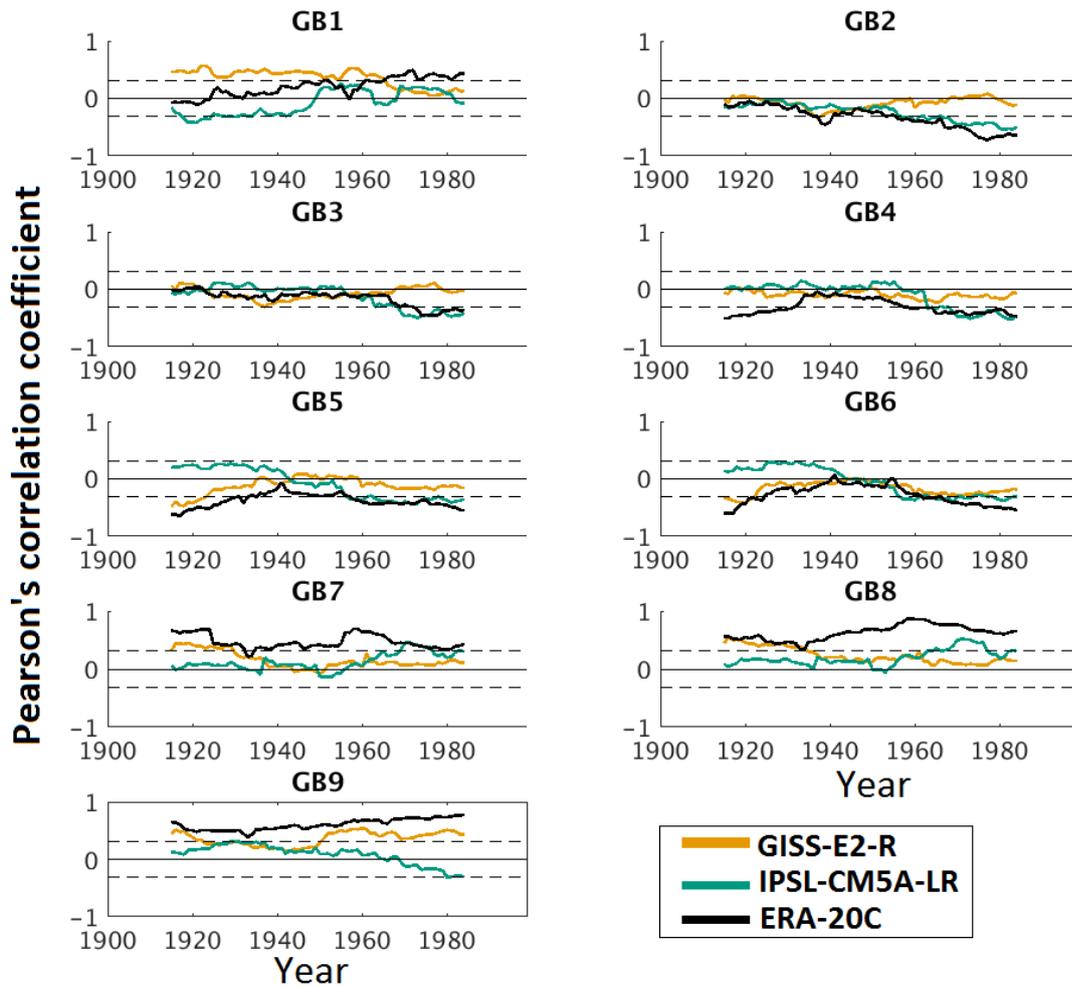


Fig. 7.17 30-year sliding/running correlation of DJF zonal wind and DJF SOI 1900-1999, GISS-E2-R (orange) and IPSL-CM5A-LR (green) climate model simulations and ERA-20C (black). All time series were detrended before calculating running correlations. The year shown is the 15th year of the window, dashed lines indicate significance at the 90% confidence levels.

7.4 Summary

The results from this section are discussed further in Chapter 8 and Chapter 9, with reference to stability of the ENSO-wind teleconnection over the LM. However, a summary of the main results from this chapter is provided here. The spatial patterns of mean zonal wind climatology were found to be similar within the three reanalyses (ERA-Interim, ERA-20C and 20CR) on a large scale (Fig. 7.5), which along with similarities in MSLP (Fig.5.4), suggests that they all capture large-scale atmospheric

circulation patterns well. The weak westerly zonal wind anomalies over the Maritime Continent are associated with the area of relatively low pressure to the south east of this region due to the ascending limb of the Walker Circulation resulting in surface convergence, as explained in Chapter 1. In contrast, in the eastern and central Pacific, there are easterly (negative) zonal wind anomalies found within the reanalysis (Fig. 7.5), supporting the schematic of normal conditions as depicted in Fig. 1.1. However, a key challenge within this chapter was the analysis of zonal wind at a smaller regional resolution, matching the resolution of the nine predictor grid boxes used in Chapter 2 for the logbook-based ENSO reconstruction. The reanalysis captured similar mean zonal values in all nine grid box regions, however it was found that mean zonal winds were stronger in ERA-20C than ERA-Interim, and in most of 20CR compared to ERA-Interim (Fig. 7.6).

Large-scale biases in mean zonal wind compared to ERA-Interim were found in the historical simulations (Fig. 7.8), and when looking at the mean zonal wind climatology of the climate models, there was reduced agreement at the grid box scale compared to the reanalysis (Fig. 7.6; Fig. 7.9). Some models even produced a mean zonal wind of the opposite direction to that from the reanalysis within certain grid boxes. It is a challenge for models to capture the mean zonal wind at this grid box scale. In terms of climatology, it is difficult for models to capture the exact locations of spatial patterns compared to observations and reanalysis data, as seen in Chapter 6, with the locations of warm and cold surface temperatures over the Pacific in the models not always matching up spatially with the patterns seen in the observations and reanalysis (Fig. 6.8). This is the case for other variables including MSLP as seen in Fig. 6.3, and for the zonal winds analysed here.

The strength of the spatial correlation between the zonal wind and the SOI were assessed over the period 1979-2004, and then over the entire twentieth century (1900-1999) within the nine predictor grid boxes from the ERA-20C and the historical climate model simulations. Fig. 7.14 showed that the general spatial pattern of correlations from the historical climate simulations were reasonably similar to those from ERA-20C (Fig. 7.13) on a large-scale, such as a region of negative correlation between SOI and zonal wind over much of the central equatorial Pacific, and positive correlation over the Maritime Continent and northern Indian Ocean. The correlations

from the models are weaker than those from the reanalysis, a common result from previous ENSO-teleconnection spatial correlations analysis (Lewis and LeGrande 2015; Brown et al. 2016). However, when these correlations were reduced to grid box scale, disagreements between the models and the reanalysis, and between the models were more apparent (Fig. 7.15). It is a challenge for models to capture the strength and in some cases the direction of the correlation between zonal wind and the SOI at this spatial resolution.

The models were ranked on their ability to capture the sign and strength of this correlation, and the mean zonal wind within the nine predictor grid boxes, as well as the DJF SOI from the rankings provided by Chapter 6. It was found that over the 1900-1999 period, GISS-E2-R ranks first out of the five climate models analysed, with MPI-ESM-P placing second and IPSL-CM5A-LR third (Table 7.6). GISS-E2-R also ranks top when using the Niño 3.4 index instead of the SOI, and again IPSL-CM5A-LR ranks third. IPSL-CM5A-LR and GISS-E2-R were ranked best out of the five climate models over the twentieth century in Chapter 6. These two models also perform well over this period in the analysis carried out in this chapter and are thus used in Chapter 8. However, GISS-E2-R is performing best when looking at the correlation between the SOI and the zonal wind. Both were used along with the ERA-20C reanalysis in the analysis of the stationarity of the teleconnections over the twentieth century, to provide consistency throughout.

The stability of the SOI-wind teleconnection within the predictor grid boxes was then assessed in order to test the assumption of stationarity which underlies paleo-climate reconstructions and the logbook-based reconstruction presented in Chapter 2. It was found that using ERA-20C, the ENSO-zonal wind teleconnection was not stable over the twentieth century within all of the nine predictor grid boxes (Fig. 7.16). It is suggested that future reconstructions could focus on regions with a higher degree of stability, such as grid boxes 7, 8 and 9. However, a source of additional uncertainty surrounding this is the use of the longer time-span reanalysis, which has fewer observations in the early period, and thus significant differences were found between the two reanalyses when compared over the earlier part of the last century. The climate models suggest that over this 100-year period the strength of the correlation weakens in certain period and is not statistically significant throughout (Fig. 7.17). It is found

that stability of the teleconnection also depends on the grid box region used, with the correlation between zonal wind and the SOI mostly over the 90% confidence threshold in grid box 7, 8 and 9, suggesting a higher degree of stationarity in the region containing these grid boxes compared to in the locations of the other predictor grid boxes. This stationarity is explored further in Chapter 8 which looks at this relationship over the Last Millennium, and then the implications of these results are discussed within Chapter 9.

8. ENSO-teleconnections and ENSO diversity over the Last Millennium, 850/51-1848/49 AD

8.1 Introduction

In this chapter, the Last Millennium climate simulations of IPSL-CM5A-LR and GISS-E2-R are used to investigate the stability of ENSO-wind teleconnections and to examine ENSO diversity in the past. In Chapter 6 it was found that IPSL-CM5A-LR ranks highly for all ENSO features analysed including the SOI, the Niño 3.4 index and ENSO diversity, placing within the top two of the five models analysed for all of these features over 1900-1999. The spatial composite plots from the model were also found to support these rankings (Fig. 6.6; 6.7; 6.10; 6.11). Chapter 7 found that IPSL-CM5A-LR ranked top over 1979-2004 for SOI-wind correlations within the predictor grid boxes, and third over the entire twentieth century. For the SOI and ENSO diversity, GISS-E2-R ranked first over the twentieth century (Table 6.8; Table 6.14). This model also ranked first for SOI-wind teleconnections over the twentieth century within the predictor grid boxes (Table 7.6). Therefore, this chapter uses both GISS-E2-R and ISPL-CM5A-LR as they performed best at the metrics under investigation and using them throughout the entire chapter provides consistency in the analysis.

These selected models are used to investigate climatic relationships over the Last Millennium (LM), which spans the period 850 to 1849 AD. The analysis here covers the DJF seasons 850/851 AD to 1848/49 AD, henceforth referred to as 850-1848 relating to the year of the December. The LM simulations are forced by natural forcings such as orbital parameters, solar and volcanic forcings, variations in greenhouse gases and changes in land use and cover, see Chapter 6, Table 6.2 (Schmidt et al. 2011). They provide a 1000-year period in which climatic variations can be investigated, thus the LM simulations are able to identify long-term changes within the natural climate system. Chapters 6 and 7 showed that the historical climate model simulations had some biases compared to reanalysis data, and are therefore not perfect representations of the climate system. For example, GISS-E2-R was found to have a warm bias in surface temperature in the eastern and central equatorial Pacific, which resulted in a mean Niño 3.4 index 1.5 K higher than that from ERSSTv4 dataset over the twentieth century (Table 6.10), whereas, IPSL-CM5A-LR has a warm surface

temperature bias which is more restricted to the far eastern Pacific (Fig. 6.8), a common bias found in CMIP5 models (Michael et al. 2013; See Fig. 6.2). However, despite their biases, the two LM simulations which are used here can provide an indication of the behaviour of natural climate phenomena such as ENSO in the past. The results from the climate model analysis can then be discussed alongside proxy records in order to provide the most comprehensive understanding of the climate within this pre-instrumental period. They can also be used to inform the logbook-based ENSO reconstructions by providing an assessment of the stability of the ENSO-wind teleconnections in the logbook predictor regions.

8.1.1 ENSO in the Last Millennium climate model simulations

A number of studies have previously assessed ENSO using the LM simulations, looking at ENSO frequency, ENSO response to external forcing and ENSO teleconnections (Coats et al. 2013; Lewis and LeGrande 2015; Brown et al. 2016; Hope et al. 2016; Le 2017; Stevenson et al. 2017). A common finding over the LM is an association between ENSO and volcanoes (Brown et al. 2016; Hope et al. 2016; Le 2017). This was previously discussed in Chapter 3 with respect to the historical and proxy record. Fig. 8.1 shows an ENSO index calculated from a range of LM simulations based on the empirical orthogonal functions of the DJF SSTA in the Niño 3.4 region (Le 2017). Changes in ENSO over the LM are highly model-dependent, however a key conclusion from Le (2017) is that there is a significant response of ENSO to volcanic forcings, with cooling associated with the major volcanic eruptions of the LM. This supports findings from previous studies and also suggests that the majority of the CMIP5 LM simulations are capable of replicating the observed ENSO response to volcanic forcing (Adams et al. 2003; Le 2017). This association between ENSO and the major volcanic eruptions of the LM has also been discussed in other studies, which support the conclusions of Le (2017) (Brown et al. 2016; Hope et al. 2016). However, these all focus on surface temperature as an indicator of ENSO, therefore, as the SOI is the key ENSO index used here, an assessment of this association between the DJF SOI and the major volcanic eruptions is carried out to investigate if, using the SOI, this association is still apparent within the climate model simulations.

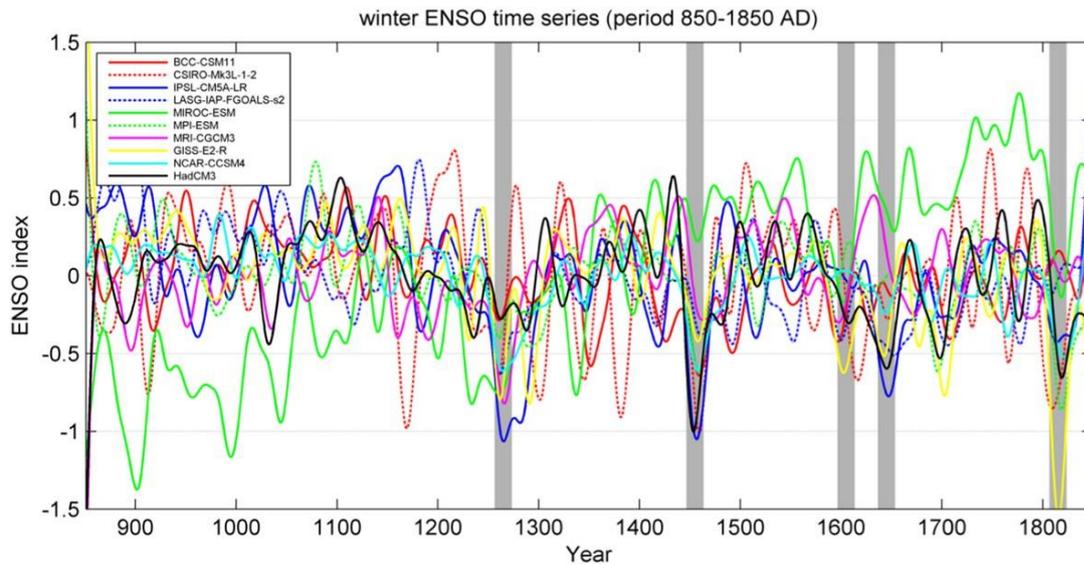


Fig. 8.1 ENSO indices derived from a number of Last Millennium climate model simulations (850-1850 AD), based on the EOF of the DJF Niño 3.4 index. The grey bars indicate major volcanic events of the LM (1258, 1452, 1815). Fig. from Le (2017).

The frequency of ENSO events over the LM was the focus of Hope et al. (2016). It was found that the spectrum of ENSO over the LM varies from model to model, with noticeable inter-model differences in ENSO frequency and variance (Hope et al. 2016). The instrumental ENSO frequency is between 2-7 years (Allan et al. 1996). During the historical simulations, Hope et al. (2016) found that GISS-E2-R and IPSL-CM5A-LR were in the top three of the six models used in their study in terms of representing the observed Niño 3.4 spectra. This is in agreement with the high ranking of these models of those used here in Chapter 6. Fig. 8.2 shows the power spectrum of Niño 3.4 from the LM simulations of these models (Hope et al. 2016), with GISS-E2-R peaking at 3-4-year band and IPSL-CM5A-LR suggesting a lower frequency of events, more in line with the frequency observed in the instrumental era.

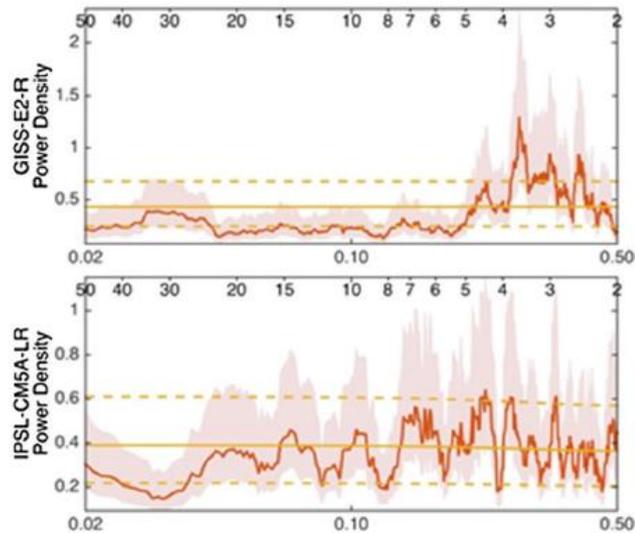


Fig. 8.2 Power spectrum density of GISS-E2-R (upper) and IPSL-CM5A-LR (lower) for the Niño 3.4 index from LM simulations (850-1850). Solid red line indicates model and the faded red shading is the 95% confidence levels. Fig. from Hope et al. (2016).

There has been little investigation into ENSO diversity in LM climate simulations (Stevenson et al. 2017). Chapter 6 highlighted some recent work which assessed ENSO diversity within the historical climate simulations from CMIP5 (Kim and Yu 2012). A number of studies have used proxy records to assess ENSO diversity (see Chapter 4). Here, the LM simulations of GISS-E2-R and IPSL-CM5A-LR are used to identify EP and CP El Niño events and to investigate the ratio of CP-EP events over the LM. The spatial patterns of EP and CP events are assessed in comparison to those from the twentieth century in Chapter 6.

8.1.2 ENSO-teleconnections in the LM simulations

As discussed in Chapter 7, a number of previous studies have assessed ENSO-teleconnections using LM simulations to assess the stability of the relationship between ENSO and variables such as precipitation and temperature (Coats et al. 2013; Lewis and LeGrande 2015; Brown et al. 2016). Lewis and LeGrande (2015) found that the surface temperature anomalies during El Niño and La Niña events were stronger in the historical period than the Last Millennium period. Using seven CMIP5 models spanning the LM, spatial composite plots of surface temperature during ENSO events showed that the models capture the spatial patterns of global surface temperature anomalies best in the tropical Pacific, with largest differences between

models found in more ENSO-remote areas such as North America and the South Pacific (Lewis and LeGrande 2015). The stability of ENSO-temperature teleconnections across the tropical Pacific were found to vary regionally, with weaker correlations between west Pacific and the Niño 3.4 index compared to the central and eastern Pacific. The western Pacific ENSO-temperature teleconnection was found to be more variable than that in the central and eastern Pacific across the ten 100-year epochs of the LM (Lewis and LeGrande 2015). The different models used vary in their representation of the magnitude of these teleconnections across the LM.

The ENSO-teleconnections in Australia were examined over the LM by Brown et al. (2016), using the 30-year running correlation between September-February rainfall and the Niño 3.4 index, and the 30-year running variance of the Niño 3.4 index (Fig. 8.3). It was found that the temporal correlation between precipitation and the Niño 3.4 index is not stable throughout the LM, with models showing periods in which the teleconnection weakens significantly and in some cases changes sign (Brown et al. 2016). ENSO variance was found to be in part associated with the strength of the teleconnection, with the teleconnection weaker when ENSO variance was low. The stability and magnitude of the teleconnection also varies depending on the model used. The spatial pattern of ENSO-rainfall teleconnection in Australia was found to be similar when comparing LM to historical simulations, although variations in the pattern over time are found within the LM simulations (Brown et al. 2016).

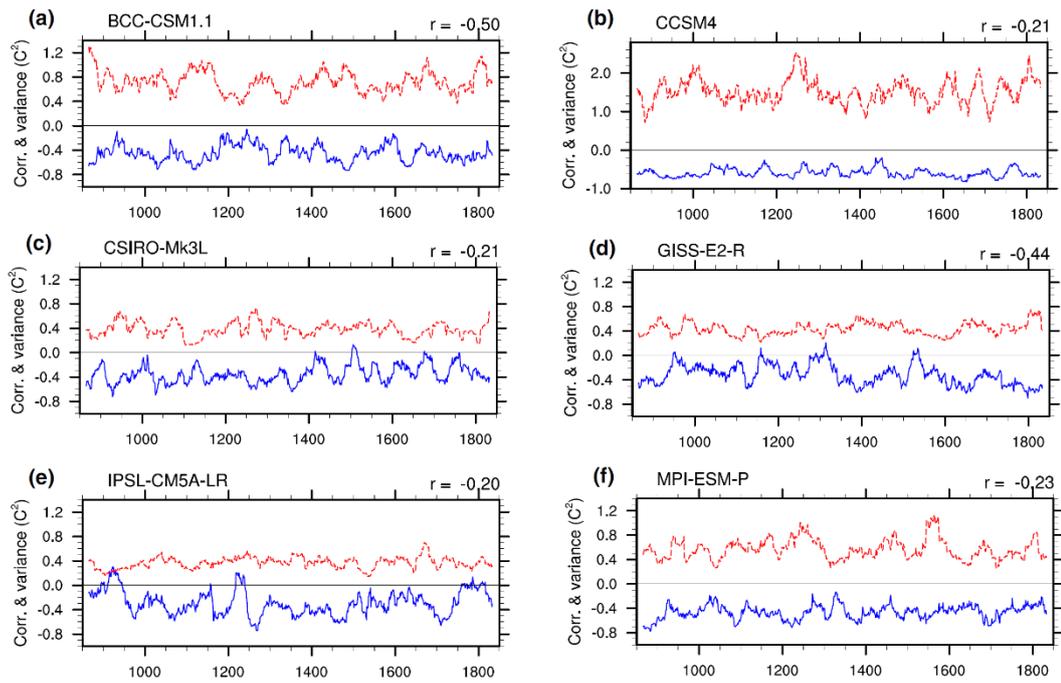


Fig. 8.3 30-year running correlations of Sept-Feb rainfall and the Niño 3.4 index (blue) and the 30-year running variance of the Niño 3.4 index (red). Correlations between the time series shown in the top-right corner of each panel. Fig. from Brown et al. (2016).

The North American ENSO-teleconnection was investigated by Coats et al. (2013) and was found to have a large variation in the strength and the character of the correlations over time within seven different LM model simulations. Simulated ENSO dynamics and the teleconnection between ENSO and North America were found to differ between models and in comparison to observations (Coats et al. 2013). A number of additional studies use LM simulations to assess various climatic phenomena, using similar approaches to those used in ENSO investigations. Klein et al. (2016) compared simulations of East African hydroclimate using LM simulations and found differences in the timings of fluctuations in hydroclimate between the models, due to internal variability of the individual models. They concluded that using a multi-model mean to analyse variation in local hydroclimate would conceal weak responses, and therefore individual models should be selected and analysed (Klein et al. 2016). The latter approach is taken within this chapter.

The LM climate simulations of seven climate models were used by Tejavath et al. (2017) to assess the relationships between the Indian Summer Monsoon Rainfall

(ISMR) and ENSO. Focusing on the two major climatic periods of the Medieval Warm Period, MWP, (1000-1199 AD) and the Little Ice Age, LIA, (1550-1749 AD), they found that the correlation between the simulated ISMR and Niño 3.4 index is higher during the MWP than the LIA. However, they found a large range in the correlation coefficients between these variables from the different model simulations used and they concluded that their results should be treated as qualitative (Tejavath et al. 2017). This highlights the difficulty of making robust conclusions from multiple climate model simulations. Tejavath et al. (2017) also found that during the MWP there were relatively higher numbers of El Niño events compared to La Niña events, and during the LIA relatively lower numbers of El Niño, associated with mean summer temperature differences during these two periods. This will be investigated further here, when using the SOI as the key indicator of ENSO.

8.1.3 ENSO variability and events, ENSO-wind teleconnection and ENSO diversity within two LM climate simulations

This chapter analyses the LM simulations of the two selected climate models in terms of their DJF SOI index and frequency of ENSO events, the spatial composite anomaly patterns of surface temperature, MSLP and zonal wind during ENSO events and finally the strength and the stability of the correlation between the DJF SOI and the zonal wind. This is followed by an analysis of ENSO diversity within the LM climate simulations.

Firstly, the analysis of the ENSO-wind teleconnections discussed in Chapter 7 will be extended back over the course of the Last Millennium simulations. As mentioned in Chapter 7, previous ENSO teleconnections over the LM have focused on temperature and precipitation over selected regions such as Northern Australia or North America (Coats et al. 2013; Brown et al. 2016). Here, a focus on winds provides a more direct link to the key meteorological variable available from the logbook data. Running correlations over the entire LM period assess the stability of the teleconnections, whereas spatial correlation plots are used to compare the strength of the correlations over two 100-year periods and provide global context to the correlations. However, the focus here, as discussed in Chapter 7, is largely on the predictor grid boxes used in Chapter 2 for the historical ENSO reconstruction, but with some discussion of the larger spatial patterns due to the challenge for climate models to fully capture spatial

patterns at the grid box scale resolution. Correlations over 100-year periods of the LM provide a direct comparison to the strength of the correlations during the LM periods compared to the historical period. The period 1750-1848, received special attention due to its overlap with the logbook-based reconstruction time period. It should be noted that this period has one less year than additional 100-year periods, but is still referred to as a 100-year period for simplicity. In addition, to provide a 100-year period in the LM for comparison to make the conclusions more robust, a period centred during the Medieval Warm Period is also used (Tejavath et al. 2017). The MWP spanned 950-1350, with previous modelling studies using the period 1000-1199 to assess the middle portion of this period (Tejavath et al. 2017). Here, the 100-year period 1050-1149 is used, covering the centre of this period.

Secondly, the diversity of ENSO in terms of the frequency of EP and CP El Niño events, and the ratio of CP to EP events is investigated. Previous studies have already used LM simulations to assess the variability and frequency of ENSO in terms of El Niño and La Niña events (Brown et al. 2016; Hope et al. 2016). Therefore, here the focus is on extending this analysis further to look at the pattern of EP and CP El Niño throughout the LM period. As well as looking at the frequency, the spatial patterns of temperature during EP and CP events will be assessed to investigate the stability of their spatial anomaly patterns, with a focus on 1750-1848 and the 100-year period within the MWP, 1050-1149.

This chapter is organised as follows: *Section 8.2* briefly summarises the methods used within this chapter, many of which are employed in previous chapters. *Section 8.3* presents the results of this Last Millennium analysis focusing on the SOI, the ENSO-zonal wind relationship and ENSO diversity within the climate model simulations. *Section 8.4* summarises the key findings of this chapter, which are then discussed fully with the key findings of the preceding chapters within Chapter 9. The work within this chapter addresses Aim 4, Objectives 4A, 4B and 4C, as outlined in Chapter 1.

8.2 Methods

This chapter uses methods similar to those described in Chapters 6 and 7, applying them to Last Millennium climate model simulations. The output from GISS-E2-R and IPSL-CM5A-LR LM simulations were used as they were found to perform best at

representing ENSO via the SOI and ENSO diversity in Chapter 6. They were also found to capture the relationship between SOI-zonal wind relatively well in Chapter 7. The use of climate models to explore the period LM is not without its limitations, as will be discussed in Chapter 9. However, they provide an indication of climatic behaviour during a period of relatively well known natural forcings and using model simulations of a suitable temporal and spatial resolution to provide a realistic representation of climate interactions. Where appropriate, data were normalised relative to the 1750-1848 AD mean and were linearly detrended prior to calculation of correlation coefficients.

Two climate models are used throughout this chapter, GISS-E2-R and IPSL-CM5A-LR, which were selected from the results of Chapters 6 and 7 as they performed consistently well in the twentieth century with respect to ENSO properties and ENSO-wind teleconnections. Therefore, the conclusions from this chapter are based around the performance of these two models and the outputs are model-dependent, thus cannot be assumed to be universally representative of all climate models. However, the results from previous chapters suggest they might provide the best indication of ENSO characteristics available from the current suite of LM climate model simulations.

8.3. Results

8.3.1 The DJF SOI in the Last Millennium

Firstly, the DJF SOI from the two models over the Last Millennium were assessed in order to investigate the temporal variability of this index within model simulations. The DJF SOI and the 30-year moving mean from GISS-E2-R and IPSL-CM5A-LR from 850/51-1848/49 AD are shown in Fig. 8.4. There is similarity between the SOI presented here and ENSO indices created from these models in previous papers (Lewis and LeGrande 2015; Brown et al. 2016). The standard deviation of the DJF SOI over the LM period is 7.31 from GISS-E2-R and 7.67 from IPSL-CM5A-LR, suggesting similar ENSO amplitudes from the two models, and similar to that of the instrumental record (1900-1999) which was found to have a standard deviation of 7.91 (Chapter 5).

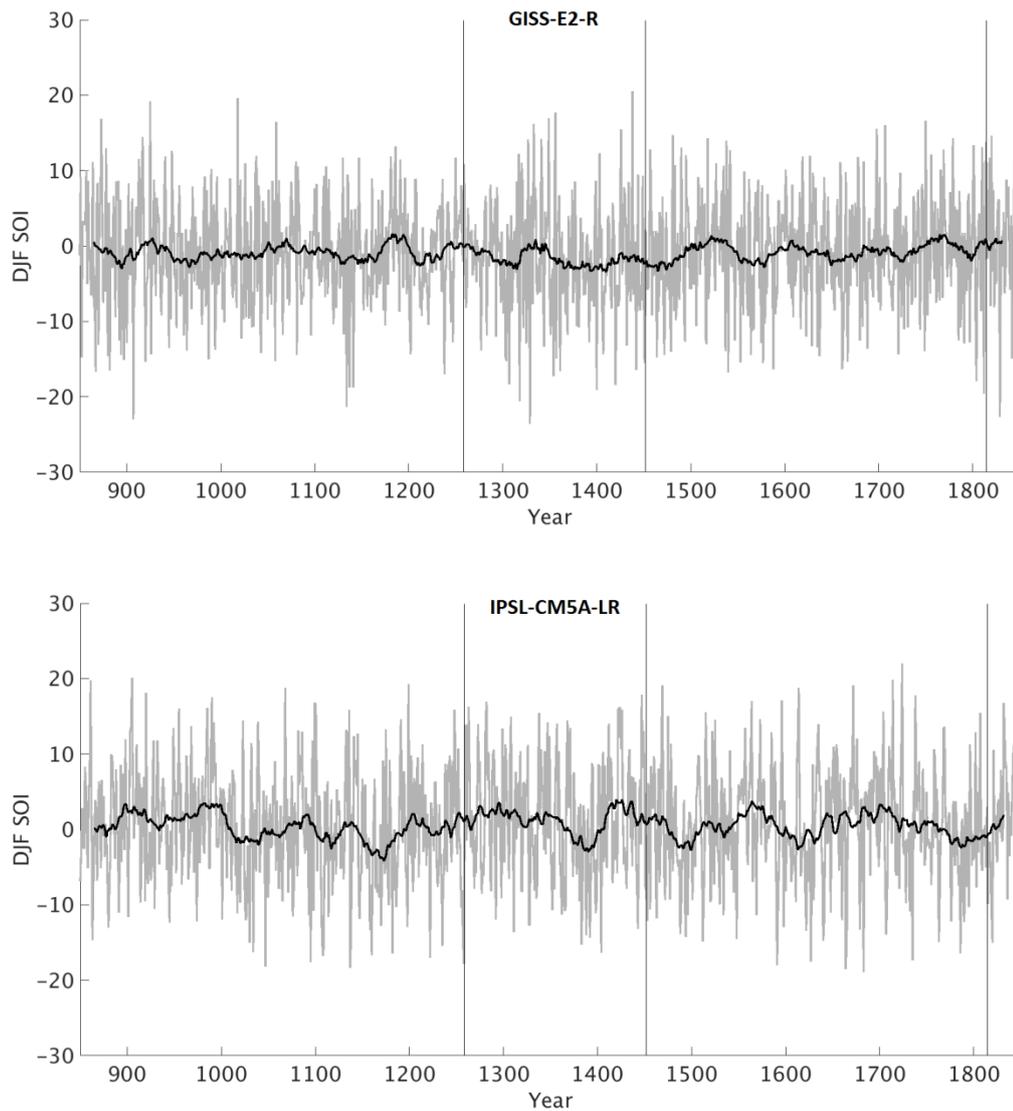


Fig. 8.4 The DJF SOI from GISS-E2-R (upper) and IPSL-CM5A-LR (lower) LM simulations relative to the 1750/51-1848/49 mean. Bold line indicates 30-year moving mean. Vertical lines indicate major volcanic eruptions 1258, 1452, 1815 as in Brown et al. (2016).

Previous studies have suggested an association between ENSO and major volcanic eruptions, with a decrease in the Niño 3.4 index suggesting La Niña conditions following major volcanic events (Brown et al. 2016; Le 2017). Using the SOI as an indicator of ENSO, this relationship is less clear and is not evident from the two model simulations shown in Fig. 8.4, which shows the three main volcanic events of the LM occurring in 1258, 1452 and 1815 (Brown et al. 2016; Le 2017). When looking at a smaller temporal resolution, focusing on the SOI in the years surrounding these three

events in more detail, no clear pattern emerges from these two models (See Appendix Fig. E1 and Fig. E2). This suggests that part of the signal found in previous studies is possibly due to the influence of a volcanic-induced drop in global temperatures, which is more likely to be identified when using temperature-based ENSO indices than the SOI. However, Brown et al. (2016) argue against this, stating that previous studies using the SOI have also seen the signal associated with volcanic events (Adams et al. 2003). However, Adams et al. (2003) used an SOI obtained from proxy records rather than from climate model simulations.

The number of El Niño and La Niña events per 100-year period of the Last Millennium from both of the models was calculated using the one standard deviation threshold from the DJF SOI, as in previous chapters (Appendix Table E1, Fig. 8.5). The frequency of ENSO events in the instrumental record is around 2-7 years (Allan et al. 1996; Hope et al. 2016). The average frequency of ENSO events in the 100-year periods of the LM is 3.2 years in both GISS-E2-R and in IPSL-CM5A-LR, therefore within the same broad frequency range as the instrumental period. Chapter 6 found that between 1900-1999, using the instrumental DJF SOI one standard deviation threshold, there were 15 El Niño events and 17 La Niña events, as outlined in black on Fig. 8.5. The number of events from the two models over the various 100-year periods of the LM ranges from 13 to 18 (12 to 20) El Niño events and 12 to 19 (13 to 19) La Niña events (note IPSL-CM5A-LR results are in brackets, GISS-E2-R out of brackets). Thus, the twentieth century frequency lies within this range. This suggests that the frequency characteristics of ENSO are similar within the LM simulations and historical instrumental record. There is also agreement with the frequency of events suggested by the power spectrum density analysis of Hope et al. (2016) for GISS-E2-R, which stated that over the LM the spectral peak of ENSO events was between 3-4 years (Fig 8.2).

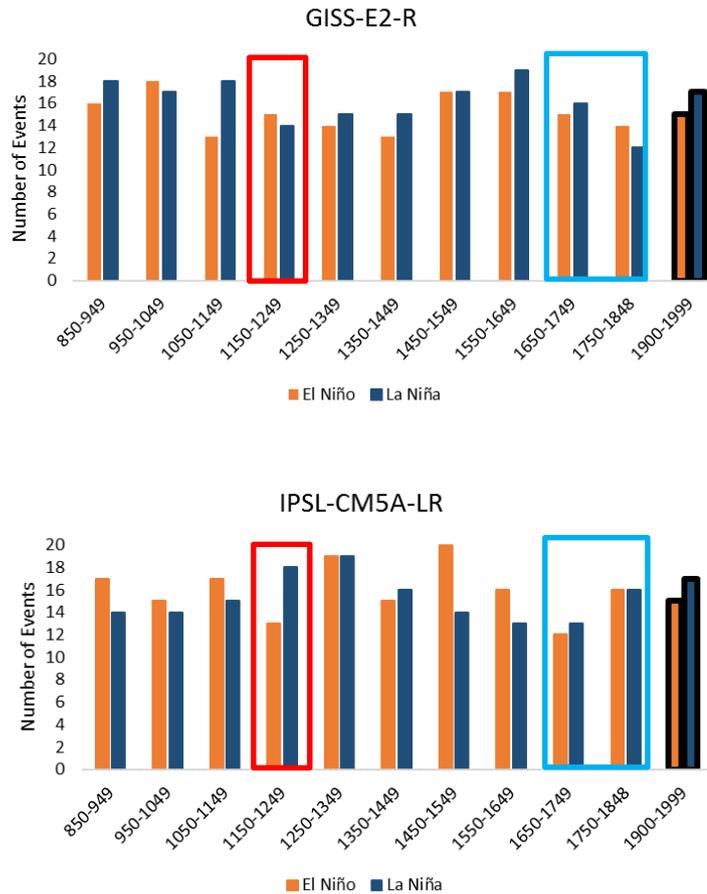


Fig. 8.5 Number of SOI-defined ENSO events per 100-year period of LM climate model simulations from GISS-E2-R (upper) and IPSL-CM5A-LR (lower), plus the number of events from the instrumental SOI over the twentieth century, indicated by data outlined in black. The 100-year period for the MWP is outlined in red, while the 200-years covering the LIA is outlined in light blue.

Comparing the LM simulations, Tejavath et al. (2017), found that during the MWP (1000-1199 AD) there were a relatively higher number of El Niño events compared to La Niña, and during the LIA (1550-1749) a relatively lower number of El Niño events than La Niña. Using a 100-year period within the MWP (1050-1150), there are more El Niño than La Niña events found in IPSL-CM5A-LR, but relatively fewer found in GISS-E2-R (Fig. 8.5). Also looking at events from 1550-1750 (LIA) GISS-E2-R suggests more La Niña than El Niño events, but IPSL-CM5A-LR does not identify this for 1550-1650. Therefore, within the two climate model simulations, the general pattern suggested by Tejavath et al. (2017) is not seen. One reason for this is the use of the SOI rather than a temperature based index, and therefore the ENSO index is less

influenced by temperature changes. The relative number of El Niño to La Niña events oscillates between century and in different ways between the two models. This suggests that on the century-scale the differences in the El Niño to La Niña ratio cannot be fully monitored.

Using the final period of the LM simulations, 1750-1848, the El Niño and La Niña events defined in this period are given in Table Appendix E1. This period is used to look in more detail at the behaviour of the models. The number of events over this period in IPSL-CM5A-LR, is close that found in the instrumental record, whereas it is lower in GISS-E2-R especially for La Niña events. Similar analysis for the MWP 100-year period can be found in Appendix E which provides useful support to the findings presented here and are discussed further later. Comparison between the two models, and with the spatial composite plots from the historical periods, 1979-2013 and 1900-1999 (Chapter 6, Fig. 6.6; Fig. 6.7 and Fig. B2: Fig. B3), provide an indication of the characteristics of ENSO events from the model simulations during this period of the LM which overlaps with the period of the historical logbook-based reconstruction. The spatial composites for ENSO events 1900-1999 from the two historical model simulations of GISS-E2-R and IPSL-CM5A-LR are shown for reference in Fig. 8.6, Fig. 8.8 and Fig. 8.10. Whereas the composite anomaly plots of surface temperature, Mean Sea Level Pressure and zonal wind are presented from the LM simulations 1750-1848 in Fig. 8.7, Fig. 8.9, Fig. 8.11.

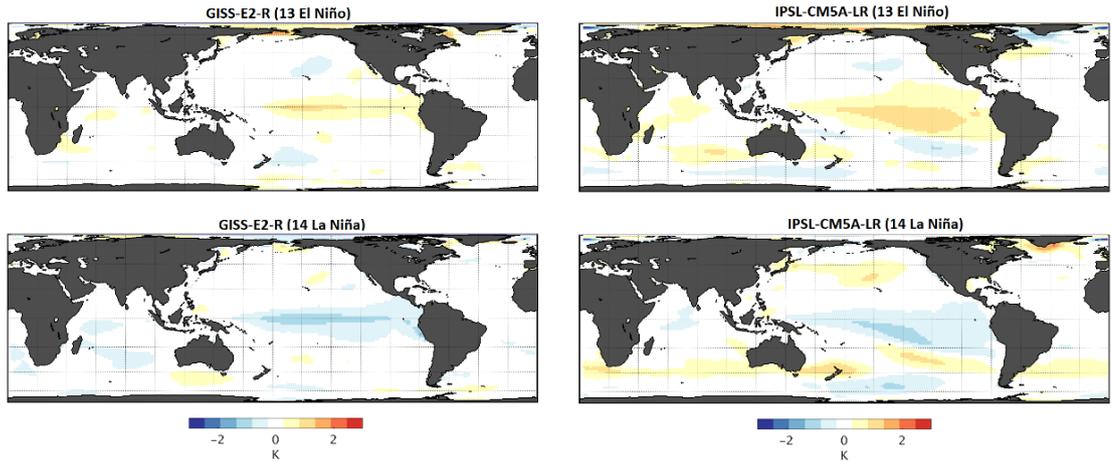


Fig. 8.6 El Niño and La Niña surface temperature (K) anomaly composites defined by one standard deviation of the DJF SOI from historical simulations, 1900-1999 from GISS-E2-R and IPSL-CM5A-LR. Anomalies based on 1900-1999 DJF mean. Numbers in brackets indicate number of events used to calculate each composite plot.

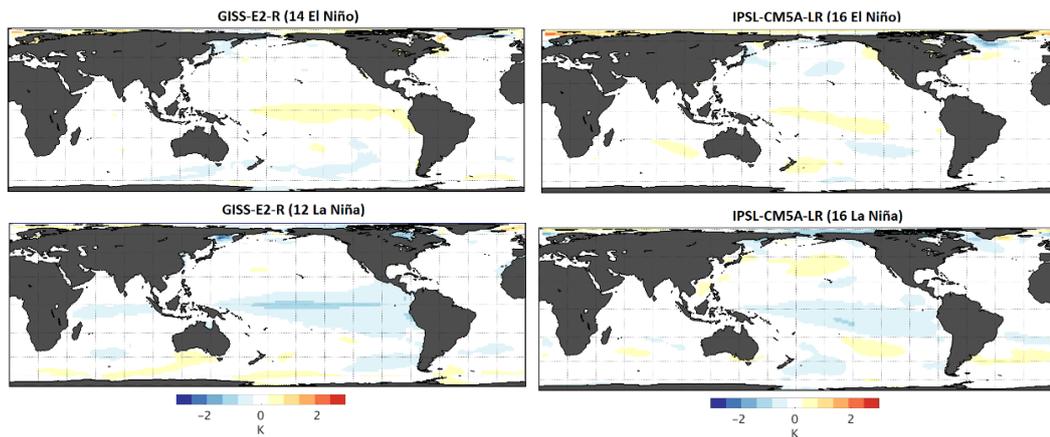


Fig. 8.7 El Niño and La Niña surface temperature (K) anomaly composites defined by one standard deviation of the DJF SOI from LM simulations, 1750-1848 from GISS-E2-R and IPSL-CM5A-LR. Numbers in brackets indicate number of events used to calculate each composite plot.

The surface temperature composite anomaly plots (Fig. 8.7) show that GISS-E2-R captures the spatial patterns of surface temperature anomalies in the equatorial Pacific expected during El Niño and La Niña events compared to IPSL-CM5A-LR over the 1750-1848 period. The anomalies are generally weaker than those found in the historical period (1900-1999) (Fig. 8.6), this is especially pronounced in IPSL-CM5A-LR. This result is similar to the findings of reduced strength in composite anomaly

plots in the LM simulations compared to the historical period in Lewis and LeGrande (2015). However, it was found in Chapter 6 that over the twentieth century the region of surface temperature anomaly in the equatorial Pacific is more zonal in GISS-E2-R than IPSL-CM5A-LR, for both El Niño and La Niña events, with strongest anomalies south of the equator in IPSL-CM5A-LR rather than on the equator as found in GISS-E2-R, a pattern which is closer to that found in the reanalysis composites (Chapter 6). This is also found to be the case in Fig. 8.7, and in the MWP (Appendix Fig. E3). In both of the models' composites, the La Niña negative surface temperature anomalies are stronger and covers a wider area than the El Niño positive anomalies. Comparing the composites to the historical period, it can be concluded that the models support similar, but weaker, spatial patterns of surface temperature anomalies associated with ENSO years 1750-1848 from the LM simulations and in the MWP presented in Appendix Fig. E3.

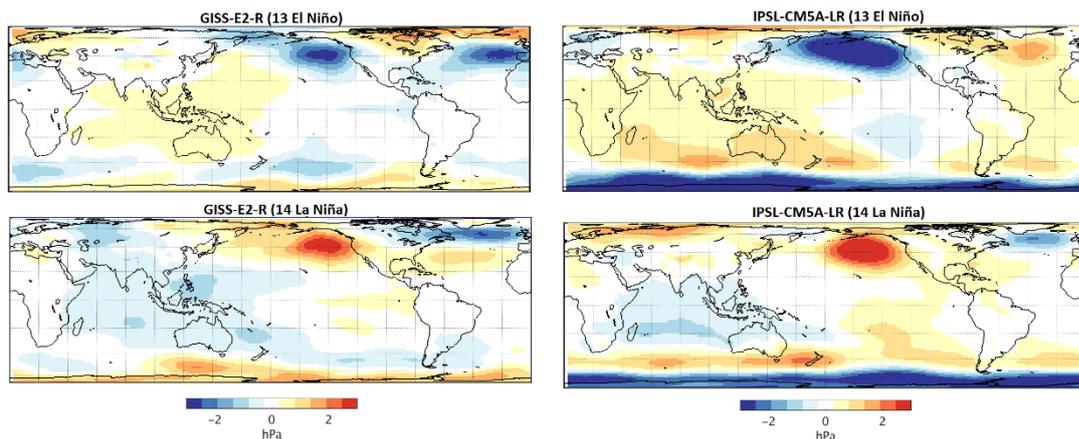


Fig. 8.8 El Niño and La Niña MSLPA (hPa) composites defined by one standard deviation of the DJF SOI from historical simulations, 1900-1999 from GISS-E2-R and IPSL-CM5A-LR. Anomalies based on 1900-1999 DJF mean. Numbers in brackets indicate number of events used to calculate each composite plot.

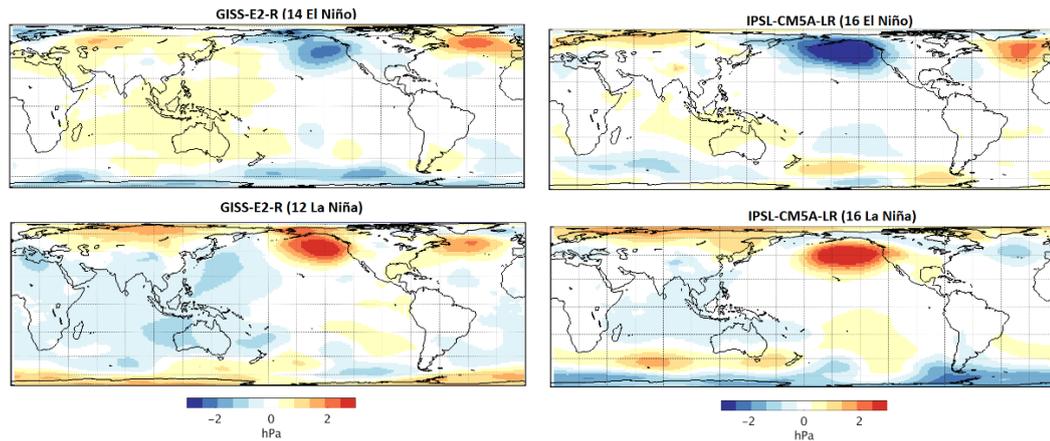


Fig. 8.9 El Niño and La Niña MSLPA (hPa) composites defined by one standard deviation of the DJF SOI from LM simulations, 1750-1848 from GISS-E2-R and IPSL-CM5A-LR. Numbers in brackets indicate number of events used to calculate each composite plot.

The MSLPA composites, Fig. 8.9, from the two models indicate positive MSLPA over Northern Australia and much of the Maritime Continent during El Niño events and negative MSLPA in this region during La Niña events, consistent with historical and reanalysis composites (Fig. 8.8; Fig. B2). Again, the strength of the anomalies are weaker in the LM simulation period in Fig. 8.9, compared to the historical period. The MWP MSLPA are generally stronger than those in Fig. 8.9 (1750-1848), Both of the models capture a large area of negative MSLP over the North East Pacific during El Niño events and positive MSLP during La Niña, associated with modulations in the Aleutian Low that are present in the historical simulations (Fig. 8.8; Wang et al. 2012). During El Niño events the Aleutian Low expands southwards and the high-pressure cell off California generally weakens, creating a large area of anomalous low pressure. In general, the MSLP spatial anomaly patterns associated with ENSO are similar, but weaker, within the periods assessed in the LM simulations compared to this historical period. This is in agreement with findings from Fig. 8.6 and Fig. 8.7 comparing surface temperature over these periods.

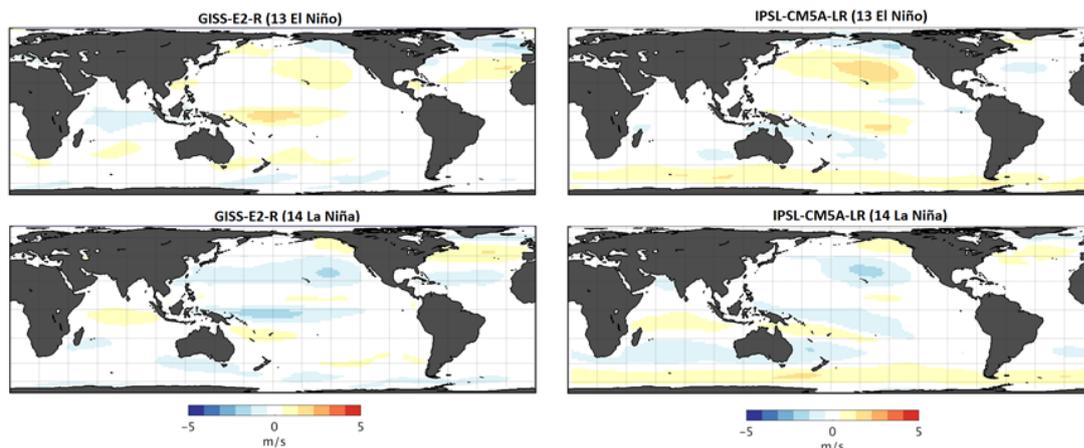


Fig. 8.10 El Niño and La Niña zonal wind (ms^{-1}) anomaly composites defined by one standard deviation of the DJF SOI from historical simulations, 1900-1999 from GISS-E2-R and IPSL-CM5A-LR. Anomalies based on 1900-1999 DJF mean. Numbers in brackets indicate number of events used to calculate each composite plot.

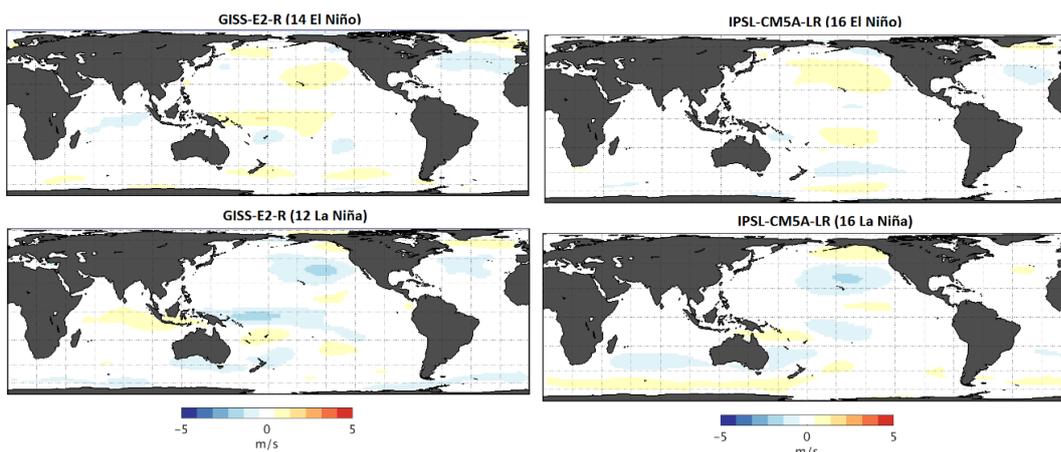


Fig. 8.11 El Niño and La Niña zonal wind (ms^{-1}) anomaly composites defined by one standard deviation of the DJF SOI from LM simulations, 1750-1848 from GISS-E2-R and IPSL-CM5A-LR. Numbers in brackets indicate number of events used to calculate each composite plot.

The spatial zonal wind patterns associated with ENSO events, 1750-1848, from the two LM model simulations are shown in Fig. 8.11. The anomaly patterns show some variation between the two models. Similar patterns are seen in the North East Pacific, with positive zonal wind anomalies during El Niño and negative zonal wind anomalies during La Niña. This is linked to the MSLPA in this region seen in Fig 8.9, with low pressure during El Niño events due to the deepening of the Aleutian Low, resulting in

more westerly zonal winds. The zonal wind pattern in the western Pacific is more prominent in GISS-E2-R compared to IPSL-CM5A-LR and is more closely comparable to the spatial anomalies found in the reanalysis over the twentieth century, however much weaker in strength compared to ERA-20C (Fig. B6).

The SOI defined events have generally similar spatial composite anomaly patterns during the LM as those in the historical and reanalysis, suggesting that there is similarity in ENSO characteristics during the historical and 1750-1848 LM period. This is also found when looking at the MWP (Appendix Fig. E3-E5). The spatial pattern of winds and surface temperature compared to the historical reanalysis are closer to observed in GISS-E2-R, whereas the spatial pressure patterns are better represented in IPSL-CM5A-LR. However, in all cases the strength of the anomalies is weaker in the LM simulations compared to the historical simulations. This supports findings of previous studies of ENSO spatial composites which concluded that the strength of the anomalies are generally weaker over the Last Millennium (Lewis and LeGrande 2015).

8.3.2 The Stability of ENSO-wind teleconnections in the Last Millennium

The DJF SOI is used as the main indicator of ENSO for this analysis, as it was for the logbook-based reconstruction, because it is the atmospheric measure of ENSO and therefore is more closely linked to the winds than the Niño 3.4 index. However, as stated in Chapter 2, initial analysis carried out using both indices found good agreement between them and therefore the choice of focus on the SOI is not expected to influence the results unduly. The two LM periods, 1050-1149 and 1750-1848 are compared with results from the historical period, 1900-1999 presented in Chapter 7. In addition, the strength of the running correlation over the entire LM is assessed in both models, focusing on the stability in the correlation within the predictor grid boxes.

The Pearson's correlation coefficient between the DJF SOI and the DJF zonal wind of GISS-E2-R and ISPL-CM5A-LR, 1750-1848 and 1050-1149 are shown in Fig. 8.12 and Fig. 8.13 respectively. Higher spatial resolution analysis and plots showing all correlations regardless of significance value can be found in Appendix E, Fig. E6-E9 for these two LM periods. The global ENSO-wind teleconnections are shown, with the nine predictor grid boxes highlighted. The predictor grid boxes for the logbook-

based reconstruction were originally chosen based on significant correlations between ERA-Interim DJF zonal wind and the instrumental DJF SOI over the period 1979-2013 (Fig. 2.5), all of which were significant at the 90% confidence level. The correlations shown in Fig. 8.12 and Fig. 8.13 are those from the LM simulation data re-gridded to the logbook resolution ($7.5^\circ \times 8^\circ$ latitude-longitude), with only those significant at the 90% confidence level shown. The correlation coefficients from the predictor grid boxes over this period are shown in Table 8.1. Similar maps for the two LM periods, at finer resolution and showing all correlation regardless of significance value can be seen in Appendix Fig. E6; Fig. E7. For reference, the correlations found in Chapter 7 (Fig. 7.14) for these two models over the historical period, 1900-1999, are shown in Fig 8.14 at the coarser $7.5^\circ \times 8^\circ$ resolution.

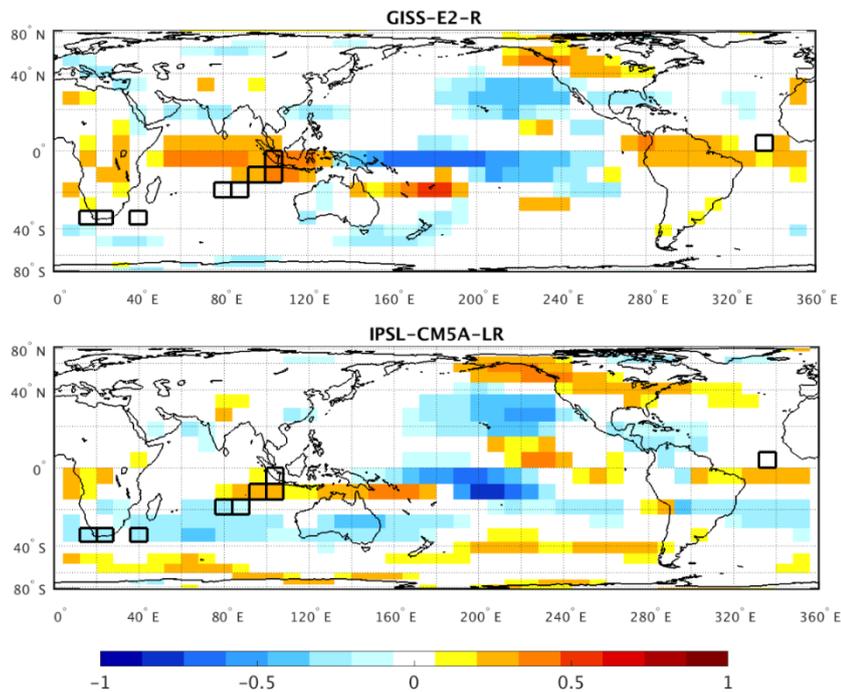


Fig. 8.12 Pearson's correlation coefficient between the mean DJF zonal wind and the DJF SOI from GISS-E2-R and IPSL-CM5A-LR for the years 1750-1848 of the LM simulations. Data are detrended prior to calculation of correlations and only those correlations significant at 90% confidence level are shown.

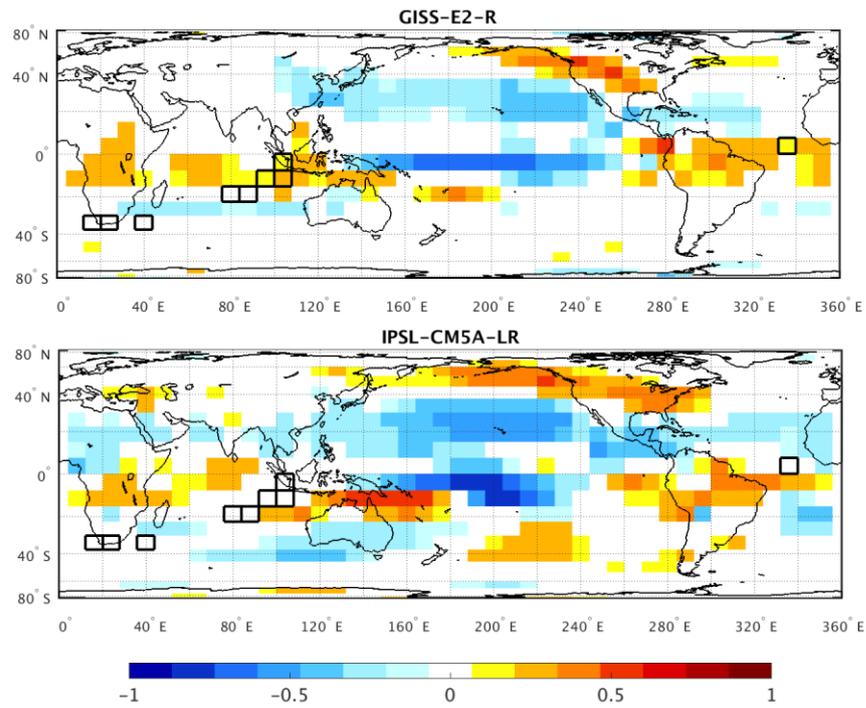


Fig. 8.13 As in Fig. 8.12 but for the years 1050-1149 from the two LM simulations.

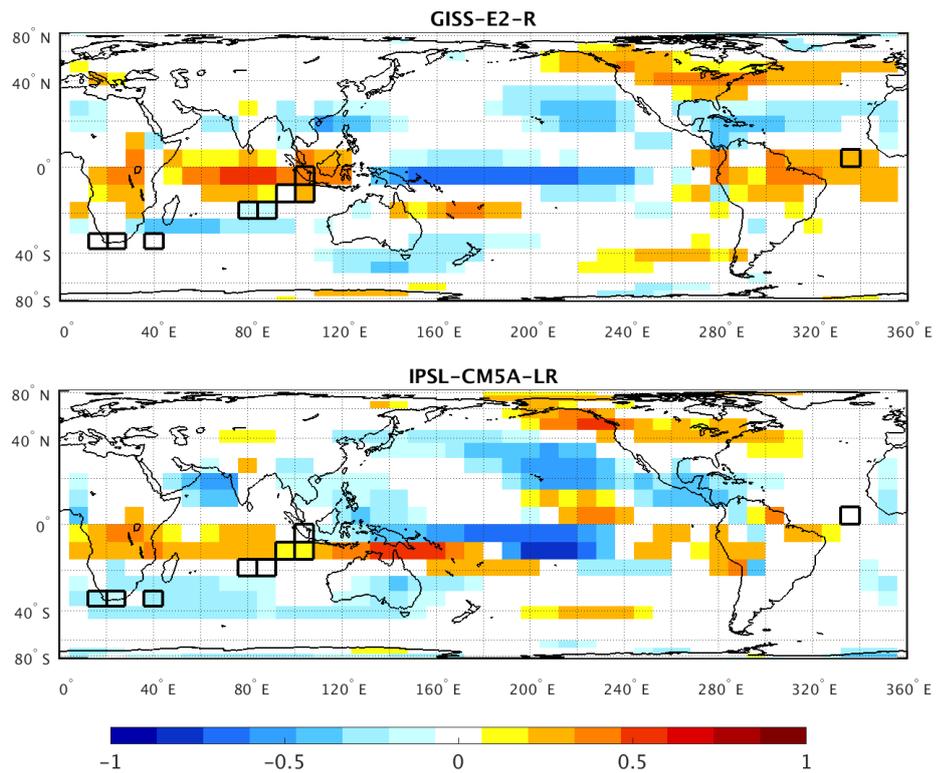


Fig 8.14 As in Fig. 8.12 but for the years 1900-1999 from the two historical simulations.

Table 8.1 Pearson’s correlation coefficient between the detrended DJF SOI and the detrended DJF zonal wind in the predictor grid boxes (1-9; 7.5° x 8° resolution), from LM simulations, 1750-1848 and 1050-1149, and the historical simulations, 1900-1999. Bold indicates significance at the 90% confidence level.

	1750-1848		1050-1149		1900-1999	
	GISS-E2-R	IPSL- CM5A-LR	GISS-E2-R	IPSL- CM5A-LR	GISS-E2-R	IPSL- CM5A-LR
1	0.13	0.04	0.20	-0.14	0.32	-0.10
2	-0.06	-0.33	-0.07	0.06	-0.06	-0.28
3	-0.08	-0.25	-0.09	0.01	-0.02	-0.18
4	-0.12	-0.21	-0.13	-0.07	-0.06	-0.20
5	-0.04	-0.29	-0.13	0.03	-0.18	-0.15
6	-0.05	-0.21	0.00	0.13	-0.20	-0.13
7	0.34	0.31	0.17	0.02	0.14	0.17
8	0.41	0.27	0.20	0.05	0.24	0.19
9	0.36	0.13	0.19	0.02	0.40	-0.04

Chapter 7 showed that within the historical climate model simulations the strength of the correlation coefficient between the SOI and zonal wind was weaker than, and in some cases opposite sign to, those found using ERA-20C and the instrumental record. GISS-E2-R was found to capture the sign of the correlation coefficient in all nine of the predictor grid boxes compared to ERA-20C, 1900-1999, while both GISS-E2-R and IPSL-CM5A-LR found significant correlations (90% confidence level) in five of the nine predictor grid boxes over this historical period. However, as seen in Fig 8.14, only in grid box 8 do both these models agree on a correlation above this significance threshold, suggesting that the two models capture the ENSO-wind teleconnection with spatial differences.

Within the 1750-1848 period of the LM simulations of GISS-E2-R and IPSL-CM5A-LR, the correlations of the predictor grid boxes are of the same sign as those found in ERA-20C, 1900-1999, in all of the nine predictor grid boxes. The correlation between the DJF SOI and the DJF zonal wind are significant at the 90% confidence level in three of the predictor grid boxes in GISS-E2-R (grid boxes 7, 8 and 9), and in seven

of those in IPSL-CM5A-LR (2-8), as seen in Table 8.1. Over the earlier 100-year period assessed, 1050-1149, the number of predictor grid boxes with significant correlations (90% confidence level) is reduced to zero in IPSL-CM5A-LR, but retained, with weaker strength than those in 1750-1848, in the three GISS-E2-R grid boxes (7-9), and in addition are significant in grid box 1.

Analysis of these three periods from within the two climate models suggests that the correlation between the DJF SOI and the DJF zonal wind is present over these time periods but the strength and locations of this teleconnection varies both within the individual models over time and when comparing the two models over the same time period. The strength of the correlations is expected to be weaker within the two climate models compared to the modern period reanalysis and instrumental record, but overall the general pattern of SOI-wind teleconnection captured by this model during this LM period (1750-1848) is reasonably similar to historical patterns. The correlations are generally weaker in the earlier period, 1050-1149, this is especially the case within the IPSL-CM5A-LR simulation. The grid boxes in the eastern Indian Ocean (7-9) show the strongest correlations over the various time periods analysed, a finding which is more obvious in the GISS-E2-R simulation. These were the grid boxes with strongest correlations in the modern period when ERA-Interim was assessed to inform the logbook-based reconstructions (Fig. 2.5).

To explore the correlation between the SOI and zonal wind further within the Last Millennium simulations, 30-year running correlations were calculated for the entire LM for each of the predictor grid boxes (following the same methodology as Chapter 7). Fig. 8.15 and Fig. 8.16 show the variation in strength of the running correlation between zonal wind and the SOI in the predictor grid boxes over the LM in GISS-E2-R and IPSL-CM5A-LR, respectively. The correlations vary within the grid boxes over the LM, they are only above the 90% confidence level during certain periods, and in some cases the sign of the correlation coefficient changes. This is similar to the changes in strength and sign of correlations between ENSO and rainfall over the LM found by Brown et al. (2016). The variability in the SOI-zonal wind correlation is found in both of the models, suggesting non-stationarity in the SOI-zonal wind teleconnection at this coarse resolution which was explored in Fig. 2.5 when carrying out the logbook-based SOI reconstruction.

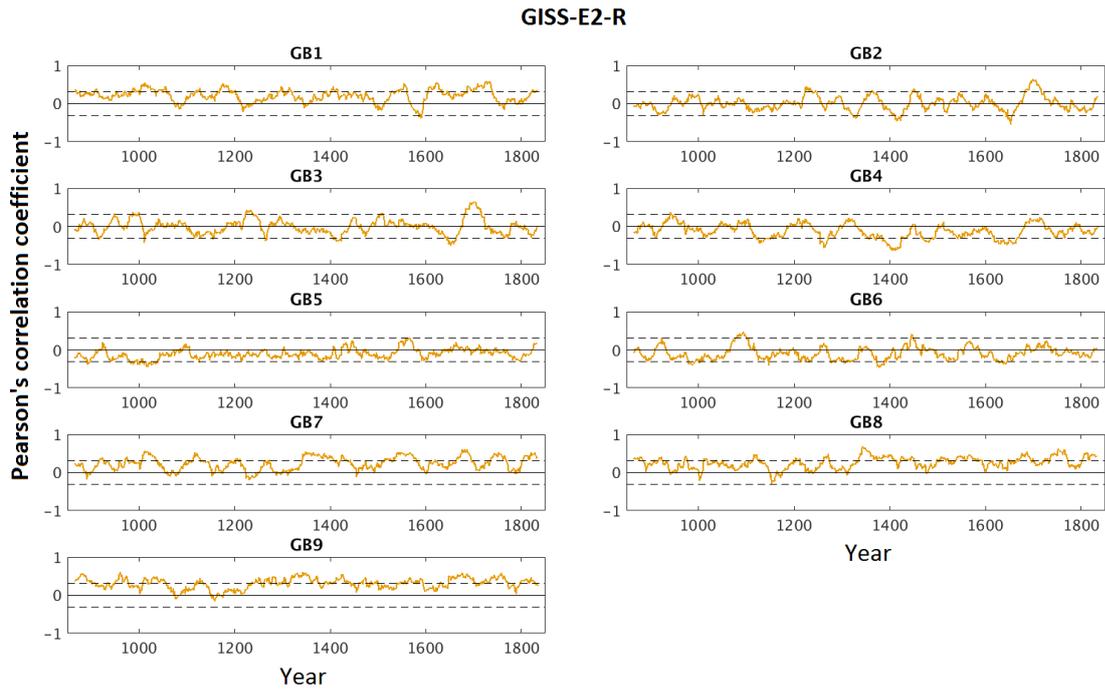


Fig. 8.15 30-year running Pearson's correlation coefficients between mean DJF zonal wind and DJF SOI from GISS-E2-R LM climate model simulations 1750-1848. Data detrended prior to calculations of correlations. The year shown is the 15th year of the window. Dashed lines indicate significance at the 90% confidence level.

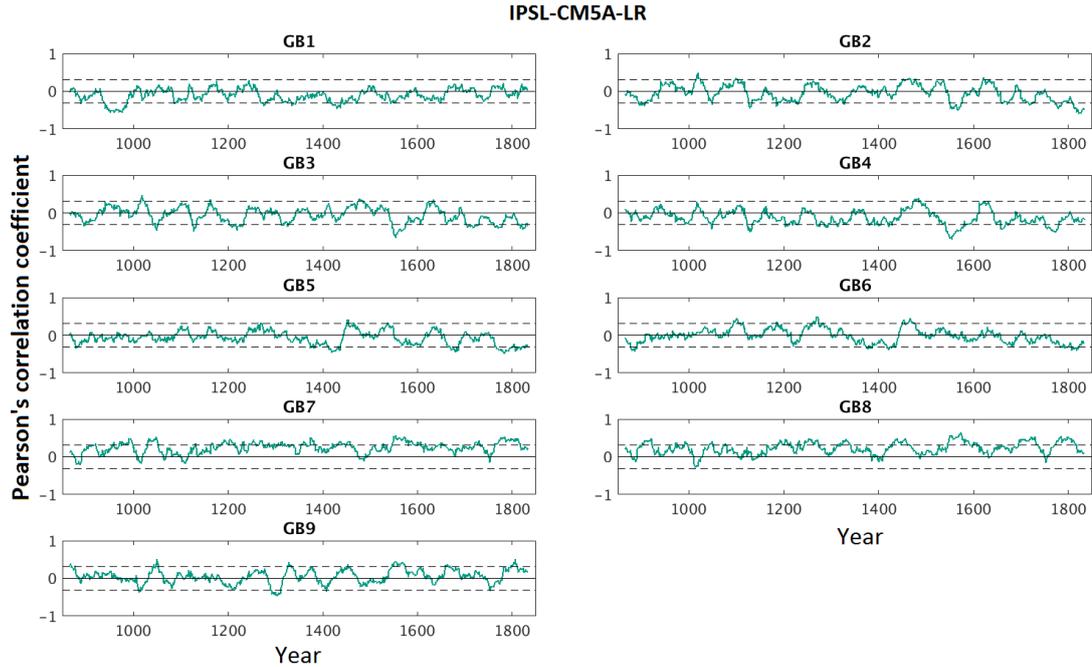


Fig. 8.16 30-year running Pearson's correlation coefficients between mean DJF zonal wind and DJF SOI IPSL-CM5A-LR climate model simulations 1750-1848. Data detrended prior to calculations of correlations. The year shown is the 15th year of the window, dashed lines indicate significance at the 90% confidence level.

The running correlations provide a similar indication of the strength of the teleconnections as provided above (Fig. 8.12 and 8.13, Table 8.1), but using a sliding window they show how the strength of the correlations vary over time. The strength of the correlation over 1750-1848 cannot be directly compared due to the running correlations over a 30-year period, therefore the critical value for the 90% confidence level is different in Fig. 8.12 and Fig. 8.13 compared to Fig. 8.15 and Fig. 8.16, and the strength of the correlation expected to be higher over a shorter period.

The extent of non-stationarity in the teleconnection can be measured from the standard deviation of the running correlation between the two variables (Brown et al. 2016). Coats et al. (2013) found that for models with a relatively larger ENSO amplitude, the non-stationary of the ENSO-teleconnection was smaller. This was also found by Brown et al. (2016), who found the ENSO amplitude of IPSL-CM5A-LR to be relatively small compared to the models they analysed and the teleconnection non-stationarity was larger, when looking at eastern Australian rainfall. Both of these studies calculated correlation over large regions such as Eastern Australia and North

America. Here, similar analysis is carried out on the scale on which the logbook-based reconstruction was derived, which covers a number of smaller spatial areas. The grid box scale within this study compared to the larger spatial scale of previous studies could be a limitation of the current results. Table 8.2 shows the standard deviation of the running correlations over the LM from the two model simulations for each predictor grid box.

Table 8.2 Standard deviation of 30-year running correlation over the LM period in each predictor grid box.

	Predictor grid box number								
	1	2	3	4	5	6	7	8	9
GISS-E2-R	0.17	0.19	0.20	0.19	0.13	0.17	0.17	0.16	0.14
IPSL-	0.17	0.21	0.20	0.19	0.17	0.19	0.15	0.16	0.18
CM5A-LR									

The standard deviation of the DJF SOI over the entire LM from GISS-E2-R and IPSL-CM5A-LR is 7.31 and 7.67, thus GISS-E2-R exhibits a slightly larger amplitude. The standard deviation of the correlation coefficients, and thus the non-stationarity of the teleconnection, are found to be fairly similar over the Last Millennium in the two model simulations. As the ENSO amplitude is relatively similar in the two models and the grid box regions are relatively small compared to the region analysed by Brown et al. (2016), the pattern they discussed is not present in the logbook regions in these two models.

Overall, GISS-E2-R and IPSL-CM5A-LR have been investigated to see if ENSO characteristics in the LM simulations are similar to those found in the historical period, 1900-1999, assessed in Chapter 7. A similar frequency of events is found, and global spatial patterns are generally weaker but overall relatively similar. Non-stationarity in the SOI-zonal wind teleconnection is found in the two simulations over the LM, with grid boxes closer to the ENSO centre of action more stable than more remote ones. Correlations were found to be similar, but weaker, in GISS-E2-R in the two LM periods compared to the historical period. There was a clear decline in the strength of the correlations in the grid boxes from 1750-1848 period compared to the 1050-1149

MWP period in IPSL-CM5A-LR. Stationarity is a key assumption of paleo-reconstructions, therefore using the results of non-stationarity here, additional caution is needed when using SOI-zonal wind teleconnection to carry out reconstructions. It is suggested that grid boxes which exhibit a lower degree of non-stationarity should be used preferentially in future reconstructions. This is discussed further in Chapter 9.

8.3.3 El Niño diversity over the Last Millennium

ENSO diversity was another key aspect of investigation within Chapter 6, assessing Eastern Pacific and Central Pacific El Niño events during the twentieth century. GISS-E2-R and IPSL-CM5A-LR are now used to assess ENSO diversity over the LM using similar methods as those used in Chapter 6. Fig. 8.17 shows the Niño 3 and Niño 4 surface temperature from the two LM model simulations.

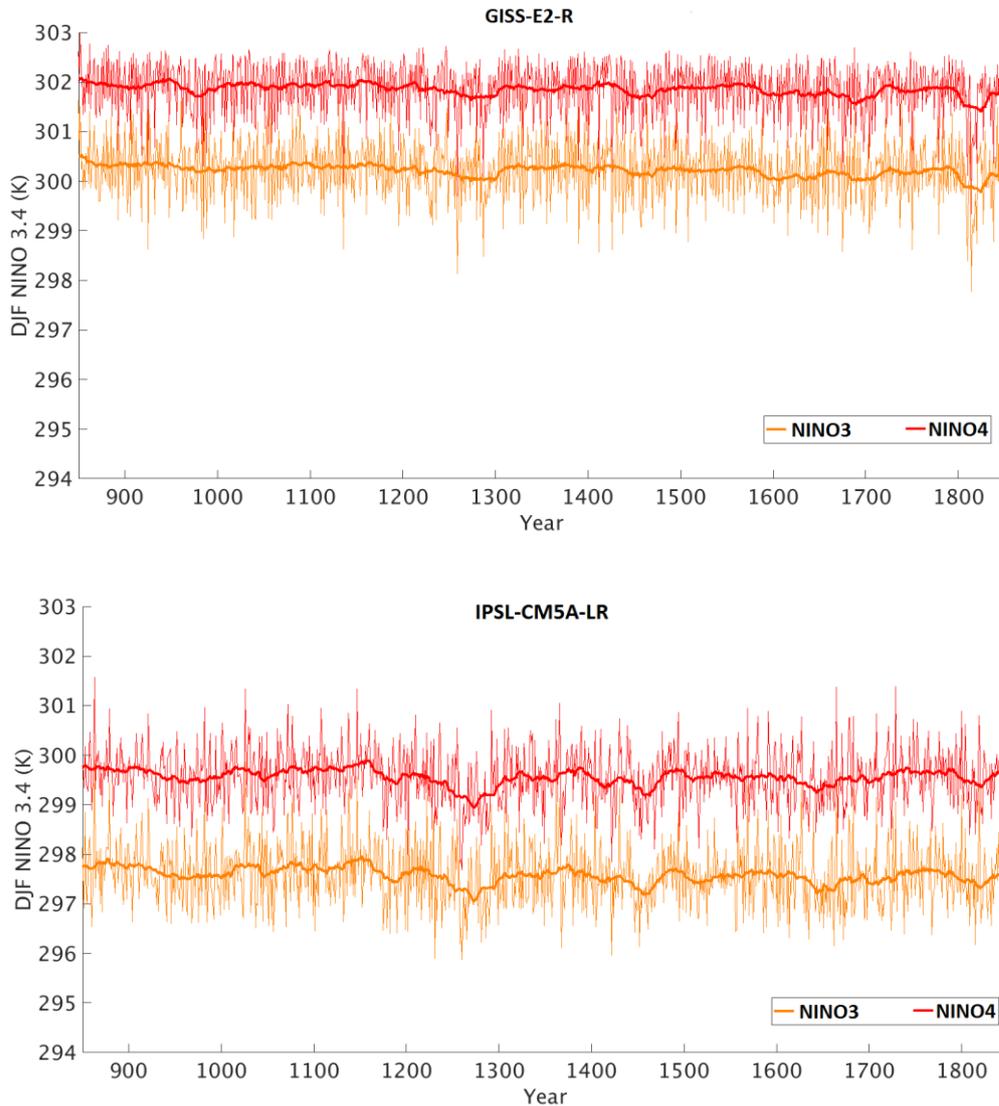


Fig. 8.17 The DJF Niño 3 and DJF Niño 4 surface temperature from GISS-E2-R and IPSL-CM5A-LR LM simulation, 850-1848. Bold lines indicate 30-year running mean. The standard deviation in Niño 3 (Niño 4) over the entire LM is 0.60 K (0.54 K) in GISS-E2-R and 0.64 K (0.57 K) in IPSL-CM5A-LR. The warm bias found in Chapter 6 in the tropical Pacific in GISS-E2-R is again found in LM simulation of this model, with the mean Niño 3 (Niño 4) 2.64 K (2.29 K) higher than that in IPSL-CM5A-LR and 1.42 K (0.55 K) higher compared to the HadISST in the twentieth century. The number of EP and CP events in each 100-year period of the LM from the two model simulations is shown in Fig. 8.18 along with the CP-EP ratio shown in Fig. 8.19, the values for which can be found in Appendix Table E2.

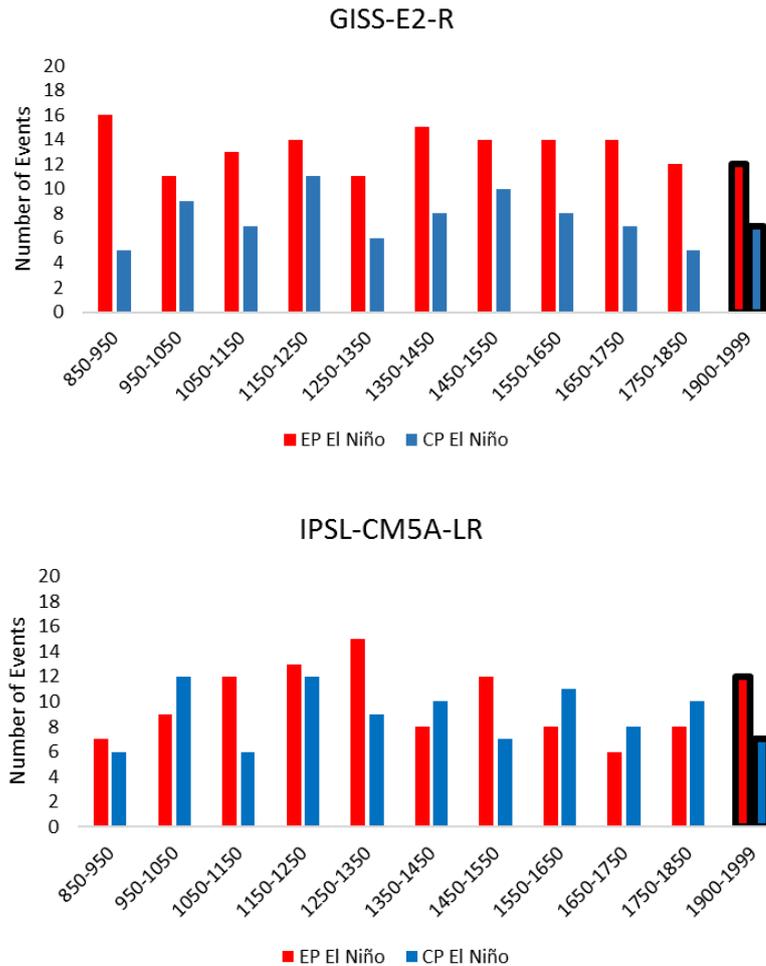


Fig. 8.18 Number of EP and CP El Niño events per 100-year period of LM simulation of GISS-E2-R (top) and IPSL-CM5A-LR (bottom), and the number of events from ERA-20C in the twentieth century, as outlined in black

The number of EP El Niño events is consistently higher than the number of CP events for the ten 100-year periods of the LM in GISS-E2-R, with the CP-EP ratio remaining lower than one. Over the twentieth century, the CP-EP ratio for GISS-E2-R was 0.64 (Chapter 6), therefore there are similar characteristics in the historical and LM simulation of this model. IPSL-CM5A-LR had one of the highest CP-EP ratios of the historical model simulations, higher than the observed ratio over the twentieth century. IPSL-CM5A-LR has a bias towards CP events due to the ENSO surface temperature anomalies being further west than observed. The CP-EP ratio over the Last Millennium in GISS-E2-R is generally lower than that found by ERA-20C in the twentieth century, 1900-1999, while it is generally higher in IPSL-CM5A-LR. The

CP-EP ratio varies across the 100-year periods of the LM. The two models show different behaviour across the LM, with the CP-EP ratio in IPSL-CM5A-LR peaking in the last three centuries of the simulation, while GISS-E2-R peaks in the second and fourth 100-year periods.

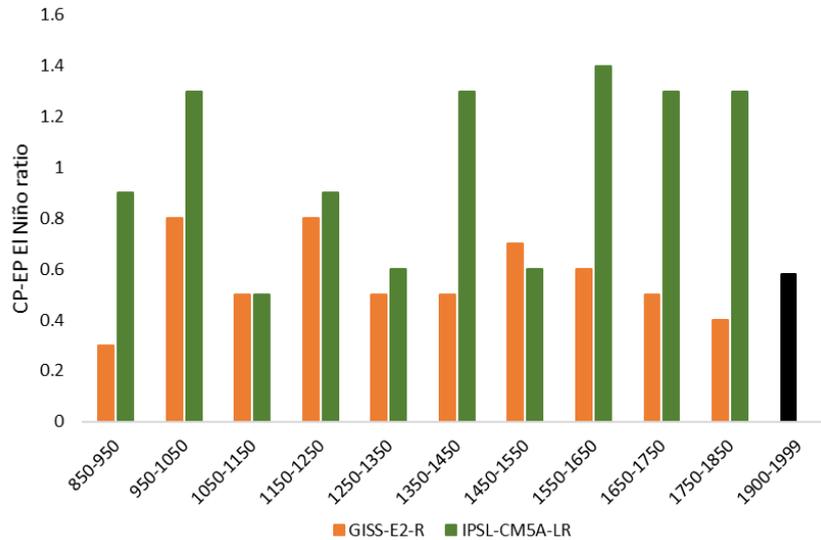


Fig. 8.19 The CP-EP Ratio for the ten 100-year periods of the LM simulations of GISS-E2-R and IPSL-CM5A-LR, and that from ERA-20C reanalysis over the last century shown in black.

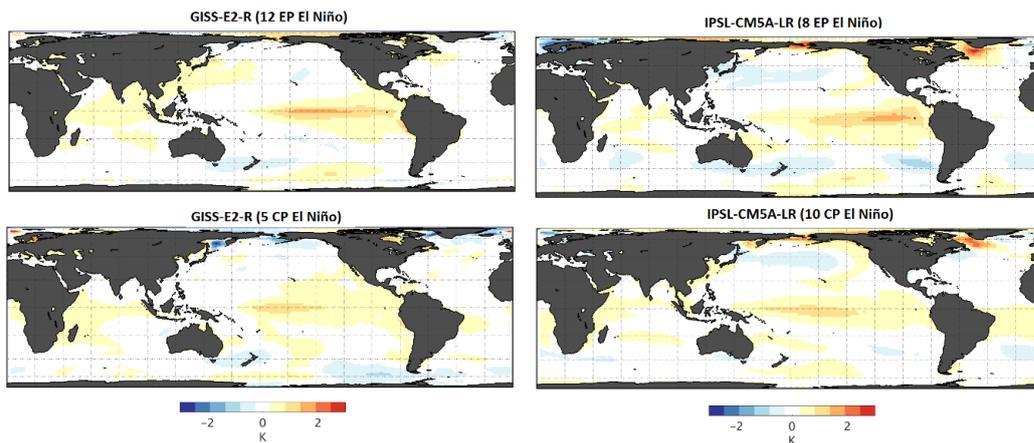


Fig. 8.20 EP and CP El Niño surface temperature (K) anomaly composites defined by the NINO method from LM simulations, 1750-1848 from GISS-E2-R and IPSL-CM5A-LR. Numbers in brackets indicate number of events used to calculate each composite plot.

In order to explore ENSO diversity further, spatial composite plots for EP and CP El Niño events during periods of the LM are explored. The surface temperature anomaly plots for EP and CP El Niño events during the last 100-year period of the LM simulations (Fig. 8.20) and the MWP (Appendix Fig. E10) show a similar spatial pattern to that seen in the twentieth century (Chapter 6), with stronger positive anomalies in EP events than CP events and with the region of surface temperature anomalies further east in EP events. The spatial patterns across the equatorial Pacific are similar in both GISS-E2-R and IPSL-CM5A-LR, suggesting similarities in the spatial temperature patterns during the different types of El Niño events in the last 100-years of the LM simulations. However, during EP events in GISS-E2-R, the surface temperature anomalies are higher further west than in IPSL-CM5A-LR, and in CP El Niño events the positive temperature anomalies expand further west in IPSL-CM5A-LR than in GISS-E2-R. The differences in the exact location of the highest anomalies during the two types of events between the two models suggest that the EP-CP ratio is partly a result of the definition of the events which selects the years with the highest temperature anomalies in these two previously defined regions. Using these set geographical regions to define EP and CP events within models with differences in climatology and in spatial composite patterns brings some bias towards the results. Future work should strive to solidify stronger methods of identification of EP and CP events from numerous climate model simulations in order to be able to assess the teleconnections associated with the types of El Niño events with reduced bias. The issues and the implications of the current lack of consensus on defining EP and CP El Niño events is discussed further in Chapter 9.

8.4 Summary

This chapter has explored ENSO over the Last Millennium using the two best performing models over the historical period assessed in Chapters 6 and 7. A ranking method was used in order to select these methods, rather than using a multi-model mean as it is thought that those which capture ENSO best over the historical period are more likely to provide a better representation of ENSO over the Last Millennium. Lewis and LeGrande (2015) concluded their study of ENSO-teleconnections in LM simulations by stating that due to issues with certain models, future work should, rather than using a range of models, aim to select one model to investigate ENSO over the

entire LM. Klein et al. (2016) suggested that the use of a multi-model mean to study climatic variation of the LM should not be advised, as averaging would lead to already weak responses becoming even weaker due to the different internal variability of the model. However, a key limitation of this approach is that the results from this chapter are therefore largely model-dependent. The two models found agreement in the frequency and spatial composites for ENSO events, but were not always consistent with each other over the LM in terms of the strength of the SOI-teleconnection and ENSO diversity. Therefore, ENSO Millennial scale variation is model-dependent, in part due to the biases within the individual models as outlined in Chapter 6.

The frequency of El Niño and La Niña events in the instrumental records over the twentieth century was found to be in the mid-range of that from the two models over the Last Millennium. This suggests that, in terms of the frequency of ENSO, the twentieth century was within the range of natural variability. Spatial composite temperature anomaly plots for El Niño and La Niña events during the two 100-year periods of the LM showed similar patterns of anomalies, but of weaker strength, to those found in the historical period. This is similar to findings from Lewis and LeGrande (2015) who found that the models showed broadly similar global impacts associated with the Niño 3.4 temperature anomalies, in the LM and historical periods, although with generally weaker anomalies in the LM simulations.

Previous studies which used temperature-based indices to analyse ENSO over the LM highlighted a potential link between volcanic eruptions and ENSO events, with a dip in temperature-based indices following the largest of the volcanic eruptions of the LM (Hope et al. 2016; Le 2017). However, using the DJF SOI this association was not found in GISS-E2-R or IPSL-CM5A-LR for the three biggest eruptions looking over the 1000-year period (Fig. 8.4), or within the years immediately surrounding the three events (Fig. E1 and Fig. E2). This could partly be model-dependent results, but also a result of using the SOI instead of the SSTA indices which would be more directly impacted by changes in global temperature associated with large volcanic events.

Tejavath et al. (2017) also noted the influence of changes on global temperature on the number of warm (El Niño) and cold (La Niña) ENSO events during certain periods of the LM, with more warm events during the MWP and more cold events during the

LIA. This was not found to be a conclusion which could be formed from analysis of the frequency of events during these periods in the two LM model simulations used in this chapter (Fig. 8.5). This could also be partly due to the use of temperature-based ENSO indices compared to the use of the SOI here.

The stability of ENSO-wind teleconnections and ENSO diversity over the Last Millennium are discussed further in Chapter 9 (Section 9.3.3), which links together findings from throughout this thesis on these two key themes. This chapter has found that the stationarity of SOI-wind teleconnections varies over the course of the LM within the predictor grid boxes, resulting in implications for the interpretation of reconstructions which are based on the assumption of stationarity. Over the two 100-year periods of the LM climate model simulations investigated in detail it was found that both models capture ENSO characteristics in a similar way to the historical simulations, but with generally weaker correlations between the SOI and zonal wind and weaker anomalies in the spatial composites of El Niño and La Niña events. GISS-E2-R found that significant correlations were found in both 1750-1848 and 1050-1149 periods (Table 8.1) in the eastern Indian Ocean region, the predictor grid boxes closest to the ENSO centre of action. This suggests that they would be the best predictor grid boxes to use for logbook-based reconstruction. However initial work in the logbook-based reconstruction found a decrease in reconstruction skill when using only grid boxes 7-9 compared to all those that were used in Chapter 2. GISS-E2-R also found a weakening of the correlation in the MWP compared to 1750-1848, and IPSL-CM5A-LR found a clear drop with no correlations meeting the 90% confidence level 1050-1149, compared to in seven of the nine in 1750-1848.

The frequency of EP and CP events during the historical period was found to be within the range of that found in the 100-year periods of the LM (Fig 8.18). The CP-EP ratio in the two models was different to each other in the LM, with IPSL-CM5A-LR having a higher CP-EP ratio than GISS-E2-R (Fig. 8.19). It is important to understand the frequency of the different types of El Niño events in the past due to increased attention and the increased frequency of CP events in recent decades reported in the literature, and the projected future changes in frequency due to anthropogenic climate change (Yeh et al. 2009; Lee and McPhaden 2010). A key issue surrounding analysis of this is the current lack of definition on which method of classifying EP and CP El Niño

events should be adopted. This currently limits the studies and conclusion which can be drawn around EP and CP events (Pascolini-Campbell et al. 2015). There are a large range of methods available for defining ENSO diversity, with new methods still being developed (Yu and Kim 2013; Capotondi et al. 2015). It is suggested that future work on drawing a consensus on this definition would help to bring together work in this area (Pascolini-Campbell et al. 2015). The results from this chapter will now be discussed within Chapter 9 which summarises the findings from the climate model analysis and discusses them alongside the results of the historical reconstruction and proxy data analysis carried out within Chapter 2-4.

9. Discussion and Conclusion

9.1 Discussion

Throughout this thesis, while addressing the main aims outlined in Chapter 1, a number of key themes have emerged. This chapter provides a discussion of these themes including the main findings relating to them and how they have built on previous work, the limitations of research within these themes and suggestions for future work to expand on these further. These themes include reconstruction methods, the use and challenges of multi-data source comparisons, the ENSO-wind teleconnection and ENSO diversity.

9.2 Reconstruction methods and data comparisons

Firstly, the differences between reconstruction methodologies and the data networks used in exploring ENSO of the past have become apparent as an important theme. This includes differences in logbook-based reconstruction methods, the impact of limited data availability on these methods, the classification of ENSO events from multiple reconstructions and the limitations of the various reconstruction data types.

9.2.1 Differences in logbook-based reconstruction methodologies

The logbook-based ENSO reconstruction, presented in Chapter 2, employed two commonly used statistical methods which have widespread application within paleoclimatology (Wilson et al. 2010). Reconstructions were produced using Principal Component Regression (PCR) and Composite-Plus-Scale (CPS). PCR has been used previously within logbook-based reconstructions (Jones and Salmon 2005; Hannaford et al. 2015), however in Chapter 2 CPS was applied to logbook observations for the first time. Previous multi-proxy reconstructions have found these two methods to be largely complementary (Mann et al. 2008; Wilson et al. 2010; Emile-Geay et al. 2013a). This was also found here, when the two methods were applied to logbook-based SOI reconstructions, with both the reconstructions having a strong agreement with the instrumental SOI in the fitting period ($r=0.89$ for PCR and $r=0.87$ for CPS) and with each other in the historical reconstruction period (PCR A and CPS A; $r=0.81$; PCR B and CPS B, $r=0.75$). Therefore, it is suggested that CPS should be considered

as an additional method which can be applied to future climate reconstructions which use the observations from ship logbooks. These two methods have led to the production of a new reconstruction of the DJF SOI from 1815-1854 which is valuable to future ENSO comparison studies and to increasing understanding of past ENSO variability.

9.2.2 Threshold test methodology

As well as the new historical logbook-based reconstructions produced by these two methods, another key output from Chapter 2, was an advance in the methodology for assessing the impact of a changing number of observations per season on the two reconstruction methods. This built upon the work of Küttel et al. (2010), providing both an initial assessment of the reconstruction skill using an idealised uniform data availability across the predictor grid boxes (e.g. 10 observations per grid box per season), and additionally a quantification of the reconstruction skill for each individual year of the historical reconstruction, 1815-1854. Threshold tests were carried out to assess the reconstruction skill of each year of the historical reconstruction and found that both PCR and CPS methods performed relatively well with a limited amount of data in the predictor grid boxes per season (Fig. 2.4). It was also found that the CPS method had better reconstruction skill compared to PCR in seasons with the most limited number of observations (Table 2.1). A key result from this analysis was that even with as few as three observations per season in each predictor grid box, using the CPS method, an RE skill score above zero can be obtained. This indicates a degree of skill within the reconstruction methods, a skill higher than if using the assumption of climatology (Table 2.1). In reality, the number of observations across the grid boxes is not normally this low in any of the reconstructed years, and thus has better skill overall. It is suggested that future logbook-based reconstructions should assess their reconstruction methods by replicating the availability of logbook data for each individual year in order to assess the impact of the changing number of observations input into their reconstruction over their temporal coverage, and for the geographical region(s) from which their predictors are taken, as results of this analysis are likely to be regionally specific. The CPS method performs better in the extreme cases of very limited data availability, whereas PCR performs better when using full ERA-Interim data availability (Table 2.1). This provides additional support to the suggestion that

both methods should be used in the future to provide the best possible reconstructions when using the observations from ship logbooks.

9.2.3 Differences in wind terms

A final important discovery from the logbook-based reconstructions was the statistically significant difference (99% confidence level) found in the Beaufort-Force equivalent terms used by the EEIC and the Dutch ships in the predictor grid box regions during this reconstruction period, 1815-1833. The adjustment in the ENSO reconstruction due to this is a new method applied in Chapter 2, as this difference had not been noted or therefore adjusted for, in previous studies that use logbook data. A number of reasons for the difference in wind force terms used were suggested in Chapter 2 such as the Dutch focus on the sails to estimate wind force (as opposed to the state of the sea more commonly used by the British), the conversion of terms in the Dutch Extract logs in the late nineteenth century and the possible differences due to conversion of terms during the digitisation process. This is highlighted as a difference that should be tested for in future logbook-based reconstructions in any region which uses observations from more than one country. Cross-checks carried out as part of the digitization process, where wind observations from ships in convoy were compared, showed good agreement between observations (Wheeler 2005). However, convoys were often made up of ships from the same country and therefore using the same type of terminology. Therefore, it would be interesting to explore this further in the future and to assess the reasons why the EEIC observations rarely use Beaufort Force 3 terminology, which, in contrast, is commonly found within the Dutch records (Koek and Können 2005). Once more observations are digitised, studies could assess differences further by carrying out climate reconstructions within the same temporal and spatial coverage using observations from individual countries and comparing the outputs of the different countries.

9.2.4 Comparisons of multiple data sources

The ENSO reconstructions from ships logbooks were compared to a range of other documentary and proxy data sources providing their own records and reconstructions of ENSO in the historical period. A strength of the logbook-based reconstruction derived in Chapter 2 is its similarity to the Jakarta rain-day and pressure-based SOI

(Können et al. 1998). Both provide an indication of historical ENSO derived from direct and routine observations of the weather. Direct observations of weather events can also be found within other historical documentary records such as those which inform ENSO chronologies produced by Quinn et al. (1987), Ortlieb (2000) and Garcia-Herrera et al. (2008). However, these were not often routine observations and therefore only provide a sporadic, discontinuous, record of events. There is often a bias toward extreme or high impact events due to the human nature of noticing and reporting on extremes (Bradley and Jones 1992). The routine logbook and rain-day observations also have the benefit of being a record which is written down daily, on the day in which the weather phenomenon being observed occurred. In contrast, often with the extreme-event documentary records the information was written down some time after the event, often from second hand accounts, bringing a degree of ambiguity to the record (Garcia-Herrera et al. 2008). Garcia-Herrera et al. (2008) stated that the use of secondary sources by Ortlieb (2000), compared to the use of primary witness accounts of the weather phenomenon used in their study, was a possible source for the difference in ENSO chronologies. Overall, there is great value in direct and routine observations from primary sources such as the logbooks due to the consistency in these records compared to the events-based documentary records.

In terms of proxies, unlike the documentary records, they are indirect records of past climate and are therefore subject to a range of limitations and uncertainties, as discussed in Chapter 3. Proxies are sensitive to a number of climatic and environmental factors making their climatic signal a result of multiple factors (Gergis et al. 2006). Most of the multi-proxy reconstructions analysed in this thesis used annually resolved proxies (e.g. Wilson et al. 2010; Li et al. 2013). There is increased uncertainty in the dating of proxies further back in time, and a reduced availability of annually resolved proxies. In addition, the proxy records are calibrated over different seasons and to different indices (Table 3.2); both these factors lead to increased disagreement between the records. Therefore, although proxies provide a useful alternative to documentary records, they have a range of uncertainties associated with them.

9.2.5 Defining ENSO events from instrumental indices

Further work is needed to provide a consensus on the methodology for defining discrete El Niño and La Niña events. There remains a limited consensus as to which index and method best defines ENSO years and the strength of events (Gergis et al. 2006). Throughout this thesis one method has been used to create ENSO chronologies from time-series indices to provide consistency throughout and to identify clearly distinguishable and strong ENSO events. The identification of events depends on the index used, the method of identification employed and the period over which the mean of the series is relative to Braganza et al. (2009). The one standard deviation threshold is used here, as commonly used in multi-proxy reconstructions (Stahle et al. 1998; Mann 2000; Braganza et al. 2009). This is higher than some alternative limits used (e.g. 0.5 standard deviations McGregor et al. 2010), however it is thought to be the most suitable for use with reconstructions using logbook data as it is more likely to capture real events rather than incorrectly interpreting noise, which would be more of a problem if a lower threshold was used. The current threshold is therefore a good balance between removing weaker events but small enough to identify a number of events. More consistent classification methodologies would prove beneficial for future ENSO reconstruction comparisons.

One limitation of the one standard deviation methodology is the non-linearity of ENSO teleconnection, with El Niño teleconnections often stronger than La Niña teleconnections, and thus more likely to provide a stronger El Niño signal in the proxy records compared to the signal from a La Niña event (Hoerling et al. 1997; Batehup et al. 2015). This was discussed in Chapter 3, which found that the logbook and most of the multi-proxy reconstructions had a better El Niño skill compared to La Niña (Table 3.4; Table 3.6), which could have been a result of the use of the one standard deviation threshold. This is in agreement with previous studies which highlight a higher skill for El Niño events as a common feature of multi-proxy reconstructions (Braganza et al. 2009; McGregor et al. 2010). This should be noted as a limitation of these reconstructions and future work on addressing this issue would be useful for the multi-proxy reconstruction community.

9.2.6 Future expansion of ENSO reconstructions

Using a range of documentary and multi-proxy reconstructions provides the most comprehensive understanding of pre-instrumental climate, but as discussed it is also challenging due to the limitations of the various datasets and in the methods of comparison used. In addition, another challenge with working with such a range of datasets is that they are being continuously updated, with new datasets and new versions regularly released. This thesis has used datasets available at the time in which the research was carried out, however future research can exploit even more databases as they are released.

Firstly, the data used for the logbook-based reconstruction has a huge potential of being expanded. Spatial and temporal availability of digitised logbook observations was a clear limitation for the logbook-based ENSO reconstruction. Chapter 2 showed that digitised observations extend back to 1750, and a small amount even further back to 1662. However, only four years between 1750 and 1815 had enough data availability in the predictor grid boxes to carry out a reconstruction. Therefore, addressing these gaps in the data availability could expand this reconstruction further. There are continuous efforts to digitise historical documents such as the ship logbooks through projects including RECLAIM (Recovery of Logbooks and International Marine data) (Wilkinson et al. 2011), ACRE (The International Atmospheric Circulation Reconstructions over the Earth) (Allan et al. 2011), and the citizen science project ‘old Weather’ (<http://oldweather.org>). There is a lack of ships’ logbook observations found over the Pacific Ocean during the ENSO reconstruction period 1815-1854. Therefore, the potential for digitisation of alternative historical records from this region would be highly valuable for the reconstruction of historical ENSO events from direct observations.

There is also potential in extended the logbook-based reconstructions up to the modern day, using datasets which provide similar data to the logbook observations. For example, ICOADS release 3.0 is a released within the last year, is a gridded dataset of surface marine observations spanning from 1662 to 2014 which uses the increasingly digitised observations available to produce updates to their database (Freeman et al. 2017). Jones and Salmon (2005) showed how ICOADS data up to 2004, could be used to provide an extension of a logbook-based reconstruction, but were hindered by gaps

in data availability in earlier versions of this product. A similar approach using this dataset has been used to create long time series of the West African Monsoon (Gallego et al. 2015) and the onset of the Indian Monsoon (Ordóñez-Pérez et al. 2016). Thus, using this recently released dataset there is potential to expand the ENSO reconstructions forward in time. Therefore, opportunity exists for future research to extend the logbook-based reconstructions both forward and back in time, expanding on the 1815-1854 period used here. The methods presented here could also be applied to a range of other climatic phenomenon such as Indian Ocean Dipole, Asian Summer Monsoon building on previous work which largely uses logbooks from the North Atlantic (Gallego et al. 2005; Barriopedro et al. 2014).

Additionally, there are continuous developments to the global proxy networks with new coral and speleothem records from the Pacific (Chen et al. 2016), continuous updates to proxy networks and databases such as the tree-rings within the Old World Drought Atlas (Cook et al. 2015) and LinkedEarth, a global multi-proxy database for the Common Era (Emile-Geay et al. *in press*). There is scope to extend reanalysis data further with observations that can be recovered from data rescue projects, including logbook observations (Compo et al. 2011). An example of this is an experimental atmospheric reanalysis carried out over Europe using a limited number of surface pressure observations and observations from ships for 1816, the ‘year without a summer’ (Brohan et al. 2012). However, there is still much work needed to extend reanalysis data further back in time. As seen in Chapter 5, historical reanalysis has a reduced number of observations and increased disagreement and uncertainty in the earlier part of the twentieth century.

There is scope to expand on the use of historical reanalyses within the reconstruction of past climates, with continuous improvements being made to reanalyses and with the expansion of reanalyses further back in time. Analysis within Chapter 5 used the ensemble output from the Twentieth Century Reanalysis Project, which represents the mean of 56 realisations (Compo et al. 2011). Further research could exploit this range of outputs by comparing climatic phenomenon within the 56 individual realisations.

Overall, it can be seen that there are a number of data sources involved in the investigations of climatic variability of the past. The logbook-based reconstruction has

provided a valuable addition to this continuously expanding collection. Future challenges remain in distinguishing the best methods for comparisons of these diverse ENSO reconstructions and in addressing the limitations of the various sources of data. This thesis has provided the most up-to-date comparison of a large range of ENSO reconstructions and suggested ways in which these can be expanded upon in the future.

9.3 Stationarity of the ENSO-wind teleconnection

The use of the teleconnection between ENSO and zonal wind is another key theme within the logbook-based reconstruction (Chapter 2), which is explored in more detail using reanalysis and climate model simulations (Chapter 7 and 8). The climate models provide a valuable tool for assessing stationarity over a longer time period and in a more holistic way than available from proxy records. Therefore, climate model simulations provide a useful addition to the proxy analysis carried out in previous chapters of the thesis. Chapter 2 used the teleconnection between zonal wind and the SOI during the boreal winter season to reconstruct ENSO. The use of zonal winds was directed by the availability of observations within the digitised logbook databases. As discussed in Chapter 7, the analysis of the ENSO-wind teleconnection is often overlooked with focus on temperature and precipitation dominating previous ENSO-teleconnection analysis. Climate model and proxy comparisons initially focused on temperature and now explore hydroclimate more widely (Smerdon et al. 2017). However, winds are still an under-explored teleconnection, but have been focused on throughout this thesis.

9.3.1 The assumption of stationarity

Significant correlations between zonal wind and the DJF SOI were found in multiple regions due to the large extent of ENSO teleconnections (Fig. 2.5). A key assumption of the logbook-based ENSO reconstruction was that of stationarity in these teleconnections. This is a common and necessary assumption of paleoclimate reconstructions (Gallant et al. 2013; Phipps et al. 2013; Batehup et al. 2015). However, it is an implicit, and somewhat tenuous assumption which can be seen as a limitation of this type of research (Gallant et al. 2013). The use of climate models over longer time-scales enables a full assessment of the long-term stability of the teleconnections associated with ENSO (Wittenberg 2009). Climate models are not perfect and the

strength and spatial structure of ENSO teleconnections can be poorly represented (Gallant et al. 2013). This can be a result of issues in the representation of mean climate and/or in the representation of ENSO and other modes of climate variability (Batehup et al. 2015). Therefore, the conclusions from climate model teleconnections cannot be expected to be fully consistent with real-world teleconnections (Batehup et al. 2015). Thus the model biases in the representation of climatology and ENSO representation were summarized in Chapters 6 and 7, and only the best two models were used to analyse the stationarity over the Last Millennium (GISS-E2-R and IPSL-CM5A-LR).

The stationarity of the ENSO-wind teleconnection was assessed over the entire twentieth century using reanalysis data (Fig. 7.16) and over the LM using the output from climate model simulations (Fig. 8.15; Fig. 8.16). The stationarity of the ENSO-wind teleconnection was assessed on a spatial scale focused on the logbook-based reconstruction within the predictor grid boxes used, and therefore is specific to this reconstruction. Further work, using similar methods to assess the stationarity of the teleconnections in other regions would be needed if carrying out a reconstruction based on different spatial coverage.

9.3.2 Stationarity over the twentieth century

The ERA-20C reanalysis suggested that the 100-year correlation (1900-1999) between DJF zonal wind and the DJF SOI (Fig. 7.13), has a similar spatial pattern to that from ERA-Interim (1979-2013) on which the logbook-based reconstruction was based (Fig 2.5), but with the strength of the correlations reduced over this longer time period. The stability of the teleconnection varied temporally over the twentieth century within the predictor grid boxes (Fig 7.16). The strength of the correlation coefficients increased with time in grid boxes 1, 2 and 3, whereas in grid boxes 4, 5, 6 there was a clear dip in the strength of the correlation focused around the 1940s. Part of this could be due to the reduced amount of data assimilated in the earlier part of the reanalysis (Poli et al. 2016), or due to a dip in availability during the Second World War. The changing nature of the quality in the reanalysis over the twentieth century hinders robust conclusions on the changes in the strength of the ENSO-wind teleconnection. Therefore, attention was then focused on climate model simulations.

Chapters 6 and 7 showed that climate model simulations have differences in their representation of climatology and in ENSO spatial patterns. Even though climate models are forced by similar natural forcings over the LM and historical periods, differences arise due to the physics and numerical algorithms of the different models used to run the simulations (Taylor et al. 2012; Schmidt et al. 2014). Of the five models analysed here, clear biases were found within some of the models such as a warm SST climatological bias in the eastern tropical Pacific in GISS-E2-R (Fig. 6.8) and low MSLP climatological anomalies in HadCM3 across the equatorial Pacific (Fig. 6.3). Despite these biases climate models provide a useful tool for examining teleconnections over long time periods (Phipps et al. 2013). The use of metrics to rank the models identified those which captured ENSO best, and thus IPSL-CM5A-LR and GISS-E2-R were used for further analysis which spanned over the Last Millennium. Differences in the running correlations between the DJF SOI and zonal wind in the predictor grid boxes were found over the twentieth century when comparing results from ERA-20C, GISS-E2-R and IPSL-CM5A-LR (Fig. 7.17), suggesting some disagreement between the model and reanalysis and also between model-model representation of this teleconnection. GISS-E2-R was found to capture the patterns of the teleconnections best out of the models over this twentieth century (Table 7.6). Nevertheless, these two models were also analysed over the LM as they provide the only data source which can provide an estimate of the changes in ENSO-wind teleconnection in the predictor grid boxes over such large temporal scale.

9.3.3 Stationarity over the Last Millennium

The analysis of the GISS-E2-R and IPSL-CM5A-LR Last Millennium simulations in Chapter 8 found that there was non-stationarity in the ENSO-wind teleconnection in the predictor grid boxes, with the strength of the teleconnection varying over time in both of the model simulations analysed (Fig. 8.15; Fig. 8.16). Two periods, 1050-1149 and 1750-1848, were investigated and the strength of the correlations were weaker in these periods than in the historical simulations (1900-1999) (Fig. 7.14; Fig. 8.12; Fig. 8.13). The correlations were also generally weaker in the MWP (1050-1149) compared to 1750-1848. However, GISS-E2-R, which captured the correlations in the predictor grid boxes better than IPSL-CM5A-LR in the historical period, found significant correlations in a number of the grid boxes in both the 100-year periods

assessed over the LM. Within GISS-E2-R the correlations in the grid boxes over the eastern Indian Ocean have periods of significant correlations over 1750-1848, whereas the grid boxes around South Africa, for example, are less stable, and had very weak or no correlation over both periods. Whereas IPSL-CM5A-LR found significant correlations in grid boxes in both the South Africa region and the Indian Ocean in 1750-1848, however in none of the predictor grid boxes in the MWP. Overall, the LM analysis identifies that teleconnections between zonal wind and ENSO are not stable over time within the predictor grid boxes, with the strength of the correlations generally remaining higher in the eastern Indian Ocean grid boxes. However, these results are limited due to them being largely model-dependent, with the two models used over the LM producing differences in results. Further work is needed to improve the LM climate simulations in order to fully utilise their output.

9.3.4 Implications of non-stationarity and future work

The stationarity of teleconnections varies temporally in different regions and thus using proxy records from various regions is more likely to capture an ENSO signal (Lewis and LeGrande 2015). Batehup et al. (2015) found that using global, randomly spaced proxy networks the pseudo-proxy ENSO reconstructions were not sensitivity to non-stationary teleconnections. Therefore, the best way to overcome the issue of non-stationarity influencing reconstructions is to use proxies from a number of teleconnection regions, so as to contain a larger ratio of climate signal to local noise (Batehup et al. 2015). The logbook-based reconstruction uses proxies from a number of teleconnection regions, including regions closer to the centre of action for ENSO, thus capturing a stronger signal from this region. This is complemented with use of additional ENSO-teleconnections from more remote regions, but regions which still have a strong ENSO signal in the modern record.

For the predictor grid boxes used in logbook-based reconstruction, Chapter 8 suggested that there was more consistent stationarity in the ENSO-wind teleconnection in the eastern Indian Ocean region (grid box 7-9). This was especially the case in GISS-E2-R simulation. However, the use of only this region would degrade the robustness of the results due to the lower number of predictors. The best way to overcome the challenge of the assumption of stationarity is to use a diverse

teleconnection region in order to capture the common signal from a number of regions. Therefore, at least one other teleconnection region should be used, however this is currently limited by data availability. The South Africa region was found to be the best supplementary region, with good data availability and a strong ENSO-teleconnection in the modern record, and a differing signal between EP and CP events (Fig. 4.1). Further digitization could focus on additional regions with strong ENSO-signal such as the northern Indian Ocean as indicated from Fig. 2.5, or Fig. 8.12-8.14 suggested stability in ENSO-zonal wind teleconnections in the Gulf of Alaska and the western Caribbean over periods of the LM. However, in reality this is limited to regions in which historical observations were taken, dictated by the old shipping routes. It is suggested that future research explores the stationarity of the ENSO-wind teleconnection within LM climate model simulation in multiple regions in order to inform further logbook-based reconstructions. For the purpose of future ENSO reconstructions, digitisation efforts should be focused on regions in which there are strong and more stable correlations between ENSO and zonal wind.

9.4 ENSO Diversity

A final key theme which has emerged within a number of chapters of this thesis is ENSO diversity. This is important as there has been an increase in the frequency of CP events recently, and this is projected to do so in future climate scenarios (Yeh et al. 2009; Lee and McPhaden 2010). It is important to understand the types of El Niño due to the differences in their teleconnections and impacts (Yeh et al. 2014). ENSO diversity remains a topic with a number of unresolved questions and is an active research area (Yeh et al. 2014). Here, methods of classifying the types of El Niño events are discussed followed by the representation of ENSO diversity within the various ENSO reconstructions. ENSO diversity within climate models are then discussed.

9.4.1. Methods of identifying and classifying El Niño types

As with defining El Niño and La Niña events from the instrumental record (Section 9.2.5), there remains a lack of consensus on the best method of classifying the different types of El Niño events. Chapter 4 discussed the range of methods currently used for defining the two types of El Niño events. Capotondi et al. (2015) listed seven

commonly used indices of ENSO diversity which define events using a range of different variables and methods, showing the range of classification criteria adopted by various research projects. Pascolini-Campbell et al. (2015) applied nine different classification methods to five different SST reconstruction/reanalysis datasets in order to compare the differences between classification schemes. Singh (2012) provides a comprehensive overview of the use of different variables and different methodologies to explore ENSO diversity. New methods are continually being introduced, for example, Yu and Kim (2013), presented a new methodology with the aim of providing an alternative to existing methods, but acknowledge that their aim was not to produce a 'best' method but rather to add an additional method to the mix. The lack of consensus on methodology is a key challenge to categorising El Niño diversity within the instrumental record, ENSO reconstructions and climate model simulations. It limits the conclusions which can be drawn from studies investigating ENSO diversity, and therefore a more universally adopted and robust classification scheme is needed (Pascolini-Campbell et al. 2015).

Within this thesis, one of these methodologies has been used throughout in order to provide consistency between analysis of ENSO reconstructions and climate model simulations. This method, here termed the NINO method, initially used by Kug et al. (2009), was chosen due to its ability to capture ENSO diversity back to 1870, because of the long record of SST (e.g. HadISST), whereas methods which use variables such as outgoing longwave radiation (Chiodi and Harrison 2013) cannot be extended as far back. It is also a commonly used method of ENSO diversity classification (Yu et al. 2012; Yeh et al. 2014; Capotondi et al. 2015). However, it must be appreciated that selection of different methodologies would result in ENSO diversity chronologies which are not exactly matched, as found by Yu et al. (2012) who tried to gain a consensus between multiple methods over the modern period 1950-onwards. In order to investigate differences between chronologies of ENSO diversity, a record from Schollaen et al. (2015), one of the first ENSO proxy reconstruction to focus on ENSO diversity, was compared with the Kug et al. (2009) method, and although differences in chronology were found, the general conclusions gained from the results of both analyses were in agreement with each other (Chapter 4). The range of methods is a

key challenge for future work in this area that needs addressing in order to create a consensus in methods which will enable more direct comparisons between datasets.

9.4.2. ENSO diversity in reconstructions

Another challenge is that it has been widely noted that EP and CP El Niño events can have different teleconnections in different regions (Yeh et al. 2014; Schollaen et al. 2015; Infanti and Kirtman 2016). Key variables such as temperature, precipitation and salinity have distinct teleconnection patterns associated with the two types of El Niño and therefore it is possible that proxies are recording climate variability associated with only one type of event, or a combination of effects, making interpretation of the signal challenging (Karamperidou et al. 2015). For example, in the western Pacific temperature teleconnections are stronger in CP events, whereas salinity effects are stronger during EP events, therefore mixed signals are present in the coral-proxy records for this region (Karamperidou et al. 2015). Thus, reconstructions which are fitted to indices which are more likely to identify EP or CP events would result in differences in the ENSO reconstruction they produce.

The two reconstruction methods applied to the logbook data were found to have different skill at reconstructing EP and CP El Niño. This is a key finding, which has not been explored using this data source prior to this thesis research. Results in Chapter 4 show that during the fitting period instrumental SOI, the PCR and the CPS fitting capture the EP El Niño events defined by the NINO method better than the CP El Niño events. The spatial zonal wind anomalies were more widespread in the predictor grid boxes during EP events rather than CP events. A bias towards capturing EP El Niño events compared to CP El Niño events would lead to a logbook-based reconstruction which could be favoured towards capturing EP events. During the period of the logbook-based reconstruction only one out of the five El Niño events identified by CPS B is suggested to be a CP event, while two out of three identified by PCR B are suggested to be CP events. Therefore, the CPS methodology appears to have a larger EP bias than the PCR method. Fig. 4.1 showed that the zonal wind anomalies in the predictor grid boxes around South Africa showed the opposite signal anomaly during EP and CP events, which could be exploited further to help distinguish between events. It is suggested that future ENSO reconstructions using data from ships' logbooks

should be aware of the potential biases in the methodologies as a result of the differences in teleconnections between EP and CP events and should take this into account in the selection of predictor grid boxes and in the selection of reconstruction methodologies.

In Chapter 4 it was found that all but one of the multi-proxy and documentary reconstructions capture EP El Niño event better than CP El Niño events (Table 4.4; Table 4.5), suggesting that they would also have a bias towards capturing EP El Niño events. This was explored further by analysing the spatial teleconnections associated with the different types of El Niño event and the locations of proxy records. It was found that the documentary El Niño chronologies, based largely on records of flooding and heavy rainfall in coastal, equatorial South America, a region in which only EP El Niño events are associated with anomalously heavy rainfall (Fig 4.6), and thus more likely to recording EP El Niño events. For the multi-proxy records, larger networks were used which covered multiple teleconnection regions. The main difference in precipitation and temperature anomalies during EP and CP El Niño events in the North American tree-ring region are increases in the strength of the anomaly during EP events compared to CP El Niño events (Fig. 4.7; Fig. 4.8). This suggests that the signal from the tree-rings in north America, commonly used in multi-proxy reconstructions, favour the capture of EP El Niño events. Looking at corals and the variables to which they are sensitive to, different teleconnection patterns were found in a number of regions for EP and CP El Niño events. Overall, the combination of signals from the different regions and the different types of proxy records used has resulted in a bias towards EP El Niño events in the majority of the ENSO reconstructions assessed in this thesis.

In order to provide the most robust ENSO reconstruction, proxy networks containing records from various teleconnections are used. However, this can still lead to an underlying bias towards one type of event if not explicitly addressed. The differences in teleconnection between the events should be considered prior to carrying out multi-proxy reconstructions. Investigations of proxy records in regions with known differences in teleconnections such as Schollaen et al. (2015), are valuable for exploiting information of ENSO diversity in the past, and should be encouraged in the

future in order to exploit differences in teleconnections within the same region and to create proxy reconstructions of EP and CP El Niño events.

9.4.3. ENSO diversity in climate models

ENSO diversity was then explored with the climate model simulations. Previous studies have investigated ENSO diversity within historical simulations (Yeh et al. 2009; Ham and Kug 2012; Kim and Yu 2012; Xu et al. 2017); this is expanded on in this thesis looking at ENSO diversity over the Last Millennium. Over the historical period there are a vast number of model simulations available compared to the limited number of LM model simulations available, therefore multi-model means are often discussed. Yeh et al. (2009) found that the CP-EP occurrence ratio in the observations was much smaller than the mean of the models they tested, 1854-2007 (i.e. more models had relatively more CP events than observations) due to cold bias which results in warm tongue anomalies being shifted further west in models. The analysis of the five model simulations in Chapter 6 over the twentieth century showed that while some of the models had a higher CP-EP ratio than observations, others had a lower CP-EP ratio (Table 6.13). Therefore, there was inconsistency in the frequency of EP and CP El Niño events within the historical climate model simulations, a spread in results which was also seen by Yeh et al. (2009). Their ability to capture the spatial extent and the strength of the MSLP and surface temperature anomalies also varied between the models. Two models were found to have the best representation of the models analysed and therefore were used to analysis over the LM, despite their biases identified in Chapter 6.

The investigation of the frequency and spatial patterns of EP and CP El Niño events using these two models (GISS-E2-R and IPSL-CM5A-LR), found that, although the two models can produce the spatial temperature anomaly patterns associated with the two types of events relatively well, showing similarity to twentieth century spatial anomaly patterns, the anomalies are weaker in the 1750-1848 period compared to the 1900-1999 period, (Fig. 6.15 And Fig. 8.20). The CP-EP ratio over the ten 100-year periods of the last millennium showed most variation in IPSL-CM5A-LR, with this model identifying a high number of CP events, a bias found in Chapter 6 in the modern period (Table E2). GISS-E2-R captured the frequency and CP-EP ratio best when

compared to ERA-20C, therefore its LM results are less likely to be representing the bias seen in IPSL-CM5A-LR. GISS-E2-R suggests that the CP-EP ratio found in the 1900-1999 period has only been higher towards CP events in 3 100-year periods in the LM (1050-1149, 1050-1249, 1450-1549 AD). Further work is needed in order to use the LM simulations to be able to explore ENSO diversity. Crucially, biases within the models need to be addressed, as only once the models can replicate recent climate well compared to observed, can they be used fully to explore further in the past and into the future.

The development of climate models is a slow and incremental process. There are continued efforts in order to try and address inconsistencies within the models and reduce climate model biases (Schmidt et al. 2014). CMIP projects are a good opportunity to take stock of current progress within climate models and to assess the development needs for the future (Schmidt et al. 2014). Therefore, following the success of CMIP5, CMIP6 was launched, which uses the most updated version of models and model forcings to address key climate questions (Eyring et al. 2016). Thus, although this thesis utilises the most up-to-date data sources available at the time of carrying out the research, due to the nature of the expanding and continuously developing research field there will always be the opportunity to expand on these comparisons further in the future.

9.4.4. Future studies on ENSO diversity

It is important to understand ENSO diversity due to a recent emergence of an increased frequency of CP events, and an increase in their intensity (Lee and McPhaden 2010). The ratio of CP- to EP- events is projected to increase under global warming scenarios investigated in CMIP5 (Kim and Yu 2012). New climate model analysis of CMIP6-PMIP4 from the PMIP Paleo-variability Working Group will include a focus on ENSO diversity and the stability of teleconnections within the climate system (Kageyama et al. 2016). This highlights that these two issues, which have been explored within this thesis, are two key areas in which further modelling work is needed. It is important to understand ENSO diversity and ENSO teleconnections both for helping to interpret ENSO reconstructions of the past and also in order to predict impacts of future events. Only once climate models have shown that they are capable of simulating climate in

the past can there be increased confidence in the ability of the models to project future climates (Cane et al. 2006).

In addition to the assessment of El Niño diversity, future work is needed to clarify the diversity of La Niña events. The existence of two distinct types of La Niña events, as argued for El Niño events, is not as clear and remains a controversial issue (Song et al. 2016). Some studies identify EP and CP La Niña using the spatial distribution of SST anomalies (Zhang et al. 2014), as is commonly applied to El Niño diversity, whereas others argue that evidence of CP-type La Niña is not apparent (Ren and Jin 2011). To answer future questions regarding the diversity of ENSO ‘further studies are needed using observational and reanalysis datasets, paleoclimate records, and model simulations’ (Capotondi et al. 2015 pp. 932). This thesis addressed some unanswered questions regarding ENSO diversity using this broad range of datasets, however further work is needed to fully understand the complexity in the variation of ENSO events.

9.5 Conclusion

This chapter has summarised the themes of this thesis, discussed the key limitations and presented the implications of this research for future research. Drawing together the findings from throughout this thesis presents an up-to-date, holistic understanding of ENSO during the early to mid-nineteenth century, a crucial period which lies prior to the advent of routine meteorological observations on land. The ENSO reconstruction from ships’ logbooks is proven to be a valuable source of information on ENSO, constructed using direct observations of the weather.

The defining aim of this thesis was to explore ENSO during the pre-instrumental, historical period in order to gain a better understanding of ENSO behaviour on longer time scales than available from the instrumental record. This is important for assessing natural changes in ENSO, to place current ENSO behaviour in a historical context and to help inform possible future changes in ENSO. It is a highly relevant topic as ENSO has large socio-environmental impacts: wildfires, flooding; smog; drought; famine. The first main aim was to reconstruct ENSO using observations from ships’ logbooks, an under-exploited source of information on the climate of the past. This is important as it has provided a new reconstruction of ENSO, highlighted the utility of this data

source for historical climatology, and the value of future digitisation in order to fully exploit this historical resource further.

The second main aim was to compare the logbook-based reconstruction to previous ENSO reconstructions. A key result was the comparison with the Jakarta rain-day and pressure-based SOI which showed good agreement between reconstructions produced using direct routine observations of the weather. There is still a lack of consensus of ENSO event chronology from the range of reconstructions assessed during the reconstruction period, 1815-1854, due to a number of possible limitations discussed throughout this thesis. Importantly the ability of the reconstructions to capture El Niño events was found to be generally higher than their ability to capture La Niña events, and in terms of ENSO diversity, the reconstructions generally capture EP El Niño events better than CP El Niño events. This is important as the various ENSO reconstructions could be providing a skewed interpretation of historical and paleo-ENSO.

ENSO variability over the twentieth century was explored using reanalysis data and historical climate model simulations. Firstly, this identified a difference in the reanalysis in the earlier half of the twentieth century which is likely due to the differences in the observations assimilated in the reanalysis and the different data assimilation methods used for the different reanalysis products. This suggests that the use of ERA-20C and 20CRv2 for investigating ENSO spanning the entire twentieth century would require additional levels of uncertainty than assessed here, with the use of ERA-Interim as the key reanalysis to inform the historical reconstruction. Differences in climatology and ENSO variability and events within historical climate model simulations were found over the twentieth century and ranking methods provided an indication of the most suitable models to use analysis over the Last Millennium.

Finally, the use of climate model simulations of the LM proved a useful tool for assessing the spatial and temporal changes in ENSO over a period which is not covered by the instrumental record and which provides a more comprehensive representation of ENSO dynamics than provided by proxy records. It was found that the assumption of stationarity in the ENSO-wind teleconnection is one that should be treated with

caution, and that the use of records from multiple regions is needed to provide the most robust ENSO signal. Further work is needed in model improvement in order to draw full conclusions on ENSO diversity over the Last Millennium.

Overall this thesis has explored El Niño Southern Oscillation during the pre-instrumental period using a wide range of data sources and methodologies. It has tackled key concepts such as ENSO diversity and ENSO-wind teleconnections, as well as investigating differences in reconstruction methodologies and developing new methods of using the observations from ships logbooks to inform pre-instrumental periods. A number of limitations have been discussed and suggestions for future research have been presented, as only once a fuller understanding of ENSO in the past is obtained can current and future ENSO be placed within its true historical context.

Bibliography

AchutaRao K, Sperber KR (2002) Simulation of the El Niño Southern Oscillation: Results from the Coupled Model Intercomparison Project. *Clim Dyn* 19:191-209.

Adams BJ, Mann ME, Ammann CM (2003) Proxy evidence for an El Niño-like response to volcanic forcing. *Nature* 426: 274-278.

Adler RF, Huffman GJ, Chang A, Ferraro R, Xie P, Janowiak J, Rudolf B, Schneider U, Curtis S, Bolvin D, Gruber A, Susskind J, Arkin P (2003) The Version 2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979-present). *J. Hydrometeorol.* 4:1147–1167.

Aldrain E, Susanto RD (2003) Identification of three dominant rainfall regions within Indonesia and their relationship to sea surface temperature. *Int J Climatol* 23:1435-1452.

Allan R, Brohan P, Compro GP, Stone R, Luterbacher J, Bronnimann S (2011) The International Atmospheric Circulation Reconstruction over the Earth (ACRE) initiative. *Bull Amer Met Soc* 92:1421-1425.

Allan R, Lindsay J, Parker D (1996) El Niño Southern Oscillation and climate variability. CSIRO: Melbourne.

Allan RJ (2000) ENSO and Climatic Variability in the Past 150 years. In *El Niño and the Southern Oscillation: Multiscale Variability and Global and Regional Impacts* [eds. Diaz HF, Markgraf V]. Cambridge University Press: Cambridge.

Ammann CM, Joos F, Schimel DS, Otto-Bliesner BL, Tomas RA (2007) Solar influence on climate during the past millennium: results from transient simulations with NCAR Climate System Model. *Proc Natl Acad Sci* 104(1):3713-3718.

Ampou EE, Johan O, Menkes CE, Nino F, Birol F, Ouillon S, Andrefouet S (2017) Coral mortality induced by the 2015-2016 El Niño in Indonesia: the effect of rapid sea level fall. *Biogeosciences* 14:817-826.

Andreoli RV, Oliveira SS, Kayano MT, Viegas J, de Souza RAF, Candido LA (2016) The influence of different El Niño types on the South American rainfall. *Int J Climatol* doi:10.1002/joc.4783.

Ashok K, Behera SK, Rao SA, Weng H, Yamagata T (2007) El Niño Modoki and its possible teleconnections. *J. Geophys. Res.* 112 C11007, doi: 10.1029/2006JC003798.

Barrett HG, Jones JM, Bigg GR (2017) Reconstructing El Niño Southern Oscillation using data from ships' logbooks, 1815-1854. Part I: Methodology and Evaluation. *Clim Dyn* DOI:10.1007/s00382-017-3644-7.

Barrett HG, Jones JM, Bigg GR (2017) Reconstructing El Niño Southern Oscillation using data from ships' logbooks, 1815-1854, Part II: Comparisons with existing ENSO reconstructions and implications. *Clim Dyn* DOI:10.1007/s00382-017-3797-4.

Barriopedro D, Gallego D, Alvarez-Castro MC, García-Herrera R, Wheeler D, Peña-Ortiz C, Barbosa, SM (2014) Witnessing North Atlantic westerlies variability from ships logbooks (1685-2008). *Clim Dyn* 43: 939-955.

Barry RG, Chorley R (2010) *Atmosphere, weather and climate*. 9th Edition. Routledge: London.

Batehup R, McGregor S, Gallant AJE (2015) The influence of non-stationary teleconnections on palaeoclimate reconstructions of ENSO variance using a pseudoproxy framework. *Clim Past* 11:1733-1749.

Bellenger, H., Guilyardi, É., Leloup, J., Lengaigne, M., & Vialard, J. (2014). ENSO representation in climate models: from CMIP3 to CMIP5. *Clim Dyn* 42(7-8):1999-2018.

Bett PE, Thornton HE, Clark RT (2017) Using the Twentieth Century Reanalysis to assess climate variability for the European wind industry. *Theor Appl Climatol* 127(1):61-80.

Bjerknes, J (1969) Atmospheric teleconnections from the equatorial Pacific. *Mon Wea Rev* 97:163-172.

Braconnot P, Harrison SP, Kageyama M, Bartlein PJ, Masson-Delmotte V, Abe-Ouchi A, Otto-Bliesner B, Zhao Y (2012) Evaluation of climate models using palaeoclimatic data. *Nature Climate Change* 1456.

Bradley RS, Jones PD (1992) *Climate Since A.D. 1500*. Routledge, New York.

Braganza K, Gergis JL, Power SB, Risbey JS, Fowler AM (2009) A multiproxy index of the El Niño-Southern Oscillation, A.D. 1525-1982. *J Geophys Res* 114 D05106.

Broad K, Pfaff ASP, Glantz MH (2002) Effective and Equitable Dissemination of Seasonal-to-Interannual Climate Forecasts: Policy Implications from the Peruvian Fishery during El Niño 1997–98. *Climatic Change*. 54:415-438.

Brohan P, Allan R, Freeman E, Wheeler D, Wilkinson C, Williamson F (2012) Constraining the temperature history of the past millennium using early instrumental observations. *Clim Past* 8:1551-1563.

Brown J, Tudhope AW, Collins M, McGregor HV (2008) Mid-Holocene ENSO: Issues in quantitative model-proxy data comparisons. *Palaeoceanography* 23: PA3202.

Brown JN, Langlais C, Maes C (2013) Zonal structure and variability of the Western Pacific dynamic warm pool edge in CMIP5, *Clim Dyn* 42, 3061–3076.

Brown JR, Hope P, Gergis J, Henley BJ (2016) ENSO teleconnections with Australian rainfall in coupled model simulations of the last millennium. *Clim Dyn* 47(1-2):79-93.

Bunge L, Clarke AJ (2010) A Verified Estimation of the El Niño Index Niño-3.4 since 1877. *J Climate* 22(14):3979-3992.

Bureau of Meteorology (2015) Southern Oscillation Index (SOI) since 1876. <http://www.bom.gov.au/climate/current/soihtml1.shtml> Accessed [21/10/2015].

Bureau of Meteorology (2017) Southern Oscillation Index (SOI) since 1876. Accessed at <http://www.bom.gov.au/climate/current/soihtm1.shtml> Accessed [06/06/2017].

Cai W et al. (2015) ENSO and greenhouse warming. *Nature Climate Change* 5, 849-859.

Cane M, Braconnot P, Clement A, Gildor H, Joussaume S, Kageyama M, Khodri M, Paillard D, Tett S, Zorita E (2006) Progress in Paleoclimate Modelling. *Journal of Climate* 19:5031–5057.

Capotondi A, Wittenberg AT, Newman M, Di Lorenzo ED, Yu JY, Braconnot P, Cole J, Dewitte B, Giese B, Guilyardi E, Jin FF, Karnauskas, Kirtman B, Lee T, Schneider N, Xue Y, Yeh SW (2015) Understanding ENSO Diversity. *BAMS*:921-938.

Cashin P, Mohaddes K, Raissi M (2014) Fair Weather or Foul? The Macroeconomic Effects of El Niño. Cambridge working papers in Economic. Accessed at <http://www.econ.cam.ac.uk/research/repec/cam/pdf/cwpe1418.pdf>. Accessed [06/06/2017].

Catchpole AJW, Faurer MA (1985) Ships' Log-Books, Sea Ice and the Cold Summer of 1816 in Hudson Bay and Its Approaches. *Arctic* 38(2):121-128.

Chen S, Hoffmann SS, Lund DC, Cobb KM, Emile-Geay J, Adkins JF (2016) A high-resolution speleothem record of western equatorial Pacific rainfall: Implications for Holocene ENSO evolution. *Earth and Planetary Science Letters* 442:61-71.

Chenworth M (2001) Two major volcanic cooling episodes derived from global marine air temperature, AD 1807-1825. *Geophysical Research Letters* 28(15):2963-2966.

Chiodi AM, Harrison DE (2013) El Niño impacts on seasonal U.S atmospheric circulation, temperature, and precipitation anomalies. The OLR-events perspective. *J. Clim* 26:822-837.

Christensen JH, Krishna Kumar K, Aldrian E, An SI, et al. (2013) Climate Phenomena and their Relevance for Future Regional Climate Change. In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Stocker TF, Qin D, Plattner GK, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Coats S, Smerdon JE, Cook BI, Seager R (2013) Stationarity of the tropical Pacific teleconnection to North America in CMIP5/PMIP3 model simulations. *Geophys Res Letts* 40:1-6.

Cobb KM, Charles CD, Cheng H, Lawrence Edwards R (2003) El Niño/Southern Oscillation and tropical Pacific climate during the Last Millennium. *Nature* 424:271-276.

- Cole JE, Rind D, Fairbanks RG (1993) Isotopic responses to interannual climate variability simulated by an atmospheric general circulation model. *Quat Sci Rev* 12: 387-406.
- Collins M, Tett SFB, Cooper C (2001) The internal climate variability of HadCM3, a version of the Hadley Centre coupled model without flux adjustments. *Clim Dyn* 17(1):61–81.
- Compo GP, Whitaker JS, Sardeshmukh PD, Matsui N, Allan RJ, Yin X, Gleason BE, et al (2011) The Twentieth Century Reanalysis Project. *Quarterly J Roy Met Soc* 137: 1-28.
- Cook ER, Briffa KR, Jones PD (1994) Spatial regression methods in dendroclimatology – a review and comparison of two techniques. *Int J Climatol* 14(4):379-402.
- Cook ER, Seager R, Kushnir Y, Briffa KR, et al. (2015) Old World megadroughts and pluvials during the Common Era. *Science Advances* 1(10).
- Cram TA, and co-authors (2015) The International Surface Pressure Databank version 2. *Geosci Data J*, 2:31–46.
- Crowley T, Zielinski G, Vinther B (2008) Volcanism and the Little Ice Age:pp. 4–5. <http://dx.doi.org/10.1029/2002GL0166335>.
- Crowley TJ, Untermann MB (2013) Technical details concerning development of a 1200 yr proxy index for global volcanism. *Earth Syst Sci Data* 5(1):187-197.
- Dai N, Arkin PA (2017) Twentieth century ENSO-related precipitation mean states in twentieth century reanalysis, reconstructed precipitation and CMIP5 models. *Clim Dyn* 48(9-10):3061-3083.
- Dee DP and co-authors (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q J Roy Meteor Soc* 137:553–597.
- Diaz HF, Hoerling MP, Eischeid JK (2001) ENSO variability, teleconnections and climate change. *Int J Clim* 21:1845-1862.
- Diaz HF, Markgraf V (1992) El Niño. Historical and Paleoclimatic Aspects of the Southern Oscillation. Cambridge, New York, Port Chester, Melbourne, Sydney: Cambridge University Press.
- Dufresne, J.-L. and coauthors (2013) Climate change projections using the IPSL-CM5 Earth System Model: from CMIP3 to CMIP5. *Clim Dyn* 40(9-10):2123-2165.
- Dunbar RB, Wellington GM, Colgan MW, Glynn PW (1994) Eastern Pacific sea surface temperature since 1600 AD: the $\delta^{18}\text{O}$ record of climate variability in Galapagos corals. *Palaeoceanography* 9:291-316.
- Emile-Geay J, Cobb KM, Mann ME and Wittenberg AT (2013a) Estimating Central Equatorial Pacific SST Variability over the Past Millennium. Part I: Methodology and Validation. *J Clim* 26:2302-2328.

Emile-Geay J, Cobb KM, Mann ME, Wittenberg AT (2013b) Estimating Central Equatorial Pacific SST Variability over the Past Millennium. Part II: Reconstructions and Implications. *J Clim* 26:2329-2352.

Emile-Geay J, McKay NP, Kaufman, DS et al. (2017) A global multiproxy database for temperature reconstructions of the Common Era. *Scientific Data*. ISSN 2052-4463 (In Press).

Emile-Geay J, Seager R, Cane MA, Cook ER and Haug, GH (2008) Volcanoes and ENSO over the Past Millennium. *J Clim* 21:3134-3148.

Evans MN, Kaplan A, Cane MA (2002) Pacific sea surface temperature field reconstruction from coral $\delta^{18}\text{O}$ data using reduced space objective analysis. *Paleoceanography* 17(1):1-13.

Eyring V, Bony S, Meehl GA, Senior CA, Stevens, B and coauthors (2016) Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development, European Geosciences Union* 9 (5):1937-1958.

Farrington AJ, Lubker S, Radok U, Woodruff S (1998) South Atlantic winds and weather during and following the Little Ice Age – A pilot study of English East India Company (EEIC) ships logs. *Meteorology and Atmospheric Physics* 67:253-257.

Ferguson CR, Villarini G (2014) An evaluation of the statistical homogeneity of the Twentieth Century Reanalysis. *Clim Dyn* 42(11):2841-2866.

Fiorino M (2004) A Multi-decadal Daily Sea Surface Temperature and Sea Ice Concentration Data set for the ERA-40 Reanalysis. ERA-40 Project report series No. 12. Accessed at <http://www.ecmwf.int/sites/default/files/elibrary/2004/9396-multi-decadal-daily-sea-surface-temperature-and-sea-ice-concentration-data-set-era-40.pdf> Accessed [28/02/2017].

Flato G, Marotzke J, Abiodun B, Braconnot P, Chou SC, Collins W, Cox P, Driouech F, Emori S, Eyring V, Forest C, Gleckler P, Guilyardi E, Jakob C, Kattsov V, Reason C, Rummukainen M (2013) Evaluation of Climate Models. In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Stocker TF, Qin D, Plattner GK, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Freeman E, Woodruff S, Worley SJ and co-authors (2017) ICOADS Release 3.0: a major update to the historical marine climate record. *Int J Clim* 37(5):211-2232.

Gagan MK, Hendy EJ, Haberle SG, Hantoro WS (2004) Post-glacial evolution of the Indo-Pacific Warm Pool and El Niño-Southern Oscillation. *Quaternary International* 118-119:127-143.

Gallant AJE, Phipps SJ, Karoly DJ, Mullan AB, Lorrey AM (2013) Nonstationary Australasian Teleconnections and Implications for Paleoclimate Reconstructions. *J Clim* 26:8827-8849.

Gallego D, Garcia-Herrera R, Ribera P, Jones PD (2005) Seasonal mean pressure reconstruction for the North Atlantic (1750–1850) based on early marine data. *Clim Past* 1:19-33.

Gallego D, Ordóñez-Pérez P, Ribera P, Peña-Ortiz C, Garcia-Herrera R (2015) An instrumental index of the West African Monsoon back to the nineteenth century. *Quart J Roy Meteor Soc* 141(693):3166-3176.

García RR and coauthors (2001) Atmospheric circulation changes in the tropical Pacific inferred from the voyages of the Manila galleons in the sixteenth-eighteenth centuries. *Bull Amer Meteor Soc* 82:2435-2456.

García-Herrera R, Barriopedro D, Hernández E, Diaz HF, García RR, Prieto MR, Moyano R (2008) A Chronology of El Niño Events from Primary Documentary Sources in Northern Peru. *J Clim* 21:1948–1962.

García-Herrera R, Können GP, Wheeler D, Prieto MR, Jones PD, Koek FB (2006) Ships' logbooks help analyse preindustrial climate. *EoS* 87(18):173-180.

García-Herrera R, Wheeler D, Können GP, Koek FB, Prieto MR (2003) CLIWOC Multilingual Meteorological Dictionary, An English-Spanish-Dutch-French Dictionary of Wind Terms Used by Mariners from 1750-1850, KNMI publication 205.

García-Herrera R, Wilkinson C, Koek FB, Prieto MR, Calvo N, Hernandez E (2005) Description and general background to ships' logbooks as a source of climatic data. *Climatic Change* 73:13-36.

Gergis J, Braganza K, Fowler A, Mooney S, Risbey J (2006) Reconstructing El Niño-Southern Oscillation (ENSO) from high-resolution palaeoarchives. *J Quaternary Sci* 21(7):707-722.

Gergis JL, Fowler AM (2009) A history of ENSO events since A.D. 1525: implications for future climate change. *Climatic Change*. 92:343-387.

Giese BS, Seidel HF, Compo GP, Sardeshmukh PD (2015) An ensemble of historical ocean reanalyses with sparse observational input. *J. Geophys. Res - Oceans*, submitted. (SODAsi.2).

Guilyardi E, Braconnot P, Jin FF, Kim ST, Kolasinski M, Li T, Musat I (2009b) Atmosphere feedbacks during ENSO in a coupled GCM with a modified atmospheric convection scheme. *J Clim* 22:5698–5718.

Guilyardi E, Wittenberg A, Fedorov A, Collins M, Wang C, Capotondi A, van Oldenborgh GJ, Stockdale T (2009a) Understanding El Niño in ocean-atmosphere general circulation models: progress and challenges. *Bulletin of the American Meteorological Society*, 90 (3): 325-340.

Ham YG, Kug JS (2012) How well do current climate models simulate two types of El Niño? *Clim Dyn* 39:383-398.

Hamilton WL (2005) Correlation of wind records and proxy wind history from tree rings at Port Angeles, Washington with sodium concentrations at summit, Greenland

and linkages with Gulf of Alaska sea level pressure forcing. *Polar Geography* 29(4):253-290.

Hannaford M, Jones JM, Bigg GR (2015) Early-nineteenth-century southern African precipitation reconstructions from ships' logbooks. *The Holocene* 25(2):379-390.

Harrison SP, Bartlein PJ, Prentice CI (2016) What have we learnt from paleoclimate simulations? *Journal of Quaternary Science* 31(4):363-385.

Heiss GA (1994) Coral reefs in the Red Sea: Growth, production and stable isotopes Tech Rep 32, GEOMAR.

Hereid KA, Quinn TM, Okumura YM (2013) Assessing spatial variability in El Niño-Southern Oscillation event detection skill using coral geochemistry. *Paleoceanography* 28:14-23.

Hersbach H, Peubey C, Simmons A, Berrisford P, Poli P, Dee DP (2015) ERA-20CM: A twentieth century atmospheric model ensemble. *Quart J Roy Meteor Soc* 141:2350–2375.

Hoerling MP, Kumar A, Xu T (1997) Robustness of the Nonlinear Climate Response to ENSO's Extreme Phases. *J Clim* 14:1277-1293.

Hope P, Henley BJ, Gergis J, Brown J, Ye H (2016) Time varying spectral characteristics of ENSO over the Last Millennium. *Clim Dyn*. DOI: 10.1007/s00382-016-3393-z.

Huang B, Banzon VF, Freeman E, Lawrimore J, Liu W, Peterson TC, Smith TM, Thorne PW, Woodruff SD, Zhang HM (2015) Extended Reconstructed Sea Surface Temperature version 4 (ERSST.v4), Part I. Upgrades and Intercomparison. *J Clim* 28:911-930.

Infanti JM, Kirtman BP (2016). North American rainfall and temperature prediction response to the diversity of ENSO. *Clim Dyn*, 46(9-10):3007–3023.

Ji D, Wang L, Feng J and co-authors (2014) Description and basic evaluation of Beijing Normal University Earth system model (BNU-ESM) version 1. *Geosci Model Dev* 7:2039-3064.

Jiang D, Yu G, Zhao P, Chen X, Liu J, Liu X, Wang S, Zhang Z, Yu Y, Li Y, Jin L, Xu Y, Ju L, Zhou T, Yan X (2015) Paleoclimate modelling in China: A review. *Advances in Atmospheric Sciences* 32(2):250-275.

Jones JM, Widmann M (2003) Instrument- and Tree-Ring-Based Estimated of the Antarctic Oscillation. *J Clim* 16:3511-3523.

Jones PD, Salmon M (2005) Preliminary reconstructions of North Atlantic Oscillation and Southern Oscillation Index measures of wind strength and directions taken during the CLIWOC period. *Climatic Change* 73:131-154.

Jones PD, Briffa KR, Barnett TP, Tett SFB (1998) High-resolution paleoclimatic records for the Last Millennium: Interpretation, integration and comparison with General Circulation model control run temperatures. *Holocene* 8:455- 471.

Jourdain NC, Gupta AS, Taschetto AS, Ummenhofer CC, Moise AF, Ashok K (2013) The Indo-Australian monsoon and its relationship to ENSO and IOD in reanalysis data and the CMIP3/CMIP5 simulations. *Clim Dyn* 41(11-12): 3073-3102.

Jungclaus J. H., and Coauthors (2013) Characteristics of the ocean simulations in MPIOM, the ocean component of the Max Planck Institute Earth System Model. *J. Adv. Model. Earth Syst.*, 5, 422–446, doi:10.1002/jame.20023 Kim ST and Yu J-Y (2012) The two types of ENSO in CMIP5 models. *Geophys Res Letts*, 39: L11704 1:6.

Kageyama M, Braconnot P, Harrison SP, Haywood AM, Jungclaus J, Otto-Blisner BL and co-authors (2016) PMIP4-CMIP6: the contribution of Paleoclimate Modelling Intercomparison Project to CMIP6. *Geoscientific Model Development Discussions*. ISSN 1991-962X.

Kao HY, Yu JY (2009) Contrasting Eastern-Pacific and Central-Pacific Types of ENSO. *J Clim* 22:615-632.

Kaplan A, Cane M, Kushnir Y, Clement A, Blumenthal M, Rajagopalan B (1998) Analyses of global sea surface temperature 1856-1991, *Journal of Geophysical Research*, 103:18567-18589.

Kaplan A, Kushnir Y, Cane MA, Blumenthal MB (1997) Reduced space optimal analysis for historical data sets: 136 years of Atlantic sea surface temperatures. *J Geophys Res* 102(C13):27835-27860.

Karamperidou C, DiNezio PN, Timmermann A, Jin FF, Cobb KM (2015) The response of ENSO flavors to mid-Holocene climate: Implications for proxy interpretation. *Palaeoceanography* 30:527-547.

Khandekar ML, Murty TS, Scott D, Baird W (2000) The 1997 El Niño, Indonesian Forest Fires and the Malaysian Smoke Problem: A deadly combination of natural and man-made hazard. *Natural Hazards* 21:131-144.

Kim ST and Yu J-Y (2012) The two types of ENSO in CMIP5 models. *Geophysical Research Letters* 39:L11704:6.

King AD, Donat MG, Alexander LV, Karoly DJ (2015) The ENSO-Australian rainfall teleconnection in reanalysis and CMIP5. *Clim Dyn* 44(9):2623-2635.

Klein F, Goosse H, Graham NE, Verschuren D (2016) Comparison of simulation and reconstructed variation in East African hydroclimate over the Last Millennium. *Clim Past Discuss* 12:1499-1518.

Knutti R, Abramowitz G, Collins M, Eyring V, Gleckler PJ, Hewitson B, Mearns L (2010) Good Practice Guidance Paper on Assessing and Combining Multi Model Climate Projections. In: Meeting report of the IPCC Expert Meeting on Assessing and Combining Multi Model Climate Projections. [Stocker T, Dahe Q, Plattner G-K, Tignor M, Midgley P (eds)]. IPCC Working Group I Technical Support Unit, University of Bern, Bern, Switzerland.

Koek FB, Können GP (2005) Determination of wind force and present weather terms: The Dutch Case. *Climatic Change* 73:79-95

- Kofeld KE, Graham RM, de Boer AM, Sime LC, Wolff EW, Le Quere C, Bopp L (2013) Southern Hemisphere westerly wind changes during the Last Glacial Maximum paleo-data synthesis. *Quat Sci Rev* 68:76-95.
- Können GP, Jones PD, Kaltofen MP, Allan RH (1998) Pre-1866 extensions of the Southern Oscillation index using early Indonesian and Tahitian meteorological readings. *J Clim* 11:2325-2339.
- Können GP, Koek FB (2005) Description of the CLIWOC database. *Climatic Change* 73(1-2):117-130.
- Krishna Kumar K, Rajagopalan B, Cane MA (1999) On the Weakening Relationship Between the Indian Monsoon and ENSO. *Science* 284(5423):2156-2159.
- Kug JS, Ham YG, Lee JY, Jin FF (2012) Improved simulation of two types of El Niño in CMIP5 models. *Environ Re Lett* 7:034002.
- Kug JS, Jin FF, An SI (2009) Two-types of El Niño events: Cold Tongue El Niño and Warm Pool El Niño. *J Clim* 22:1499-1515.
- Kumar KK, Rajagopalan B, Hoerling M, Bares G, Cane M (2006) Unravelling the mystery of Indian Monsoon failure during El Niño. *Science* 314:115-119.
- Küttel M, Xoplaki E, Gallego D, Luterbacher J, García-Herrera R, Allan R, Barriendos M, Jones P, Wheeler D, Wanner H (2010) The importance of ship log data: reconstructing North Atlantic, European and Mediterranean Sea level pressure fields back to 1750. *Clim Dyn* 34:1115-1128.
- Lachlan-Cope T, Connolley W (2006) Teleconnections between the tropical Pacific and the Amundsen-Bellingshausen Sea: Role of the El Niño/Southern Oscillation. *JGR Atmos* 111:D23101.
- Laing A, Evans JL (2015) Introduction to tropical meteorology, 2nd Edition. URL: http://www.met.ed.ucar.edu/tropical/textbook_2nd_edition/index.htm.
- Lamb HH (1995) *Climate, History and the Modern World* (London: Routledge).
- Larkin NK, Harrison DE (2005) Global seasonal temperature and precipitation anomalies during El Niño autumn and winter. *Geophys Res Lett* 32:L16705.
- Le T (2017) ENSO response to external forcing in CMIP5 simulations of the Last Millennium. *Global and Planetary Change* 148:105-112.
- Lee T, McPhaden MJ (2010) Increasing intensity of El Niño in the central-equatorial Pacific. *Geophys Res Lett* 37:L14603.
- Lewis SC, LeGrande AN (2015) Stability of ENSO and its tropical Pacific teleconnections over the Last Millennium. *Clim fdatPast* 11:1347-1360.
- Li G, Xie SP (2014) Tropical Biases in CMIP5 Multi-model Ensemble: The Excessive Equatorial Pacific Cold Tongue and Double ITCZ Problems. *J Clim* 27:1765-1780.

Li G, Xie SP, Du Y, Luo Y (2016) Effects of excessive equatorial cold tongue bias on the projections of tropical Pacific climate change. Part I: the warming pattern in CMIP5 multi-model ensemble. *Clim Dyn* 47(12):3817-3831.

Li J, Xie, SP, Cook ER, Morales MS, Christie DA, Johnson NC, Chen F, D'Arrigo R, Fowler AM, Gou X, Fang K (2013) El Niño modulation over the past seven centuries. *Nature Climate Change* 3(9):822-826.

Liang YC, Chou CC, Yu JY, Lo MH (2016) Mapping the locations of asymmetric and symmetric discharge responses in global rivers to the two types of El Niño. *Environ Res Lett* 11:L044012.

Linsley BK, Dunbar, RB, Wellington GM, Mucciarone DA (1994) A coral-based reconstruction of Intertropical Convergence Zone variability over Central America since 1707. *J Geophys Res* 99:9977– 9994.

Linsley BK, Ren L, Dunbar RB, Howe S (2000) El Niño southern Oscillation (ENSO) and decadal-scale climate variability at 10°N in the eastern Pacific from 1893-1994: a coral-based reconstruction from Clipperton Atoll. *Paleoceanography* 15:322-335.

Lowe JJ and Walker MJC (2013) *Reconstructing Quaternary Environments*. Routledge: Oxon.

Mann ME, Bradley R, Hughes M (2000b) Long-term variability in the El Niño/Southern Oscillation and associated teleconnections. In: Diaz HAMV, Markgraf VADH (eds) *El Niño and the Southern Oscillation; multi-scale variability and global and regional impacts*. Cambridge University Press, Cambridge, pp 327-372.

Mann ME, Cane MA, Zebiak SE, Clement A (2005) Volcanic and solar forcing of the Tropical Pacific over the past 1000 years. *J Clim* 18(3):447–456.

Mann ME, Gill E, Bradley RS, Hughes MK, Overpeck J, Keimig FT, Gross W (2000a) Global Temperature patterns in the past centuries: An interactive presentation. *Earth Interactions* 4(4) 1-1.

Mann ME, Rutherford S, Wahl E, Ammann C (2005) Testing the Fidelity of Methods Used in Proxy-Based Reconstructions of Past Climate. *J Clim* 18:4097-4107.

Mann ME, Zhang Z, Hughes MK, Bradley RS, Miller SK, Rutherford S, Ni F (2008) Proxy-based reconstructions of hemispheric and global surface temperature variations over the past two millennia. *PNAS* 105(36):13252-13257.

McGregor S, Timmermann A, Timm, O (2010) A unified proxy for ENSO and PDO variability since 1650. *Clim of the Past* 6(1):1-17.

Michael JP, Misra V, Chassignet EP (2013) The El Niño and Southern Oscillation in the historical centennial integrations of the new generation of climate models. *Reg Environ Change* 13:121-130.

Mooney PA, Mulligan FJ, Fealy R (2010) Comparison of ERA-40, ERA-Interim and NCEP/NCAR reanalysis data with observed surface air temperatures over Ireland. *International Journal of Climatology* 31(4):545-557.

Nash DJ, Adamson GCD (2014) Recent advances in the historical climatology of the tropics and subtropics. *Bull Amer Meteor Soc* 95(1):131-146.

National Institute of Water and Atmospheric Research (2017) Southern Oscillation and El Niño. Accessed at <https://www.niwa.co.nz/climate/information-and-resources/el-nino/el-nino-and-southern-oscillation> Accessed on [06/06/2017].

Neukom R, Nash DJ, Endfield GH, Grad SW, Grove CA, Kelso C, Vogel CH, Zinke J (2013) Multi-proxy summer and winter precipitation reconstructions for southern Africa over the last 200 years. *Clim Dyn* 42(9-10):2713-2726.

NOAA (2017a) Schematic diagrams of El Niño, Normal and La Niña conditions. Accessed at <https://www.pmel.noaa.gov/el-nino/schematic-diagrams> Accessed on [06/06/2017].

NOAA (2017b) El Niño climate impacts. Accessed at https://www.pmel.noaa.gov/el-nino/sites/default/files/thumbnails/image/Nino_winter_andsummer_620_from_climate.gov_0.jpg Accessed on [06/06/2017].

NOAA (2017c) La Niña climate impacts. Accessed at https://www.pmel.noaa.gov/el-nino/sites/default/files/thumbnails/image/Nina_winter_andsummer_620_from_climate.gov.jpg Accessed on [06/06/2017].

NOAA (2017d) NINO regions diagram. Accessed at http://www.cpc.noaa.gov/products/analysis_monitoring/ensostuff/nino_regions.shtml Accessed on [06/06/2017].

NOAA (2017e) Niño 3 SST Index. Accessed at http://www.esrl.noaa.gov/psd/gcos_wgsp/Timeseries/Nino3/, Accessed on [10/03/2017].

NOAA (2017f) Niño 4 SST Index Accessed at http://www.esrl.noaa.gov/psd/gcos_wgsp/Timeseries/Nino4/ Accessed on [10/03/2017].

NOAA (2017g) Monthly/Seasonal Climate Composites Accessed at <https://www.esrl.noaa.gov/psd/cgi-bin/data/composites/printpage.pl> Accessed on [02/08/2017].

Oort AH, Yienger JJ (1996) Observed Interannual Variability in the Hadley Circulation and Its Connection to ENSO. *J Clim* 9:2751–2767.

Ordóñez-Pérez P, Gallego D, Ribera P, Peña-Ortiz C, García-Herrera R (2016) Tracking the Indian Summer Monsoon Onset Back to the Pre-Instrumental Period. *J Clim* 29:8115-8127.

Ortlieb L (2000) The documented historical record of El Niño events in Peru: An update of the Quinn record (sixteenth through nineteenth centuries) in Diaz HF and Markgraf V (eds) *El Niño and the Southern Oscillation: Multi-scale variability and global and regional impacts*. Cambridge University Press, Cambridge pp. 207-295.

PAGES 2K consortium (2013) Continental-scale temperature variability during the past two millennia. *Nature Geosci* 6:339-346.

Pascolini-Campbell M, Zanchettin D, Bothe O, Timmreck C, Matei D, Jungclaus JH, Graf HF (2015) Toward a record of Central Pacific El Niño events since 1880. *Theor Appl Climatol* 119:379-389.

Philander SG (1990) *El Niño, La Niña and the Southern Oscillation*. London: Academic Press.

Phipps SJ, McGregor HV, Gergis J, Gallant AJE, Neukom R, Stevenson S, Ackerly D, Brown JR, Fischer MJ, van Ommen TS (2013) Paleoclimate Data-Model Comparisons and the Role of Climate Forcings over the Past 1500 years. *J Clim* 26(18):6915-6936.

Poli P et al. (2013) The data assimilation system and initial performance evaluation of the ECMWF pilot reanalysis of the 20th century assimilating surface observations only (ERA-20C). ERA Report series. Accessed at https://rda.ucar.edu/datasets/ds626.0/docs/ERA-20C.era_report_series_14.pdf. Accessed on [07/08/2017].

Poli P and co-authors (2016) ERA-20C: An Atmospheric reanalysis of the twentieth century. *J Clim* 29:4083-4097.

Quinn TM, Crowley TJ, Taylor FW, Henin C, Jannot P, Join Y (1998) A multicentury stable isotope record from a New Caledonia coral: Interannual and decadal sea surface temperature variability in the southwest Pacific since 1657 A.D. *Paleoceanography* 13:412-426.

Quinn TM, Taylor FW, Crowley TJ (1993) A 173-year stable isotope record from a tropical South Pacific coral. *Quat Sci Rev* 12:407-418.

Quinn, WH, Neal V, Antunez De Mayolo SE (1987) El Niño occurrences over the past four and a half centuries. *J Geophys Res* 92(13):14449-14461.

Raible CC, Brönnimann S, Renate A (2016) Tambora 1815 as a test case for high impact volcanic eruptions: Earth system effects. *WIREs Climate Change* 7:569-589.

Rayner N, Brohan P, Parker D, Folland C, Kennedy J, Vanicek M, Ansell, Tett S (2006) Improved analyses of changes and uncertainties in sea surface temperature measured in situ since the mid-nineteenth century: The HadSST2 data set. *J Clim* 19:446-469.

Rayner NA, Parker DE, Horton EB, Folland CK, Alexander LV, Rowell DP, Kent EC, Kaplan A (2013) Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J Geophys Res* 108:4407(D14).

Ren HL, J FF (2011) Niño indices for two types of ENSO. *Geophys Res Lett* 38(4): L04704.

Ropelewski CF, Halpert MS (1987) Global and Regional Scale Precipitation Patterns Associated with the El Niño/Southern Oscillation. *Mon Weather Rev* 115:1606-1626.

Ropelewski CF, Jones PD (1987) An Extension of the Tahiti–Darwin Southern Oscillation Index. *Mon Weather Rev* 115:2161–2165.

- San Gupta A, Jourdain NC, Brown JN, Monselesan D (2013) Climate Drift in the CMIP5 Models. *J Clim* 26:8597-8615.
- Sandeep S, Stordal F, Sardeshmukh PD, Compo GP (2014) Pacific Walker Circulation variability in coupled and uncoupled climate models. *Clim Dyn* 43:103-117.
- Sato M, Hansen JE, McCormick MP, Pollack JB (1993) Stratospheric aerosol optical depths, 1850-19990. *J. Geophys Res Atmos* 98(D12):22987-22994.
- Schiermeier Q (2015) Hunting the Godzilla El Niño. *Nature* 526:490-491.
- Schmidt GA, Annan JD, Bartlein PJ, Cook BI, Guilyardi E, Hargreaves JC, Harrison SP, Kageyama M, LeGrande AN, Konecky B, Lovejoy S, Mann ME, Masson-Delmotte V, Risi C, Thompson D, Timmermann A, Tremblay LB, Yiou P (2014) Using paleo-climate comparisons to constrain future projections in CMIP5. *Climate of the Past* 9:775-835.
- Schmidt GA, Jungclauss JH, Ammann CM, Bard E and co-authors (2012) Climate forcing reconstructions for use in PMIP simulations of the Last Millennium (v1.1). *Geosci Model Dev* 5:185–191.
- Schmidt GA and co-authors (2011) Climate forcing reconstructions for use in PMIP simulations of the Last Millennium (v1.0). *Geosci Model Dev* 4:33–45.
- Schollaen K, Karamperidou C, Krusic P, Cook E, Helle G (2015) ENSO flavors in a tree-ring $\delta^{18}\text{O}$ record of *Tectona grandis* from Indonesia. *Clim Past* 11:1325-1333.
- Singh A (2012) Contrasting the flavors of ENSO and related trends in the tropical Pacific Ocean in recent decades. *Oceanography*. Universite Paul Sabatier - Toulouse III. English. Accessed at <https://tel.archives-ouvertes.fr/tel-00795552>. Accessed on [08/06/2017].
- Smerdon JE, Luterbacher J, Phipps SJ, and co-authors (2017) Comparing proxy and model estimates of hydroclimate variability and change over the Common Era. *Clim Past* <https://doi.org/10.5194/cp-2017-37>.
- Smith T, Reynolds R, Peterson TC, Lawrimore J (2008) Improvements to NOAA's historical merged land-ocean surface temperature analysis (1880-2006) *J Clim* 21:2283-2296.
- Song L, Chen S, Chen W, Chen X (2016) Distinct impacts of two types of La Niña events on Australian summer rainfall. *Int J Clim* 37(5):2532-2544.
- Sperber KR, Annamalai H, Kang IS, Kitoh A, Moise A, Turner AG, Wang B, Zhou T (2013) The Asian summer monsoon: An Intercomparison of CMIP5 versus CMIP3 simulations of the late 20th century. *Clim Dyn* 41:2711–2744.
- Stahle DW, D'Arrigo RD, Krusic PJ, Cleaveland MK, Cook ER, Allan RJ, Cole JE, Dunbar RB, Therrell MD, Gay DA, Moore MD, Stokes MA, Burns BT, Villanueva-Diaz J, Thompson LG (1998) Experimental dendroclimatic reconstruction of the Southern Oscillation. *Bull Am Meteorol Soc* 79:2137-2152.

Stevenson S, Capotondi A, Fasullo J, Otto-Bliesner B (2017) Forced changes to twentieth century ENSO diversity in a Last Millennium context. *Clim Dyn* doi:10.1007/s00382-017-3573-5.

Sueyoshi T, Ohgaito R, Yamamoto A, Chikamoto MO, Hajima T, Okajima H, Yoshimori M, Abe M, O'Ishi R, Saito F, Watanabe S, Kawamiya M, Abe-Ouchi A (2013) Set-up of the PMIP3 paleoclimate experiments conducted using an earth system model, MIROC-ESM. *Geosci Model Dev* 6(3):819–836

Taylor KE, Stouffer RJ, Meehl GA (2012) An Overview of CMIP5 and the Experiment Design. *Bull Amer Meteor Soc* 93:485–498.

Taylor MH, Losch M, Wenzel M, Schröter J (2013) On the Sensitivity of Field Reconstruction and Prediction Using Empirical Orthogonal Functions Derived from Gappy Data. *J Clim* 26:9194–9205.

Tejavath CT, Ashok K, Chakraborty S, Ramesh R (2017) The Indian summer monsoon climate during the Last Millennium as simulated by the PMIP3. *Clim Past Discuss* doi:10.5194/cp-2017-24.

Thompson LG (1993) Reconstructing the paleo ENSO records from tropical and subtropical ice cores. *Bull Inst fr etudes andines* 22(1):65-83.

Toh, YY, Turner AG, Johnson SJ and co-authors (2017) Maritime Continent seasonal climate biases in AMIP experiments of the CMIP5 multimodel ensemble *Clim Dyn* doi:10.1007/s00382-017-3641-x.

Tudhope AW, Shimmield GB, Chilcott CP, Jebb M, Fallick AW, Dalglish AN (1995) Recent changes in climate in the far western equatorial Pacific and their relationship to the Southern Oscillation: Oxygen isotope records from massive corals, Papua New Guinea. *Earth Planet Sci Lett* 136:575-590.

Turner AG, Innes PM, Slingo JM (2005) The role of the basic state in the ENSO-monsoon relationship and implications for predictability. *Q J R Meteorolo Soc* 131:781-804.

Ubilava D (2011) El Niño, La Niña, and world coffee price dynamics. *Agricultural economics* 43(1):17-26.

Vierira L, Solanki S (2009) Evolution of the solar magnetic flux on time scales of years to millennia. *Astron Astrophys*. DOI:10.1051/0004-6361/200913276.

Walker GT, Bliss EW (1932) World weather V. *Mem Roy Meteor Soc* 4:53-84.

Wang C, Deser C, Yu JY, DiNezio P, Clement A (2012) El Niño southern Oscillation (ENSO): A Review. In Glynn P, Manzello D, Enochs I (Eds.), *Coral Reefs of the Eastern Pacific*. Springer-Verlag.

Wang C, Fiedler PC (2006) ENSO variability and the eastern tropical Pacific: A review. *Progr Oceanogr* 69:239-266.

Wang G, Hendon HH (2007) Sensitivity of Australian rainfall to inter-El Niño variations. *J. Clim* 20:4211-4226.

- Wang H, Kumar A, Wang W, Xue Y (2012) Influence of ENSO on Pacific Decadal Variability: An Analysis Based on the NCEP Climate Forecast System. *J Clim.* 25:6136-6151.
- Wang YM, Lean JL, Sheeley NR Jr (2005) Modelling the Sun's magnetic field and irradiance since 1713. *Astrophys J* 625(1):522-638.
- Watanabe S, Hajima T, Sudo K, Nagashima T, Takemura T, Okajima H, Nozawa T, Kawase H, Abe M, Yokohata T, Ise T, Sato H, Kato E, Takata K, Emori S, Kawamiya M (2011) MIROC-ESM 2010: model description and basic results of CMIP5-20c3m experiments. *Geosci Model Dev* 4:845-872.
- Weng H, Wu, G, Liu Y, Behera SK, Yamagata T (2011) Anomalous summer climate in China influenced by the tropical Indo-Pacific Oceans. *Clim Dyn* 36:769-782.
- Wheeler D (2001) The weather of the European Atlantic seaboard during October 1805: An exercise in climatology. *Climatic Change* 48:316-385.
- Wheeler D (2005) An examination of the accuracy and consistency of ships' logbook weather observations and records. *Climatic Change* 73: 97-116.
- Wheeler D (2014) Hubert Lamb's treasure trove – ships logbooks in climate research. *Weather* 69:133-139.
- Wheeler D, García-Herrera R, Wilkinson CW, Ward C (2010) Atmospheric circulation and storminess derived from royal navy log books; 1685 to 1750. *Climatic change.* 101:257-280.
- Wheeler D, Wilkinson C (2005) The Determination of Logbook Wind Force and Weather Terms: The English Case. *Climatic Change* 73:57-77.
- Whetton PH, Allan RJ, Rutherford I (1996) Historical ENSO teleconnections in the eastern hemisphere: Comparison with latest El Niño series of Quinn. *Clim Change* 32:103-109.
- Wiedermann M, Radebach A, Donges JF, Kurths J, Donner RV (2016) A climate network-based index to discriminate different types of El Niño and La Niña. *Geophys Res Lett* 43:7176.
- Wilkinson C (2009) British Logbooks in UK Archives, 17th-19th Centuries – a survey of the range, selection and suitability of British logbooks and related documents for climatic research. CRU RP12.
- Wilkinson C, Woodruff SD, Brohan P, Claesson S, Freeman E, Koek F, Lubker SJ, Marzin C, Wheeler D (2011) Recovery of logbooks and international marine data: the RECLAIM project. *Int J Climatol* 31:968–979.
- Wilks DS (2005) *Statistical methods in the atmospheric sciences*. Elsevier, London.
- Wilks DS (2011) *Statistical Methods in Atmospheric Sciences: An Introduction*. 3rd Edition. Academic Press, San Diego, CA.

Wilson R, Cook E, D'Arrigo R, Riedwyl N, Evans MN, Tudhope A, Allan R (2010) Reconstructing ENSO: the influence of method, proxy data, climate forcing and teleconnections. *J Quaternary Sci* 25(1):62-78.

Wittenberg AT (2009) Are historical records sufficient to constrain ENSO simulations? *Geophysical Res letters* 36:L12702.

Wolff C, Haug GH, Timmermann A, Damste' JSS, Brauer A, Sigman DM, Cane MA, Verschuren D (2011) Reduced interannual rainfall variability in East Africa during the last ice age. *Science* 333:743–747.

Woodruff SD and co-authors (2011) ICOADS release 2.5: Extensions and enhancements to the surface marine meteorological archive. *Int J Climatol* 31:951–967.

Woodruff SD, Diaz HF, Worley SJ, Reynolds RW, Lubker SJ (2005) Early ship observational data and ICOADS. *Climatic Change* 73:169-194.

Xu K, Tam CY, Zhu C, Liu B, Wang W (2017) CMIP5 Projections of Two Types of El Niño and their related tropical precipitation in the twenty-first century. *J Clim* 30: 849-864.

Yeh SW, Kug JS, An SI (2014) Recent progress on two types of El Niño: Observations, dynamics, and future changes. *Asia-Pacific J Atmos Sci* 50(1):69-81.

Yeh SW, Kug JS, Dewitte B, Kirtman B, Jin FF (2009) Recent changes in El Niño and its projection under global warming. *Nature* 461:511-515.

Yu JY, Kim ST (2013) Identifying the types of major El Niño events since 1870. *Int J Climatol* 33:2105-2112.

Yu JY, Zou Y, Kim ST, Lee T (2012) The changing impact of El Niño on US winter temperatures. *Geophy Res Letters* 39:L15702.

Zhang T, Perlwitz J, Hoerling MP (2014) What is responsible for the strong observed asymmetry in teleconnections between El Niño and La Niña? *Geophy Res Letters* 41(3):1091-1025.

Zhou T, Wu B, Dong L (2014) Review: Advances in Research of ENSO Changes and the Associated Impacts on Asian-Pacific Climate. *Asia-Pac J Atmos Sci* 50(4):405-422.

Zinke J, Dullo WC, Heiss GA, Eisenhauer A (2004) ENSO and Indian Ocean subtropical dipole variability is recorded in a coral record off southwest Madagascar for the period 1659 to 1995. *Earth Planet Sci Lett* 228:177-194.

Appendix A

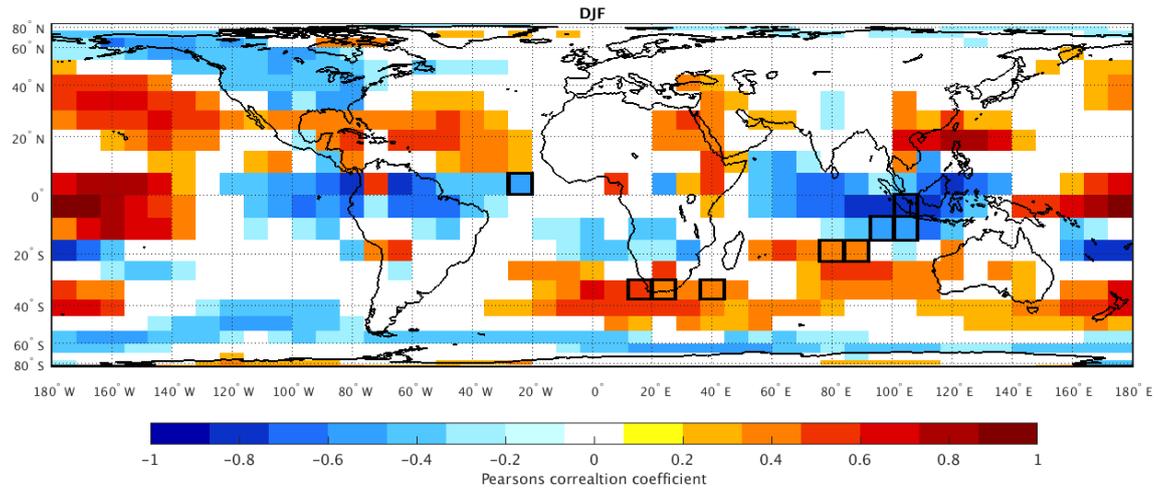


Fig. A1 Pearson's correlation coefficient between de-trended mean local noon DJF zonal wind and the DJF Niño 3.4 index, 1979/80-2013/14. Only those significant at the 90% confidence level are shown. Grid boxes used in the reconstructions are outlined and are numbered 1 to 9 from West to East.

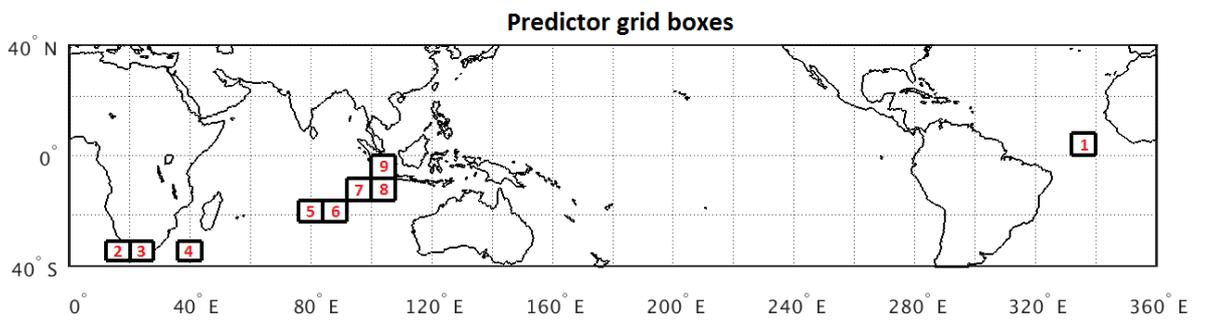


Fig. A2 Numbered labels for the nine predictor grid boxes used in the logbook-based reconstruction.

Table A1 Location and labels of the nine predictor grid boxes used in the logbook-based reconstructions.

Grid box Number	Latitude limits	Longitude limits
1	0°N - 7.5°N	20°W - 24°W
2	30°S - 37.5°S	12°E - 20°E
3	30°S - 37.5°S	20°E - 28°E
4	30°S - 37.5°S	36°E - 44°E
5	15°S - 22.5°S	76°E - 84°E
6	15°S - 22.5°S	84°E - 92°E
7	7.5°S - 15°S	92°E - 100°E
8	7.5°S - 15°S	100°E - 108°E
9	0°N - 7.5°S	100°E - 108°E

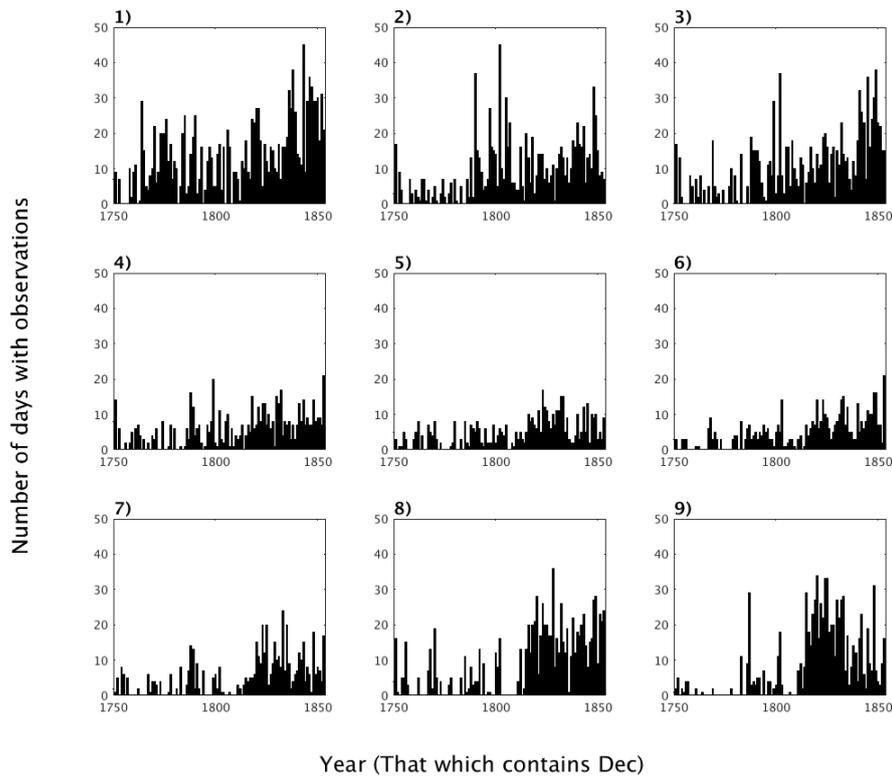


Fig. A3 Number of days with both wind force and direction observations from logbooks per DJF season in the nine predictor grid boxes, 1750/51-1853/54.

Table A2 The DJF SOI reconstructions for 1788/89, 1791/92, 1802/03 and 1811/12, using PCR and CPS. Climatology taken from 1834-1853.

	DJF SOI PCR Reconstruction	DJF SOI CPS Reconstruction
1788/89	-4.6	0.90
1791/92	-26.36	-0.95
1802/03	-12.84	-0.10
1811/12	3.47	0.01

Table A3 Linear regression equations for the four historical reconstructions and their R^2 values, statistically significant ($p < 0.05$) trends are shown in bold.

	PCR A	PCR B	CPS A	CPS B
Linear fit line	$y = -0.039x + 71$	$y = -0.0009x + 0.0001$	$y = 0.05x + 93$	$y = -0.018x + 33$
R^2	0.20 P < 0.01	0.0001 P = 0.95	0.34 P < 0.01	0.04 P = 0.21

Table A4 Mean normalised SOI and two-sampled t-test results for the two periods, 1815-1833 (Time period I) and 1834-1853 (Time period II) for the four SOI reconstructions.

	PCR A		PCR B		CPS A		CPS B	
	I	II	I	II	I	II	I	II
Mean SOI	0.50	-0.47	-0.02	0.01	0.48	-0.45	0.01	-0.01
Two-sampled t-test	t(35) = 3.36, p = 0.0019		t(35) = -0.09, p = 0.93		t(35) = 3.16, p = 0.0033		t(35) = 0.07, p = 0.95	

Table A5 Pearson’s correlation coefficient between logbook DJF SOI reconstructions and the two early instrumental Jakarta based SOI records. Those significant at the 95% confidence level are indicated in bold.

	Jakarta rain-day SOI (June-Nov) (1830-1850)	Jakarta single site SOI (DJF average) (1841/42-1853/54)
PCR A	0.17	0.36
PCR B	0.20	0.36
CPS A	0.56	0.45
CPS B	0.55	0.45

Table A6 The logbook-based DJF SOI reconstructions, 1815/16 – 1853/54. NaN indicated years with missing data.

	PCR A	PCR B	CPS A	CPS B
1815	1.12	-7.62	-0.14	-0.58
1816	6.88	-1.87	0.88	0.38
1817	5.39	-3.36	-0.01	-0.64
1818	18.99	10.24	1.85	1.62
1819	18.78	10.04	1.49	1.21
1820	3.21	-5.53	0.56	0.29
1821	6.86	-1.88	-0.01	-0.59
1822	7.74	-1.00	1.46	1.16
1823	3.56	-5.18	-0.12	-0.75
1824	-9.31	-18.05	-0.45	-1.11
1825	19.90	11.15	0.82	0.34
1826	17.97	9.23	0.99	0.61
1827	18.57	9.83	0.82	0.43
1828	19.87	11.13	0.88	0.74
1829	NaN	NaN	NaN	NaN
1830	-7.04	-15.78	-0.54	-1.35
1831	27.86	19.12	0.47	0.32
1832	5.60	-3.14	0.41	-0.04
1833	-0.37	-9.11	-0.75	-1.82
1834	1.72	10.24	0.60	1.23
1835	-31.84	-23.33	-1.02	-0.60
1836	45.55	54.06	2.13	2.80
1837	-15.12	-6.61	1.02	1.40
1838	-14.50	-5.99	-1.80	-1.37
1839	-9.65	-1.14	0.04	0.48
1840	-4.40	4.11	-0.35	0.08
1841	-5.66	2.85	-0.29	0.13
1842	-28.19	-19.68	-1.15	-0.73
1843	-8.90	-0.38	-0.22	0.15
1844	NaN	NaN	NaN	NaN
1845	-8.73	-0.21	-1.15	-0.75
1846	4.01	12.52	0.19	0.74
1847	-26.34	-17.82	-1.24	-0.80
1848	9.57	18.09	0.22	0.69
1849	-5.38	3.14	-0.83	-0.48
1850	-3.57	4.94	-1.16	-0.78
1851	-43.31	-34.80	-2.16	-1.84
1852	-2.72	5.79	-0.94	-0.49
1853	3.12	11.63	-0.48	-0.06

Appendix B

Table B1 Two sampled t-test comparing the DJF SOI from the instrumental record to that from the three reanalysis products, the DJF SOI is relative to the 1979-2004 mean. Comparisons made over four periods; 1979-2004, 1900-1949, 1950-1999 and 1900-1999.

	1979-2004	1900-1949	1950-1999	1900-1999
ERA-INT	H=0	---	----	---
ERA-20C	H=0	H=1	H=0	H=1
20CR	H=0	H=1	H=0	H=1

*H=1 indicates that the test rejects the null hypothesis at the 95% confidence level

Table B2 The years of El Niño and La Niña events defined from by one standard deviation of the DJF SOI from the instrumental record and the two historical reanalysis, 1900-1999. Bold indicates events found in the instrumental record and the selected reanalysis.

	El Niño		La Niña			
	Instrumental	ERA-20C	20CR	Instrumental	ERA-20C	20CR
1905		1902	1902	1903	1916	1909
1911		1905	1904	1916	1917	1916
1914		1911	1905	1917	1942	1938
1918		1914	1911	1920	1950	1950
1925		1925	1913	1928	1954	1954
1940		1930	1914	1938	1955	1955
1941		1940	1918	1942	1961	1961
1957		1941	1925	1949	1970	1966
1958		1957	1930	1950	1973	1970
1972		1958	1940	1955	1975	1973
1977		1968	1941	1961	1988	1975
1982		1972	1982	1970	1998	1988
1986		1982	1991	1973	1999	1996
1991		1991	1997	1975		1998
1997		1997		1988		1999
				1998		
				1999		
Total:	15	15	14	17	13	15

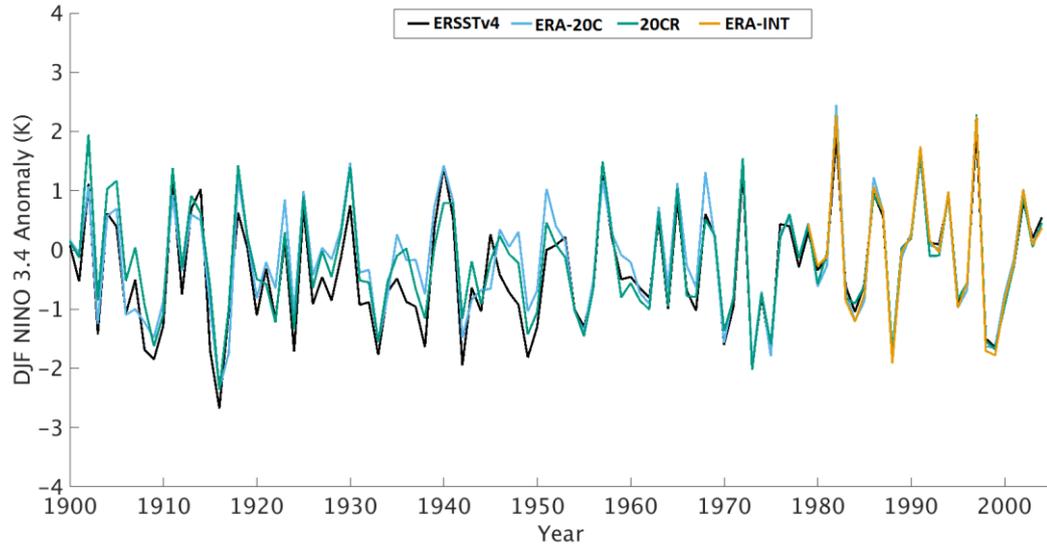


Fig. B1 The DJF Niño 3.4 anomaly from the three reanalyses (ERA-20C, 20CR and ERA-Interim) and ERSSTv4, 1900-2004. Anomaly relative to the 1979-2004 base period.

Table B3 Two sampled t-test comparing the DJF Niño 3.4 index from the instrumental record to that from the three reanalysis products, the DJF Niño 3.4 index is relative to the 1979-2004 mean. Comparisons made over four periods; 1979-2004, 1900-1949, 1950-1999 and 1900-1999.

	1979-2004	1900-1949	1950-1999	1900-1999
ERA-INT	H=0	---	----	---
ERA-20C	H=0	H=0	H=0	H=0
20CR	H=0	H=0	H=0	H=0

*H=1 indicates that the test rejects the null hypothesis at the 95% confidence level

Table B4 The years of El Niño and La Niña events defined from by one standard deviation of the DJF Niño 3.4 index from the instrumental record (ERSST) and the two historical reanalysis, 1900-1999. Bold indicates events found in the instrumental record and the selected reanalysis.

El Niño			La Niña		
ERSST	ERA-20C	20CR	ERSST	ERA-20C	20CR
1902	1902	1902	1903	1903	1909
1904	1911	1904	1908	1908	1916
1911	1918	1905	1909	1909	1922
1913	1923	1911	1915	1916	1924
1914	1925	1913	1916	1917	1933
1918	1930	1918	1924	1933	1938
1925	1940	1925	1933	1942	1942
1930	1941	1930	1938	1955	1949
1940	1951	1940	1942	1970	1955
1957	1957	1941	1949	1973	1970
1965	1965	1957	1970	1975	1973
1968	1968	1965	1973	1984	1975
1972	1972	1972	1975	1988	1988
1982	1982	1982	1988	1998	1998
1986	1986	1986	1998	1999	1999
1991	1991	1991	1999		
1994	1994	1994			
1997	1997	1997			
Total: 18	18	18	16	15	15

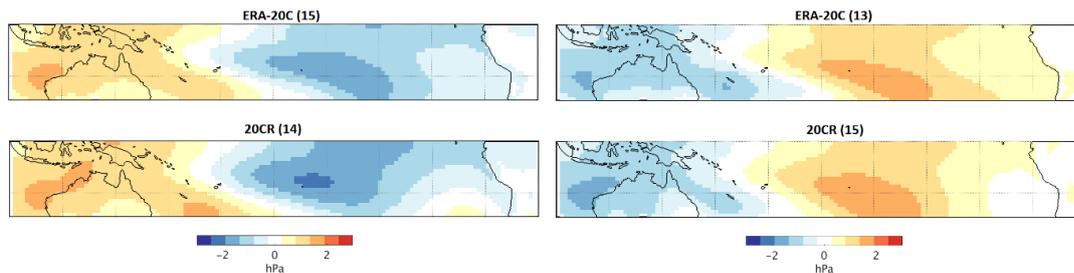


Fig. B2 DJF MSLP (hPa) composite anomaly for El Niño events (left) and La Niña events (right) from the two historical reanalysis using the SOI one standard deviation threshold, 1900-1999. Numbers in brackets indicate the number of events used to calculate each composite plot.

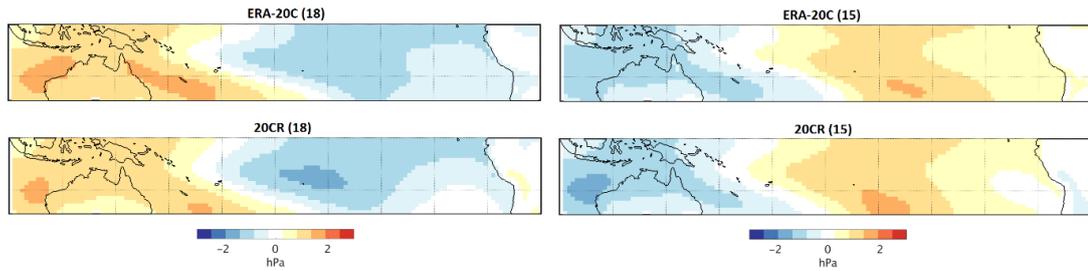


Fig. B3 DJF MSLP (hPa) composites anomaly for El Niño events (left) and La Niña events (right) from the two historical reanalysis using the Niño 3.4 index one standard deviation threshold, 1900-1999. Numbers in brackets indicate the number of events used to calculate each composite plot.

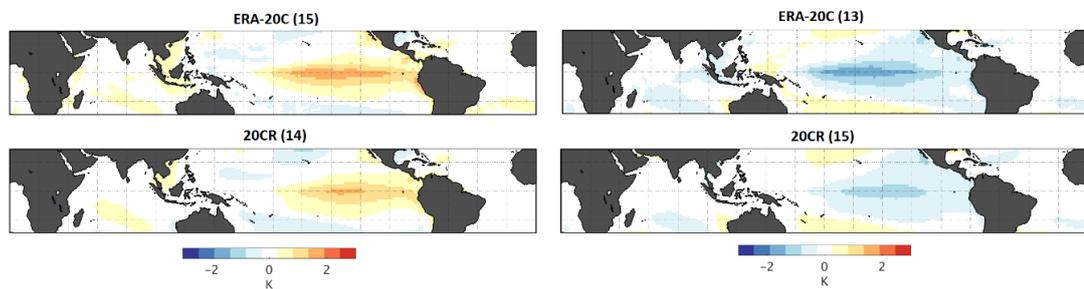


Fig. B4 Surface temperature (K) composite anomaly for El Niño events (left) and La Niña events (right) from the two historical reanalysis using SOI one standard deviation, 1900-1999. Numbers in brackets indicate the number of events used to calculate each composite plot.

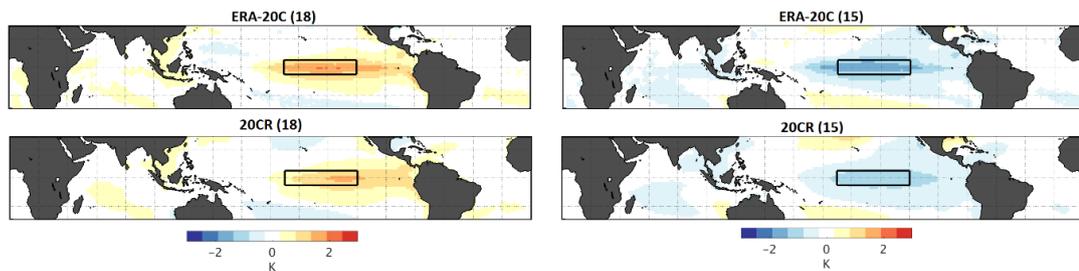


Fig. B5 Surface temperature (K) composite anomaly for El Niño events (left) and La Niña events (right) from the two historical reanalysis using the Niño 3.4 index one standard deviation, 1900-1999. Numbers in brackets indicate the number of events used to calculate each composite plot.

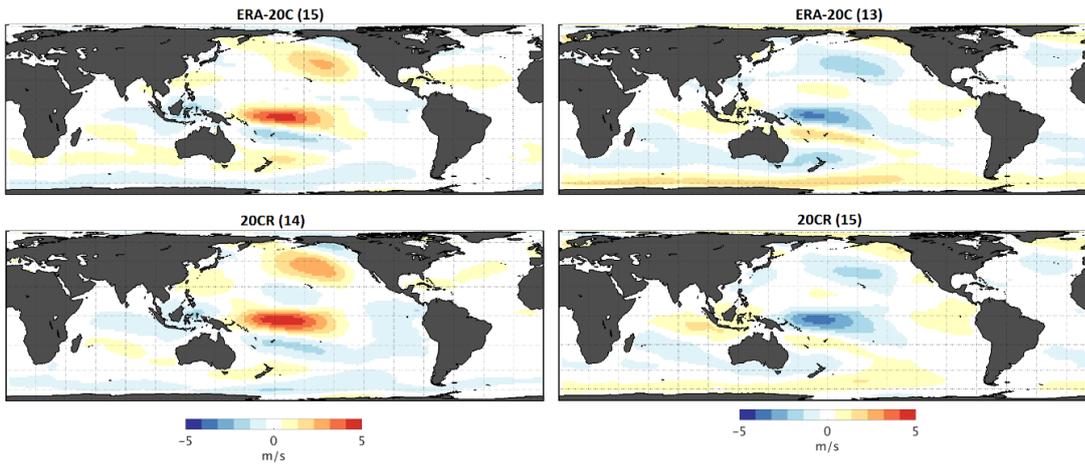


Fig. B6 Zonal wind (ms^{-1}) composite anomaly for El Niño events (left) and La Niña events (right) from the two historical reanalysis using the SOI one standard deviation threshold, 1900-1999. Numbers in brackets indicate the number of events used to calculate each composite plot.

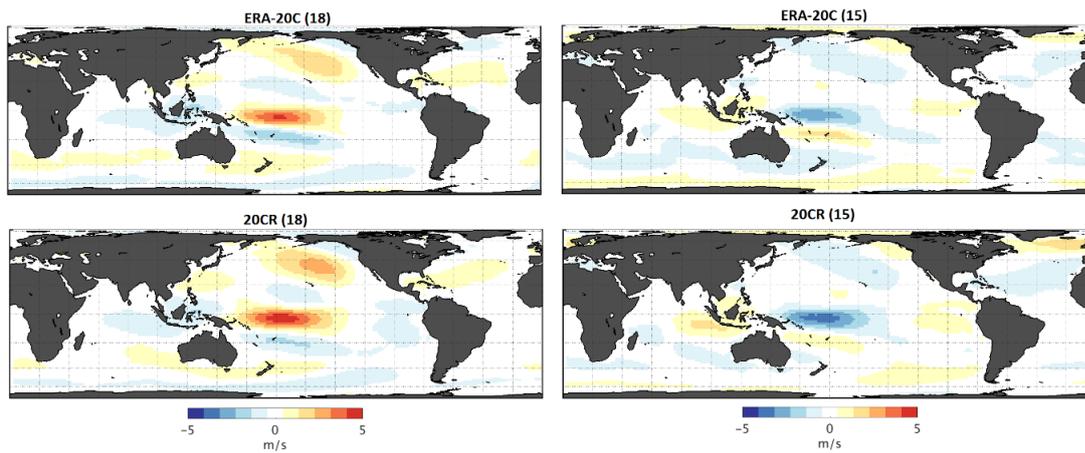


Fig. B7 Zonal wind (ms^{-1}) composite anomaly for El Niño events (left) and La Niña events (right) from the two historical reanalysis using the Niño 3.4 index one standard deviation threshold, 1900-1999. Numbers in brackets indicate the number of events used to calculate each composite plot.

Appendix C

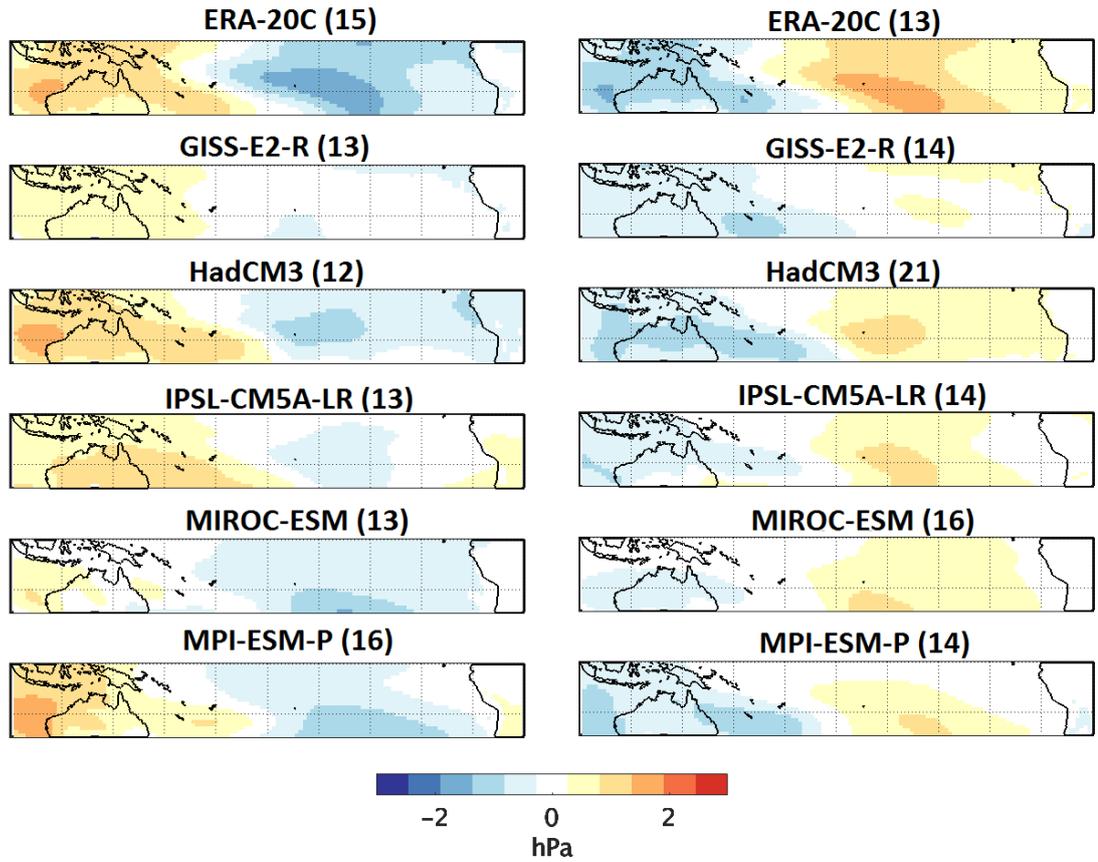


Fig. C1 DJF MSLP (hPa) composite anomaly from ERA-20C reanalysis and climate model simulations during El Niño (left) and La Niña events (right) defined using the SOI one standard deviation, 1900-1999. Anomalies based on 1900-1999 DJF mean. Numbers in brackets indicate the number of events used to calculate each composite plot.

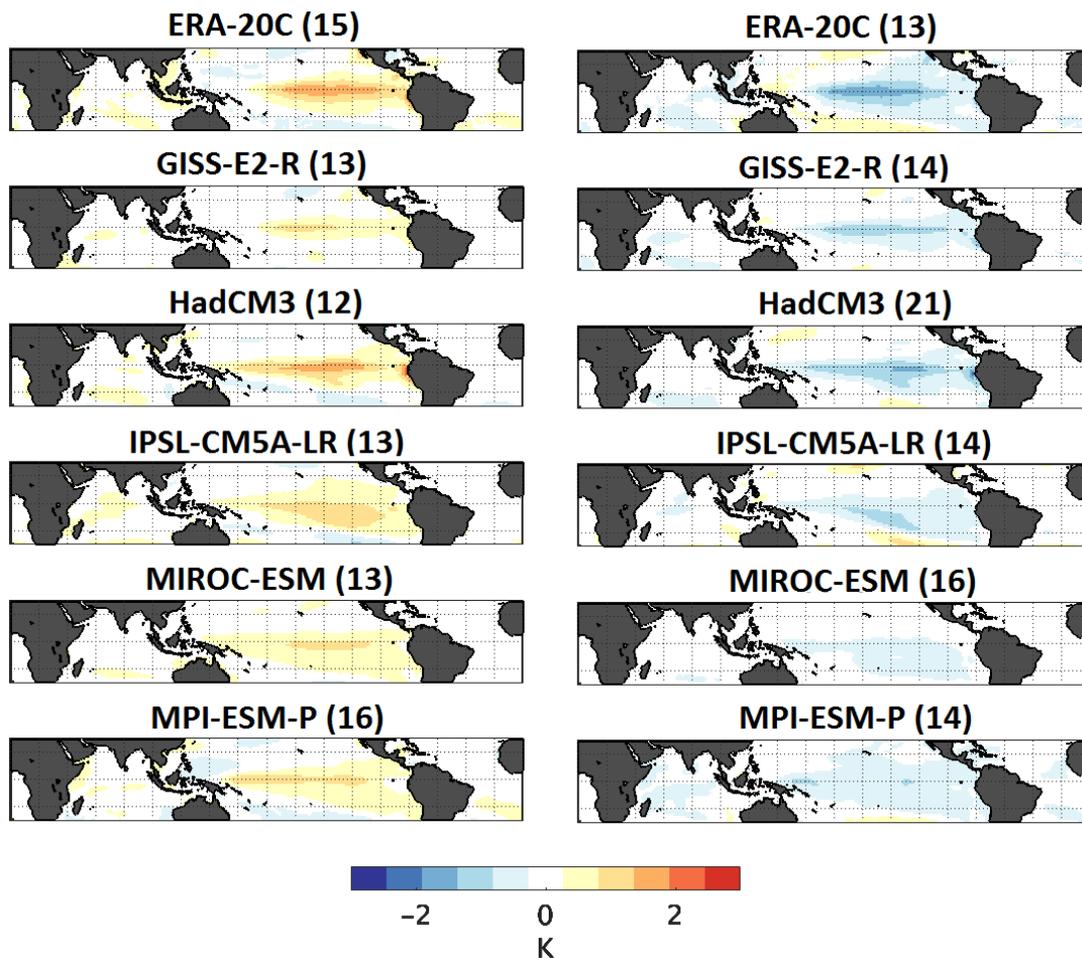


Fig. C2 Mean DJF surface temperature (K) composite anomaly from ERA-20C reanalysis and climate model simulations during El Niño (left) and La Niña events (right) defined using the SOI one standard deviation, 1900-1999. Anomalies based on 1900-1999 DJF mean. Numbers in brackets indicate the number of events used to calculate each composite plot.

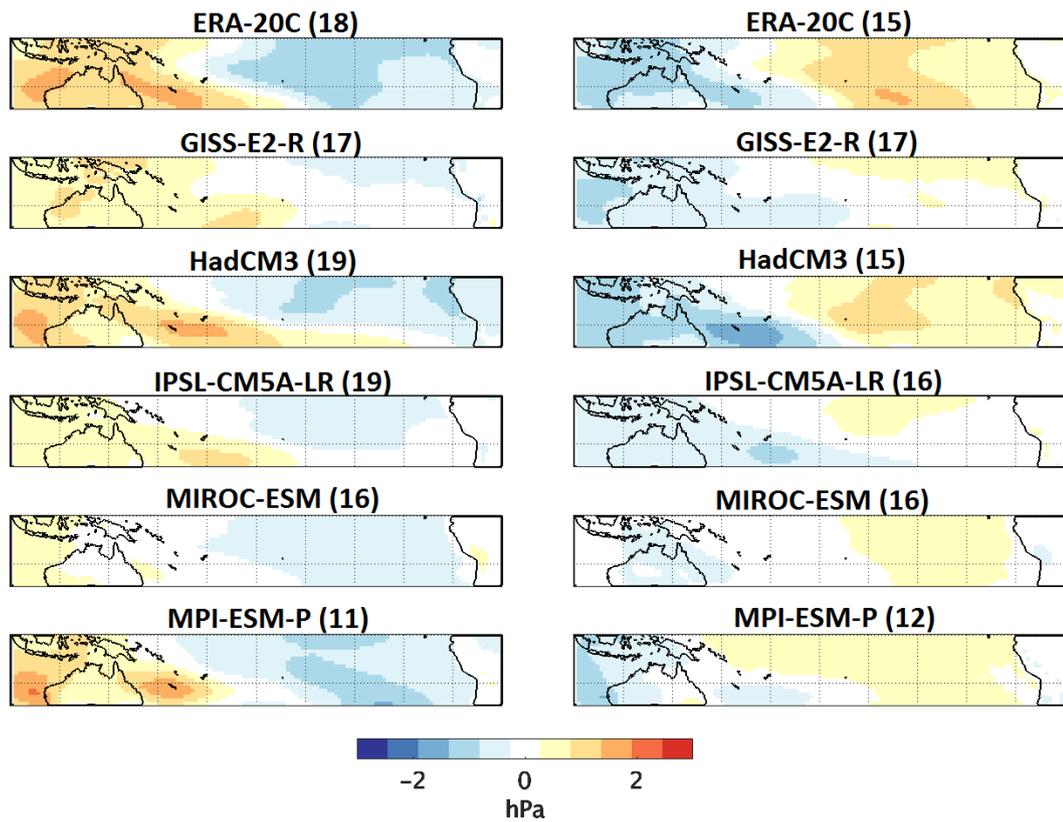


Fig. C3 Mean DJF MSLP (hPa) composite anomaly from ERA-20C reanalysis and climate model simulations during El Niño (left) and La Niña (right) years defined using the Niño 3.4 one standard deviation, 1900-1999. Anomalies based on 1900-1999 DJF mean. Numbers in brackets indicate the number of events used to calculate each composite plot.

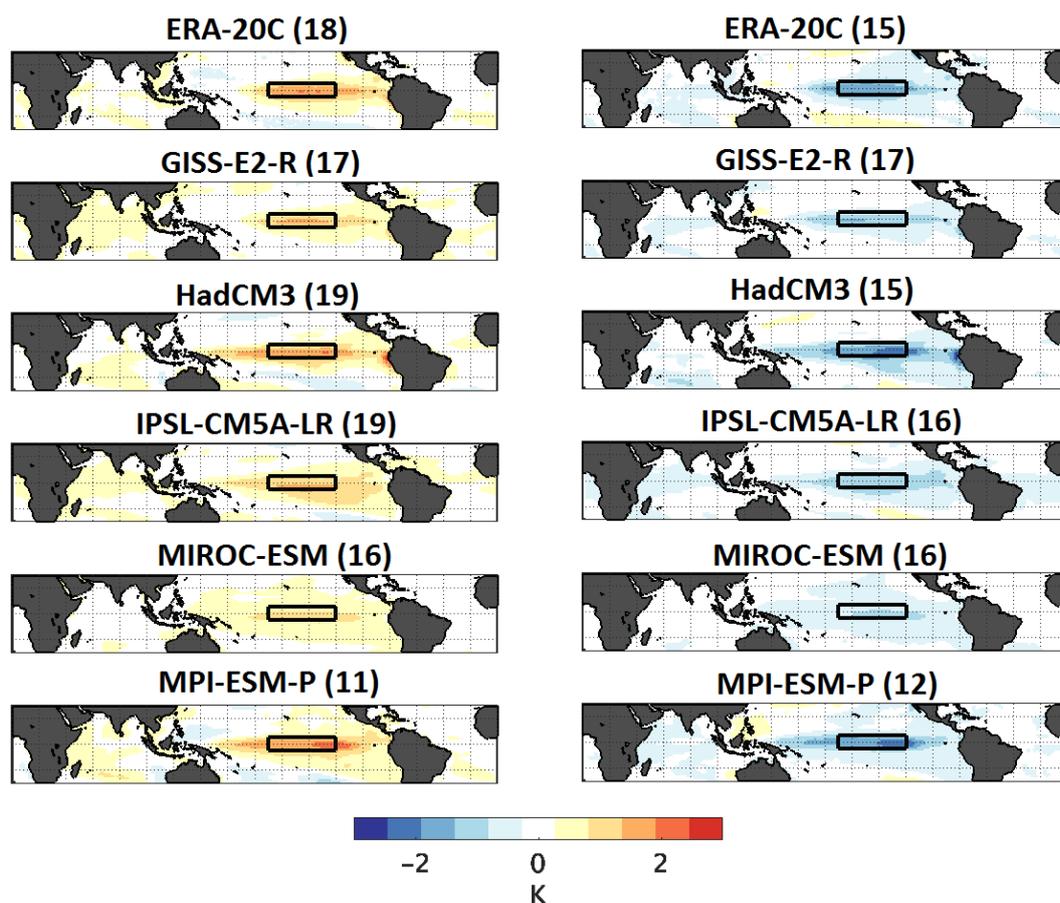


Fig. C4 Mean DJF surface temperature (K) anomaly from ERA-20C reanalysis and climate model simulations during El Niño (left) and La Niña events (right) defined using the SOI one standard deviation, 1900-1999. Anomalies based on 1900-1999 DJF mean. Numbers in brackets indicate the number of events used to calculate each composite plot.

Table C1 A two-sampled t-test comparing the DJF SOI from the instrumental record (BoM) to that from the climate model simulations, 1979-2004 and 1900-1999. Bold indicates that the null hypothesis cannot be rejected at the 95% confidence level.

	1979-2004	1900-1999
GISS-E2-R	H= 0, p=1.0	H=0, p=0.14
HadCM3	H= 0, p=1.0	H=1, p=1.67x10 ⁻⁴
IPSL-CM5A-LR	H= 0, p=1.0	H=1, p=0.0016
MIROC-ESM	H= 0, p=1.0	H=0, p=0.167
MPI-ESM-P	H= 0, p=1.0	H=1, p=0.0065

*H=1 indicates that the test rejects the null hypothesis at the 95% confidence level

Table C2 Ranking of historical model simulations for SOI-based metrics compared to the DJF SOI instrumental record (BoM) (1979-2004).

	Rank for MSLP Gradient	Rank for SOI standard deviation	Rank for El Niño frequency	Rank for La Niña frequency	Overall Ranking
GISS-E2-R	2	2	=2	=3	2
HadCM3	3	4	5	1	4
IPSL-CM5A-LR	1	1	=2	=3	1
MIROC-ESM	5	3	1	2	3
MPI-ESM-P	4	5	=2	5	5

Table C3 Ranking of historical model simulations for SOI-based metrics compared to the DJF SOI from ERA-20C (1900-1999).

	Rank for MSLP Gradient	Rank for SOI standard deviation	Rank for El Niño frequency	Rank for La Niña frequency	Overall Ranking
GISS-E2-R	=3	2	=2	=1	=1
HadCM3	2	1	5	5	4
IPSL-CM5A-LR	1	4	=2	=1	=1
MIROC-ESM	5	3	=2	4	5
MPI-ESM-P	=3	5	1	=1	3

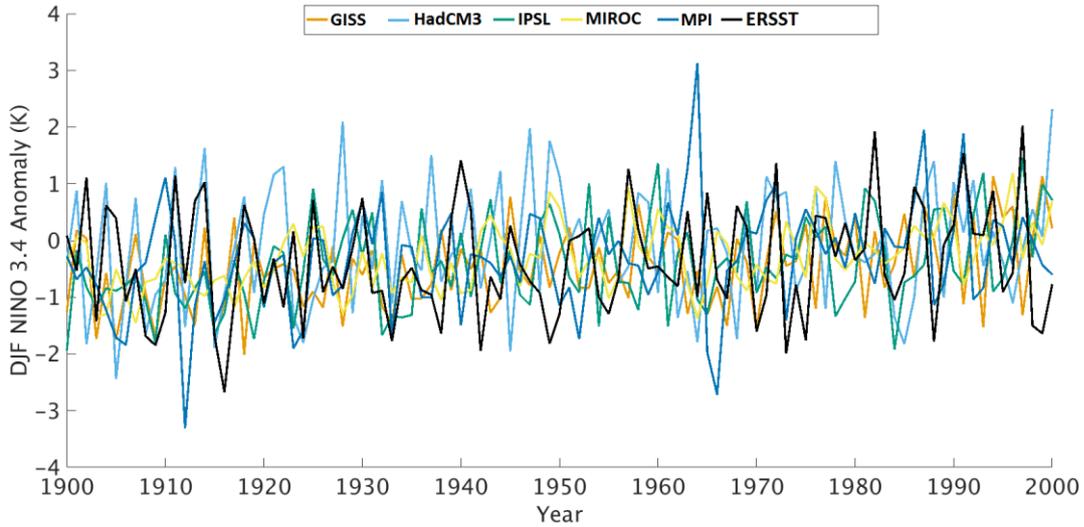


Fig. C5 The DJF Niño 3.4 Anomaly (K) (1979-2004 base period) from historical climate model simulations and ERSST, 1900-2004.

Table C4 Results from t-test comparing the Niño 3.4 anomaly indices from ERSST to the climate model simulations, 1979-2004 and 1900-1999. Bold indicates that the null hypothesis cannot be rejected at the 95% confidence level.

	1979-2004	1900-1999
GISS-E2-R	H= 0, p=1.0	H=0, p=0.34
HadCM3	H= 0, p=1.0	H=1, p=0.05
IPSL-CM5A-LR	H= 0, p=1.0	H=0, p=0.66
MIROC-ESM	H= 0, p=1.0	H=0, p=0.55
MPI-ESM-P	H= 0, p=1.0	H=0, p=0.63

*H=1 indicates that the test rejects the null hypothesis at the 95% confidence level

Table C5 Ranking of historical model simulations for Niño 3.4 metrics compared to ERRST (1979-2004).

	Rank for Niño 3.4 mean	Rank for Niño 3.4 standard deviation	Rank for El Niño frequency	Rank for La Niña frequency	Overall Ranking
GISS-E2-R	4	=2	5	=2	5
HadCM3	2	1	=3	5	=3
IPSL-CM5A-LR	1	4	=3	=2	2
MIROC-ESM	3	5	=1	=2	=3
MPI-ESM-P	5	=2	=1	1	1

Appendix D

Table D1 Two sample t-test to compare DJF zonal mean climatologies from ERA-Interim and ERA-20C/20CR for each grid box (1-9), 1979-2004. An H value of 0 indicates that the null hypothesis cannot be rejected at the 95% level.

	Predictor grid box									Number of H=0
	1	2	3	4	5	6	7	8	9	
ERA-20C	H=1	H=1	H=1	H=0	H=1	H=0	H=0	H=0	H=0	5
20CR	H=1	H=0	H=1	H=1	H=0	H=1	H=0	H=0	H=0	5

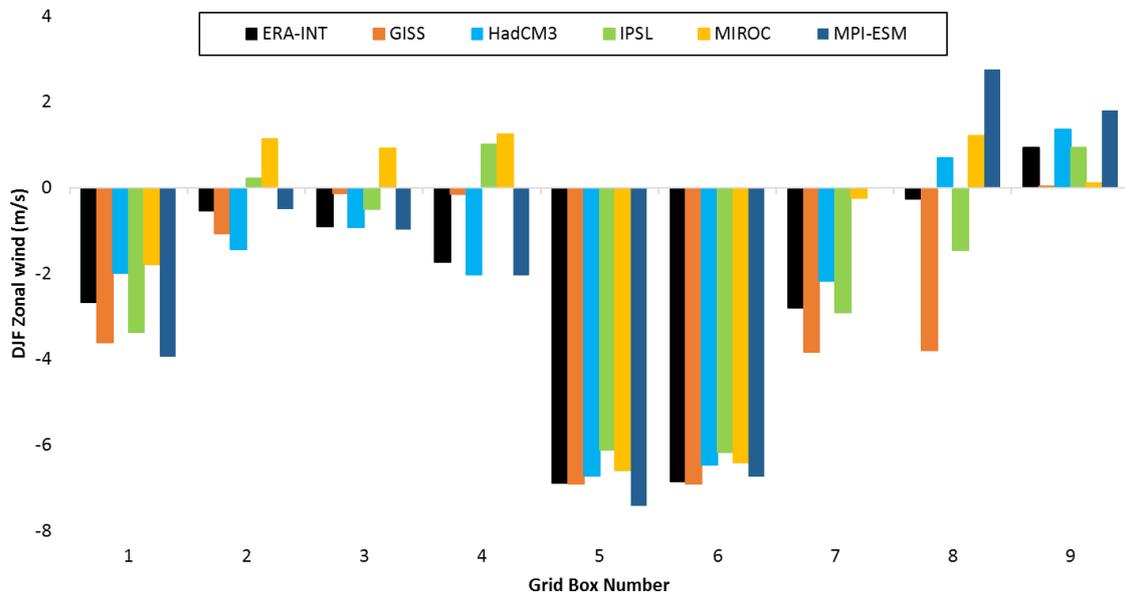


Fig. D1 DJF Mean zonal wind (ms^{-1}) per grid box (1-9), 1979-2004 from reanalysis and historical climate model simulations at $7.5^\circ \times 8^\circ$ resolution.

Table D2 Two sample t-test to analyse difference between model simulation DJF zonal mean climatologies compared against ERA-Interim for each grid box (1-9), 1979-2004. Bold numbers indicate the number of logbook grid boxes where the null hypothesis cannot be rejected at the 95% level.

	Grid box Number									Sum
	1	2	3	4	5	6	7	8	9	H=0
GISS-E2-R	1	1	1	1	0	0	1	1	1	2
HadCM3	1	1	0	0	0	1	0	1	1	4
IPSL-CM5A-LR	1	1	1	1	1	1	0	1	0	2
MIROC-ESM	1	1	1	1	0	1	1	1	1	1
MPI-ESM-P	1	0	0	0	1	0	1	1	1	4

Table D3 Two sample t-test to analyse difference between model simulation DJF zonal mean climatologies compared against ERA-20C for each grid box (1-9), 1900-1999. Bold numbers indicate the number of logbook grid boxes where the null hypothesis cannot be rejected at the 95% level.

	Grid box Number									Sum
	1	2	3	4	5	6	7	8	9	H=0
GISS-E2-R	1	1	1	1	1	0	1	1	1	1
HadCM3	1	0	1	1	1	1	1	1	1	1
IPSL-CM5A-LR	1	1	1	1	1	1	1	1	0	1
MIROC-ESM	1	1	1	1	0	1	1	1	1	1
MPI-ESM-P	1	1	1	1	1	0	1	1	1	1

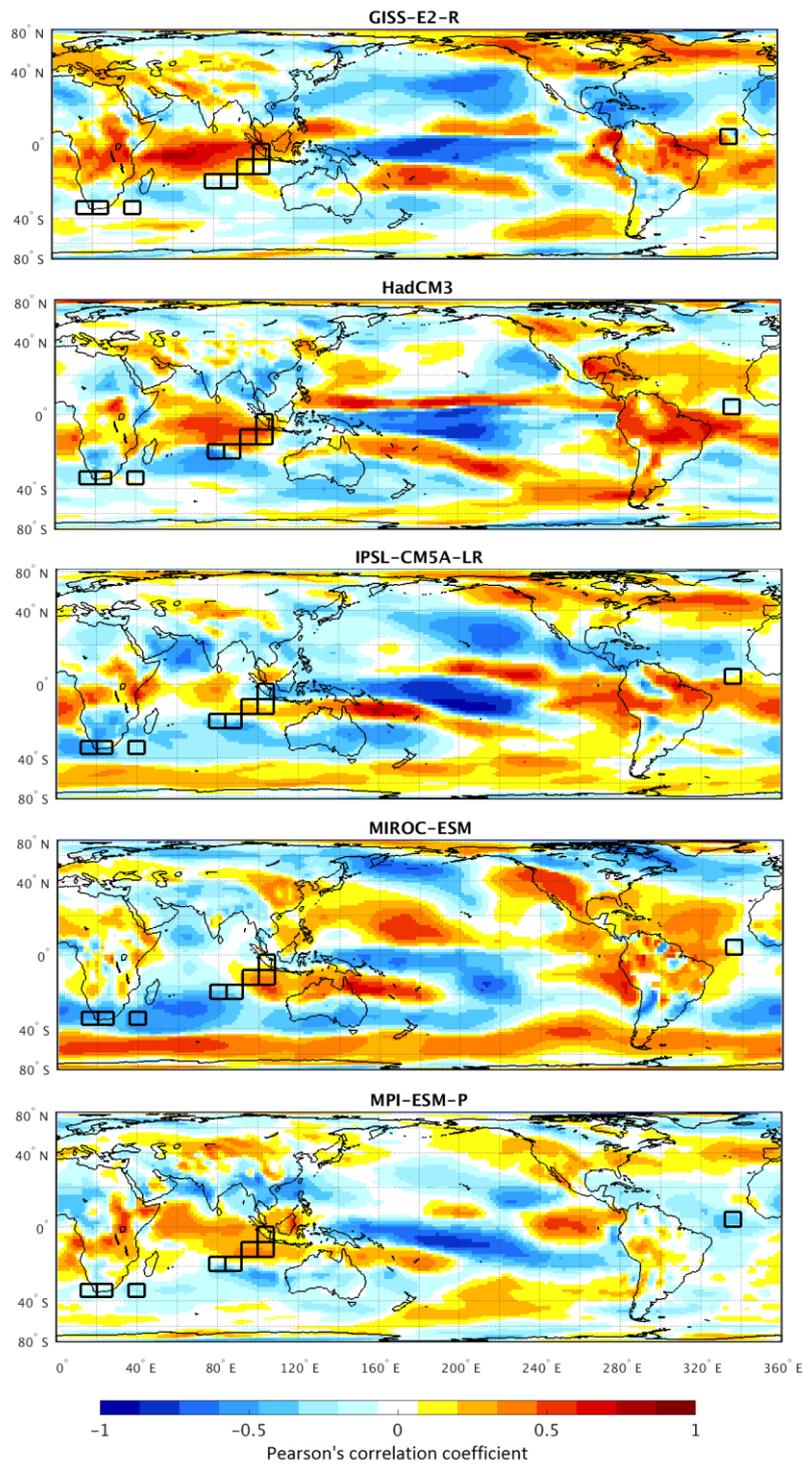


Fig. D2 Pearson's correlation coefficient between DJF zonal wind and the DJF SOI from climate model simulations, 1979-2004. All time series detrended prior to calculation of correlations. Black rectangles indicate logbook predictor grid boxes.

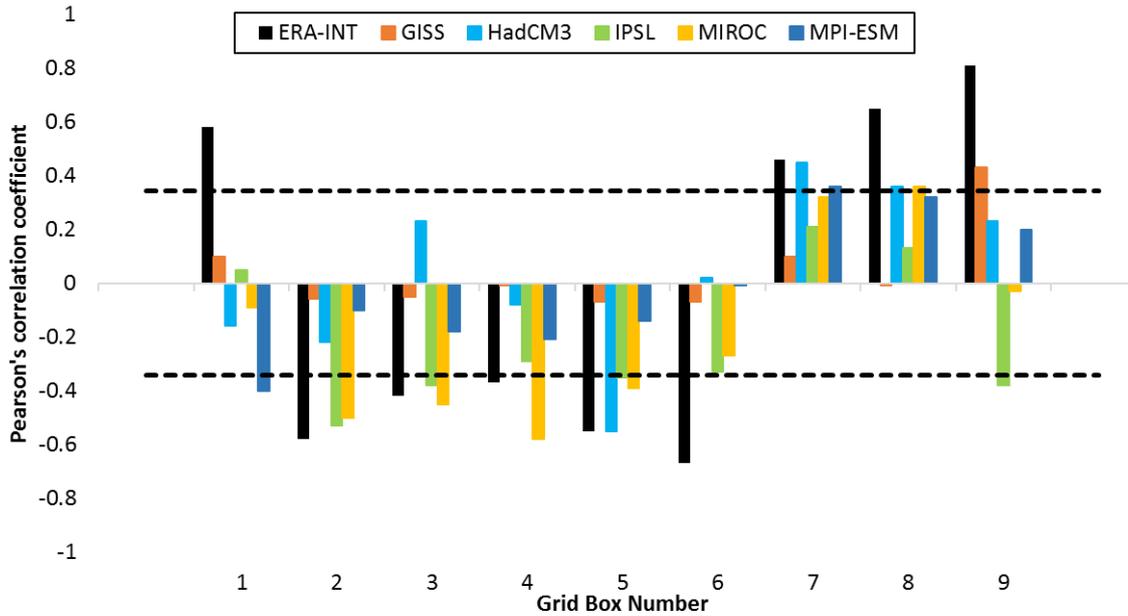


Fig. D3 Pearson's correlation coefficient between DJF zonal wind per grid box (1-9) and the DJF SOI from reanalysis and climate model simulations. Dashed lines indicates 90% significance value. Time series detrended prior to calculations of correlations.

Table D4 Number of logbook grid boxes with significant correlation (90% confidence level) of correct sign, and the number of grid boxes with the same sign correlation as ERA-Interim, 1979-2004, between the DJF zonal wind and the DJF SOI.

	Number of grid boxes with significant correlations	Number of grid boxes with same sign correlation as ERA-Interim
GISS-E2-R	1	8
HadCM3	3	6
IPSL-CM5A-LR	4	8
MIROC-ESM	5	7
MPI-ESM-P	2	8

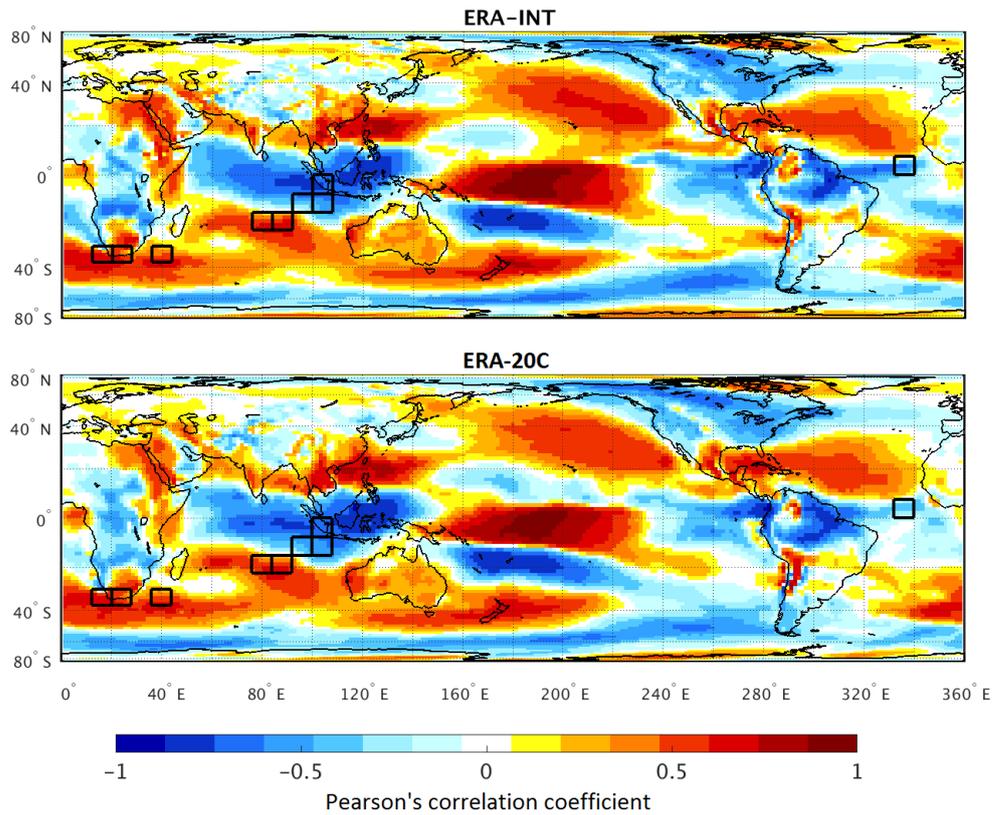


Fig. D4 Pearson's correlation coefficients between mean DJF zonal wind and the DJF Niño 3.4 index from ERA-Interim (top) and ERA-20C (below), 1979-2004. Data detrended prior to calculations of correlations. Black rectangles indicate logbook regions. All values shown.

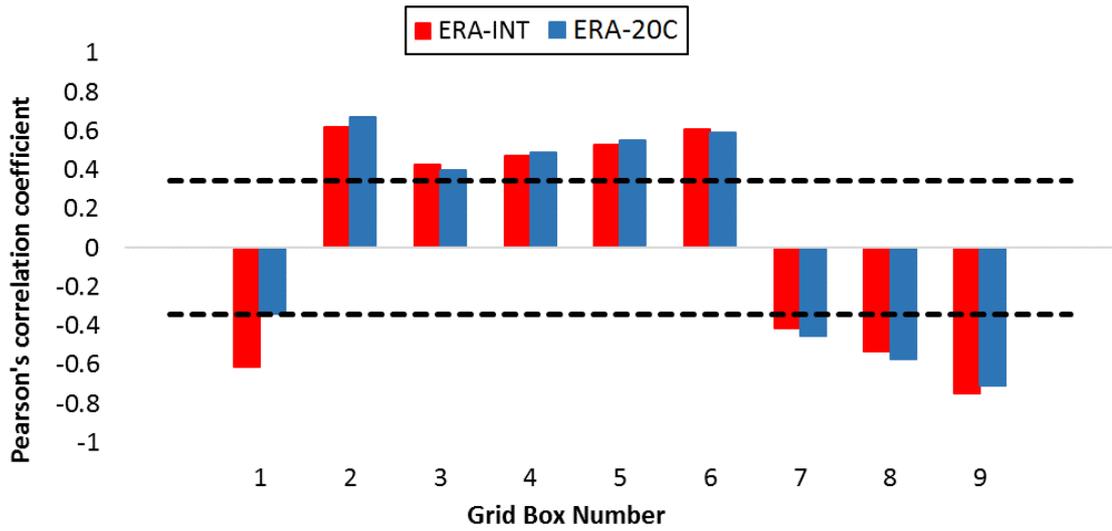


Fig. D5 Pearson's correlation coefficient between DJF zonal wind per grid box (1-9) and the DJF Niño 3.4 index from reanalysis. Dashed lines indicates significance at the 90% confidence level. Time series detrended prior to calculations of correlations.

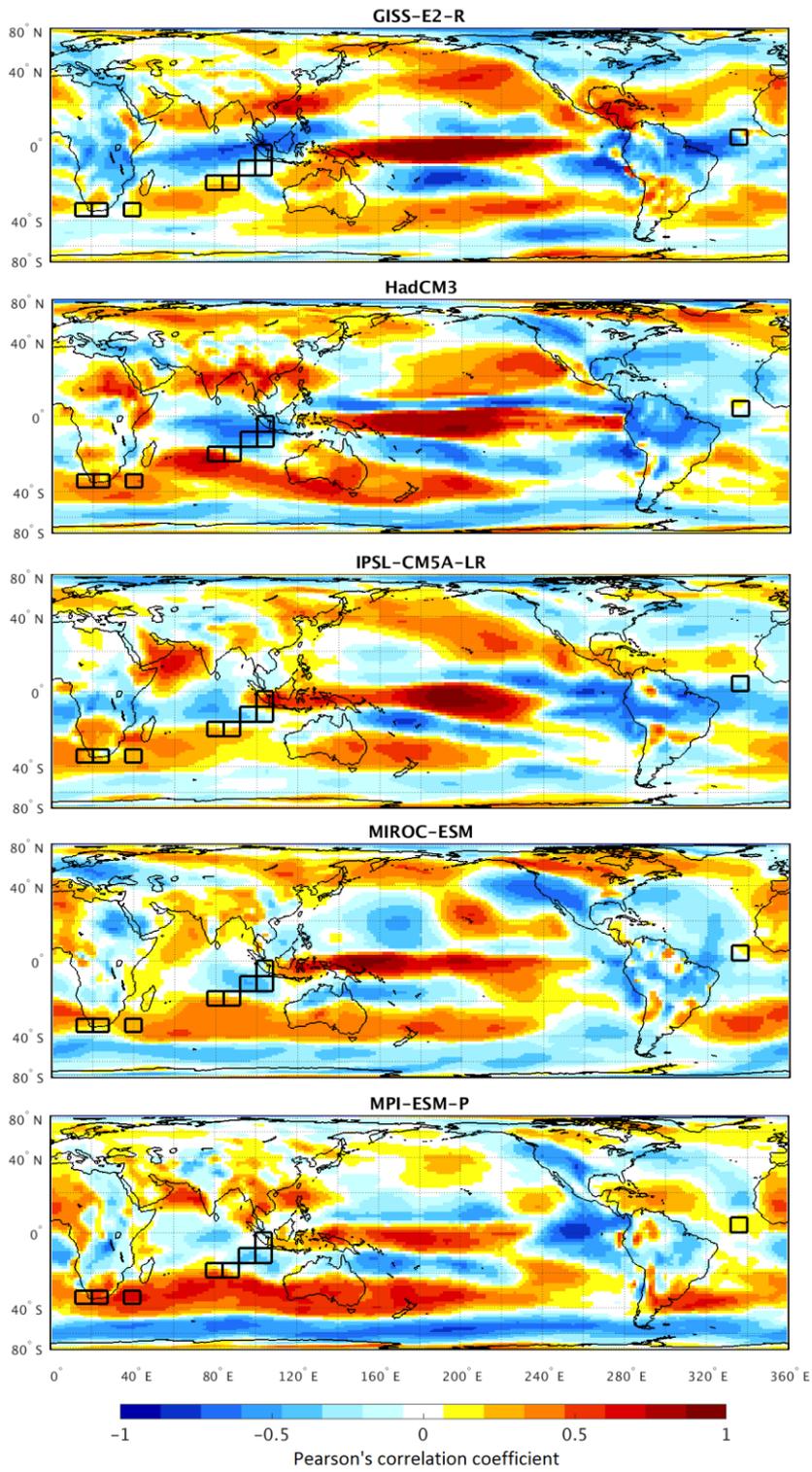


Fig. D6 Pearson's correlation coefficient between DJF zonal wind and the DJF Niño 3.4 index from climate model simulations, 1979-2004. All time series were detrended prior to calculation of correlations. Black rectangles indicate logbook grid boxes.

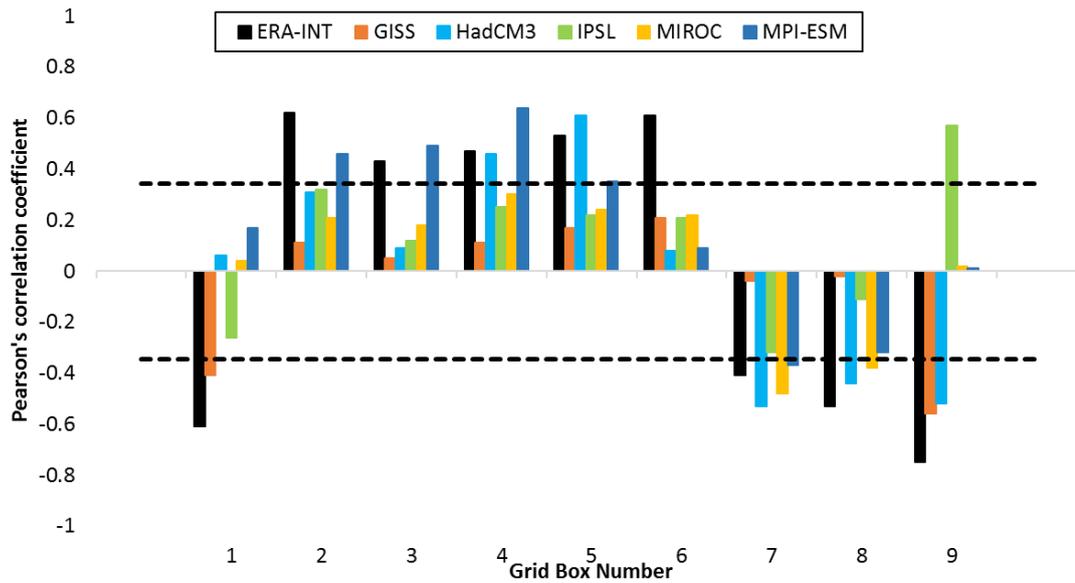


Fig. D7 Pearson's correlation coefficient between DJF zonal wind per grid box (1-9) and the DJF NINO3.4 index from reanalysis and climate model simulations. Dashed lines indicates significance at the 90% confidence level. Time series detrended prior to calculations of correlations.

Table D5 Number of logbook grid boxes with significant correlation (90% confidence level) of correct sign, and the number of grid boxes with the same sign correlation as ERA-Interim, 1979-2004, between DJF zonal wind and the DJF Niño 3.4 index.

	Number of grid boxes with significant correlations	Number of grid boxes with same sign correlation as ERA- Interim
GISS-E2-R	2	9
HadCM3	5	8
IPSL-CM5A-LR	0	8
MIROC-ESM	2	7
MPI-ESM-P	5	7

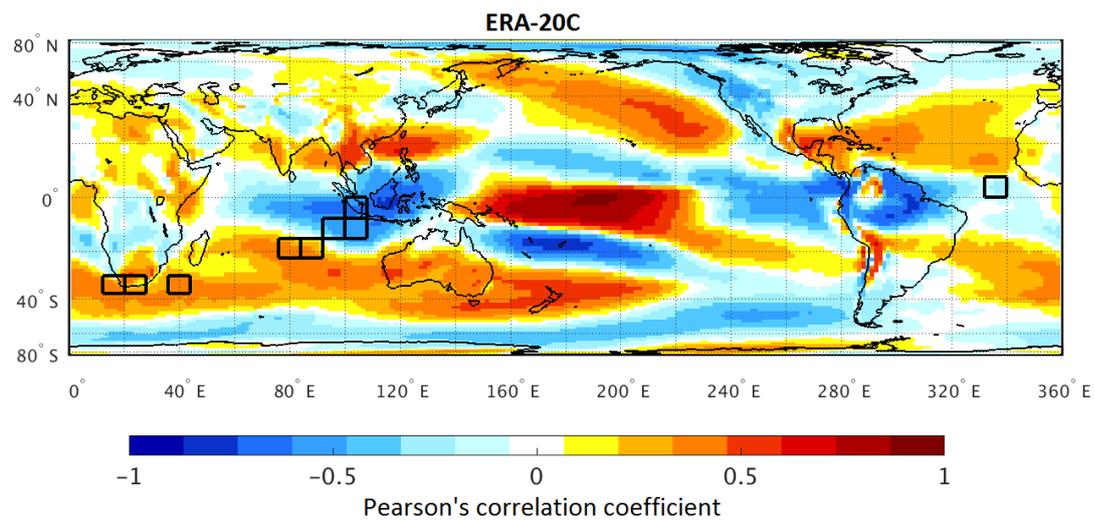


Fig. D8 Pearson's correlation coefficients between mean DJF zonal wind and the DJF Niño 3.4 index from ERA-20C 1900-1999. Data detrended prior to calculations of correlations. Black rectangles indicate logbook regions. All values shown.

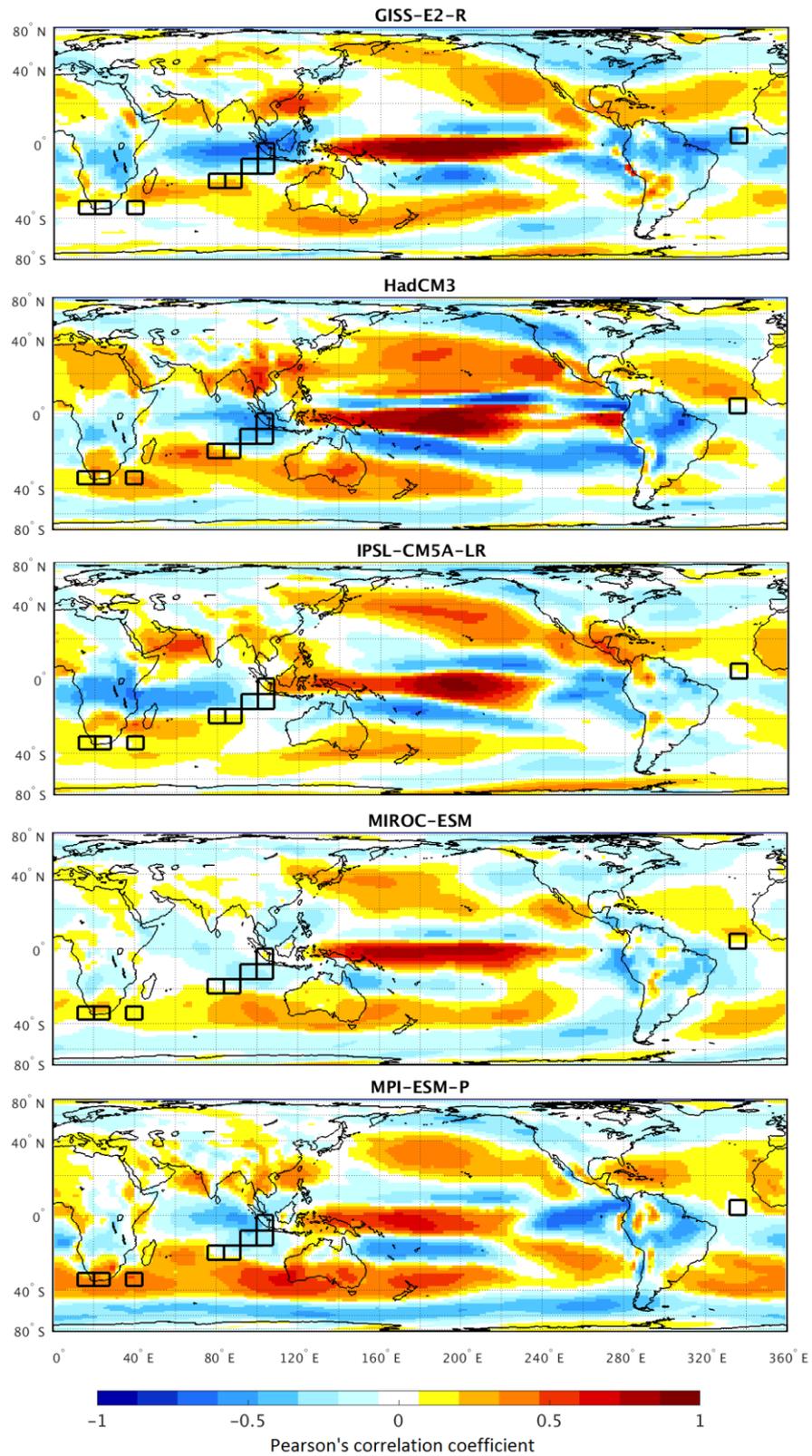


Fig. D9 Pearson's correlation coefficient between DJF zonal wind and the DJF Niño 3.4 index from climate model simulations, 1900-1999. All time series detrended prior to calculation of correlations. Black rectangles indicate logbook predictor grid boxes.

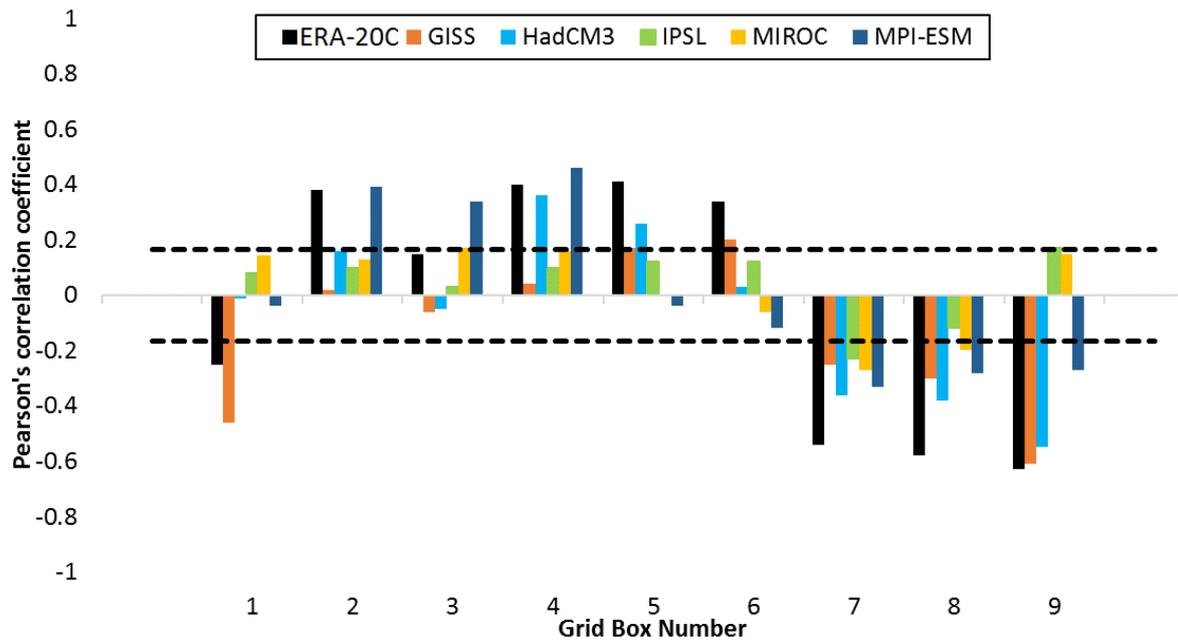


Fig. D10 Pearson's correlation coefficient between DJF zonal wind per grid box (1-9) and the DJF Niño 3.4 index from reanalysis and climate model. Dashed lines indicates significance at the 90% confidence level. Time series detrended prior to calculations of correlations.

Table D6 Number of logbook grid boxes with significant correlation (90% confidence level) of correct sign, and the number of grid boxes with the same sign correlation as ERA-20, 1900-1999, between DJF zonal wind and the DJF Niño 3.4 index.

	Number of grid boxes with significant correlations	Number of grid boxes with same sign correlation as reanalysis
GISS-E2-R	6	7
HadCM3	5	8
IPSL-CM5A-LR	2	7
MIROC-ESM	3	5
MPI-ESM-P	6	7

Table D7 Ranking of historical climate models for zonal wind and DJF Niño 3.4 index metrics compared to the ERA-20C, 1900-1999.

	Results from grid-box t-test of mean zonal wind	Rank for Niño 3.4	Rank for grid boxes with significant correlations	Rank for grid boxes with same sign correlation	Overall Ranking
GISS-E2-R	H=1	3	=1	=2	1
HadCM3	H=1	2	3	1	2
IPSL-CM5A-LR	H=1	1	5	=2	=3
MIROC-ESM	H=1	4	4	5	5
MPI-ESM-P	H=1	5	=1	=2	=3

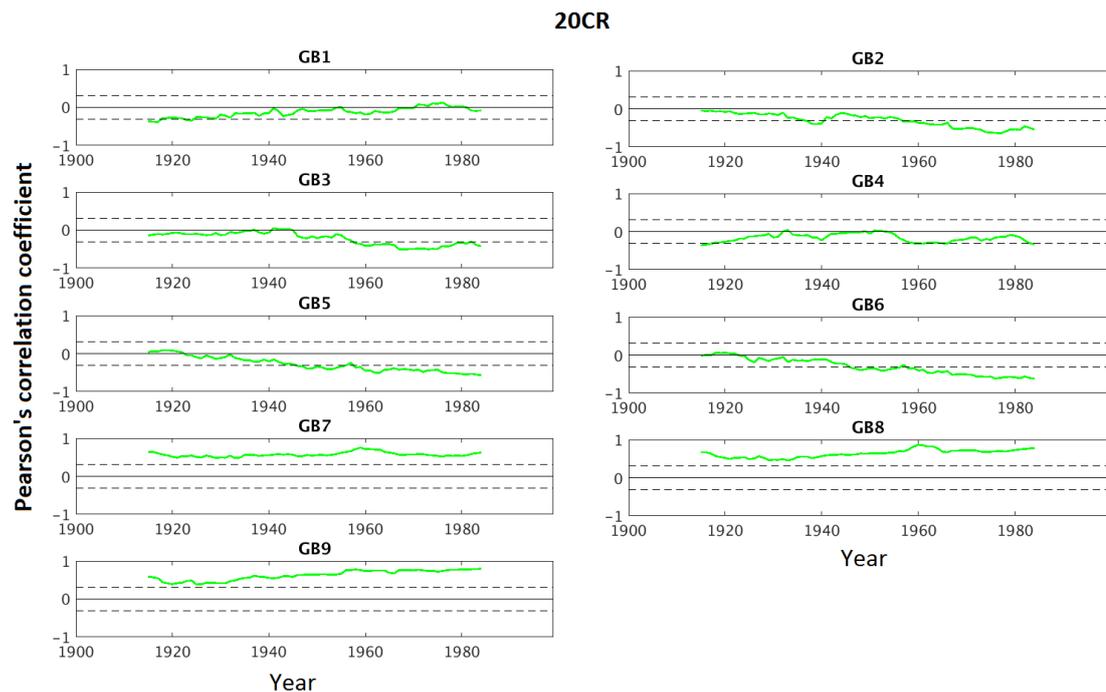


Fig. D11 30-year running correlation of DJF zonal wind and DJF SOI 1900-1999, 20CR. All time series were detrended before calculating running correlations. The year shown is the 15th year of the window, dashed line indicates significance at the 90% confidence level.

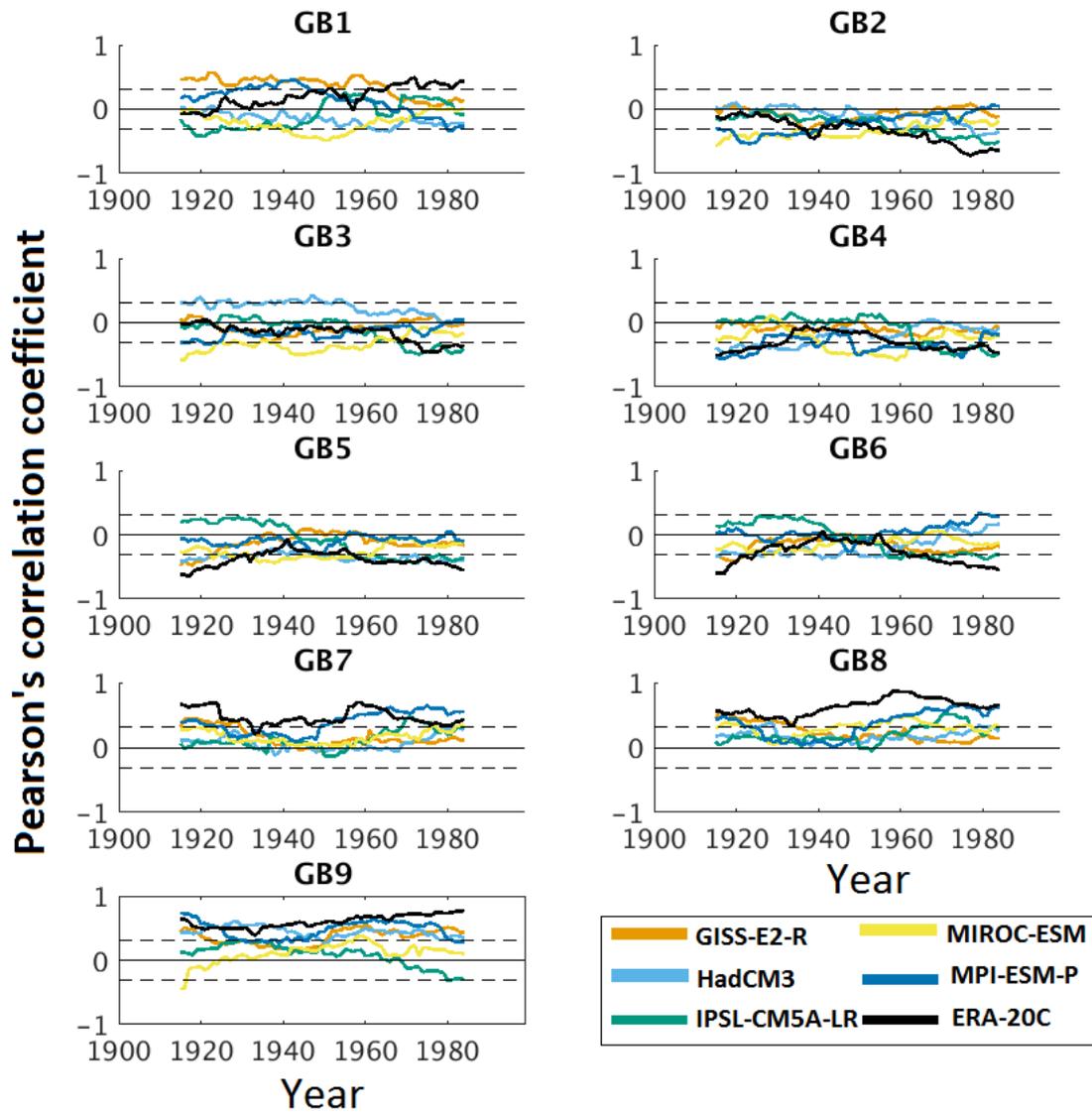


Fig. D12 30-year running correlation of DJF zonal wind and the DJF SOI 1900-1999, climate model simulations and ERA-20C. All time series were detrended before calculating running correlations. The year shown is the 15th year of the window, dashed line indicates significance at the 90% confidence level.

Appendix E

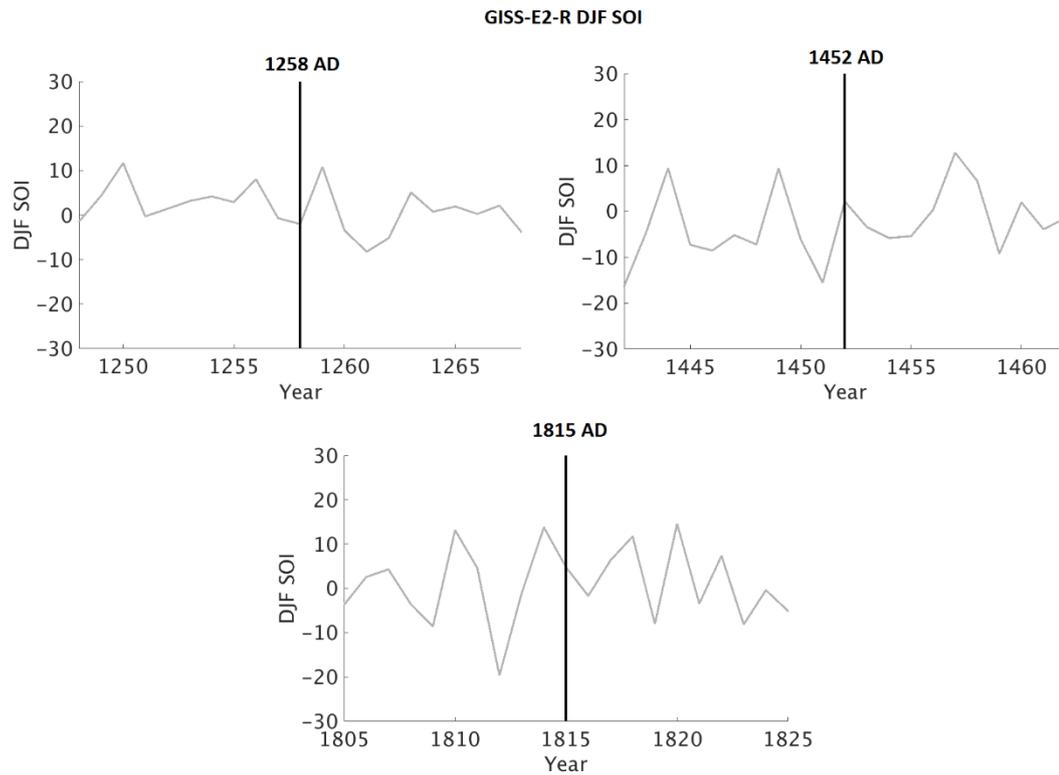


Fig. E1 The DJF SOI from GISS-E2-R LM simulation in the 20 years surrounding the three major volcanic events of the LM, 1258, 1452 and 1815 AD. Vertical lines indicate years of eruption.

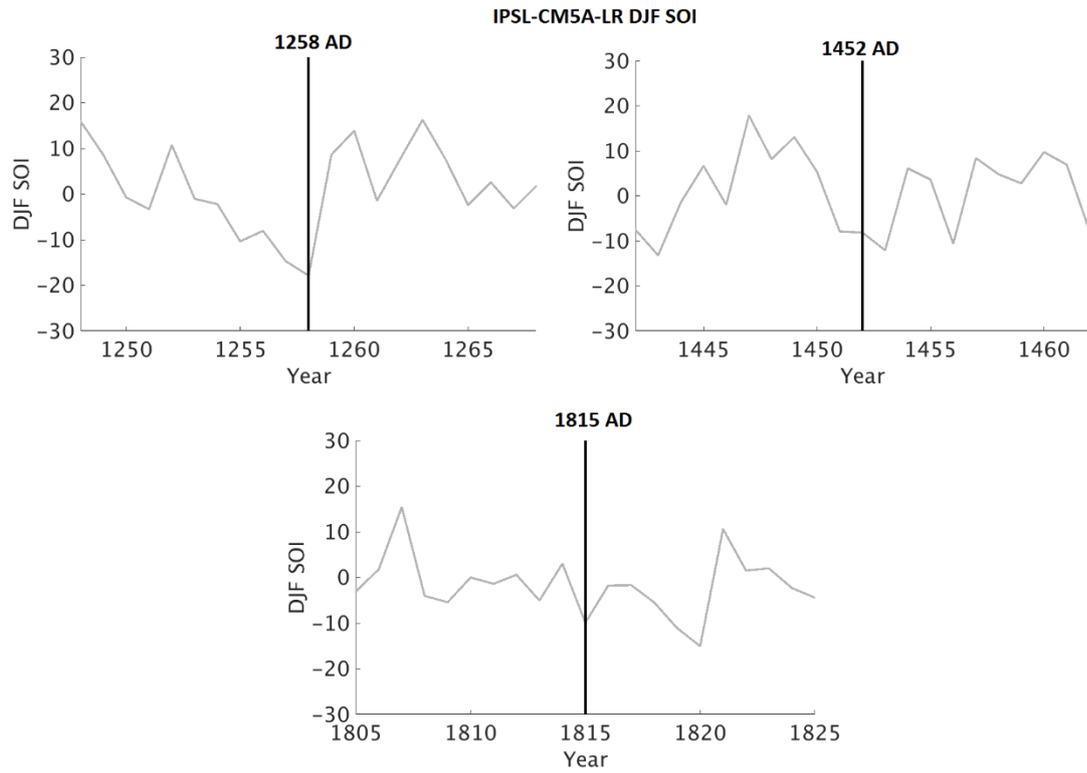


Fig. E2 The DJF SOI from IPSL-CM5A-LR LM simulation in the 20 years surrounding the three major volcanic events of the LM, 1258, 1452 and 1815 AD. Vertical lines indicate years of eruption.

Table E1 Number of El Niño and La Niña events during 100-year periods from Last Millennium simulations. (based on one standard deviation of the SOI from the mean of the 100-year period; SOI normalised over 1750-1848), and ENSO frequency (years).

	GISS-E2-R			IPSL-CM5A-LR		
	El Niño	La Niña	ENSO frequency	El Niño	La Niña	ENSO frequency
850-949	16	18	2.9	17	14	3.2
950-1049	18	17	3.1	15	14	3.4
1050-1149	13	18	3.2	17	15	3.1
1150-1249	15	14	3.4	13	18	3.2
1250-1349	14	15	3.4	19	19	2.6
1350-1449	13	15	3.6	15	16	3.2
1450-1549	17	17	2.9	20	14	2.9
1550-1649	17	19	2.8	16	13	3.4
1650-1749	15	16	3.2	12	13	4
1750-1849	14	12	3.8	16	16	3.1

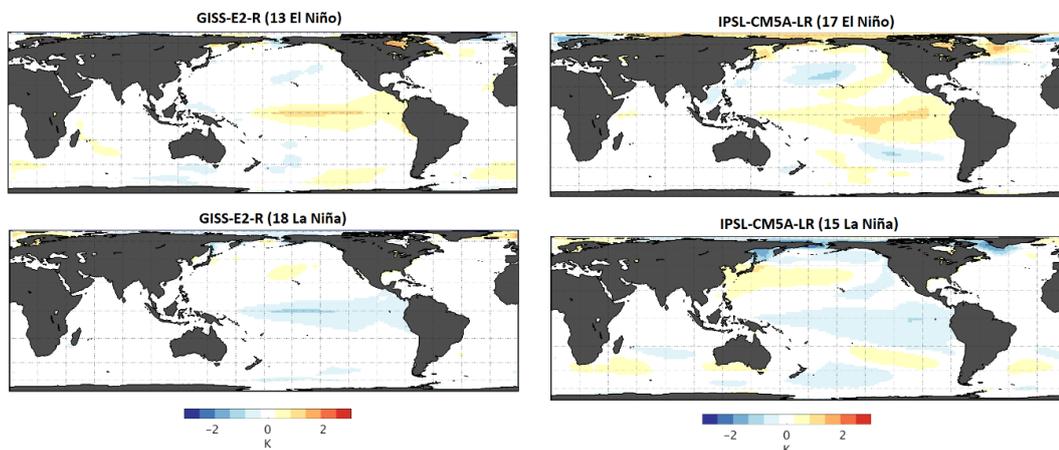


Fig. E3 El Niño and La Niña surface temperature (K) anomaly composites defined by one standard deviation of the DJF SOI from LM simulations, 1050-1149 from GISS-E2-R and IPSL-CM5A-LR. Numbers in brackets indicate the number of events used to calculate each composite plot.

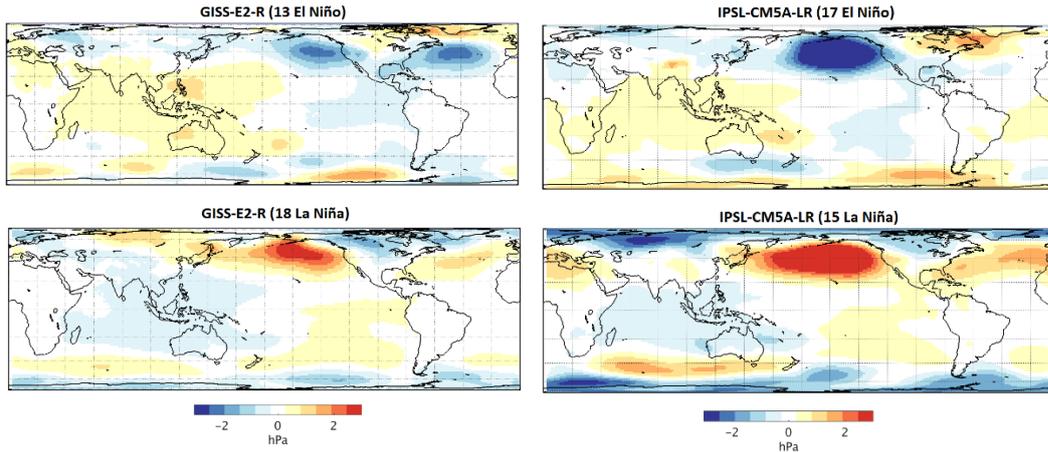


Fig. E4 El Niño and La Niña MSLPA (hPa) composites defined by one standard deviation of the DJF SOI from LM simulations, 1050-1149 from GISS-E2-R and IPSL-CM5A-LR. Numbers in brackets indicate the number of events used to calculate each composite plot.

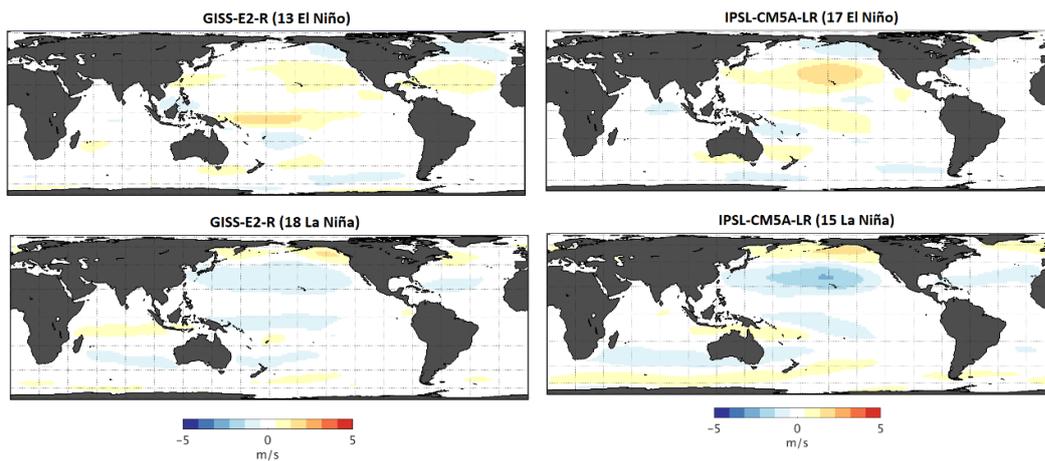


Fig. E5 El Niño and La Niña zonal wind (ms^{-1}) anomaly composites defined by one standard deviation of the DJF SOI from LM simulations, 1050-1149 from GISS-E2-R and IPSL-CM5A-LR. Numbers in brackets indicate the number of events used to calculate each composite plot.

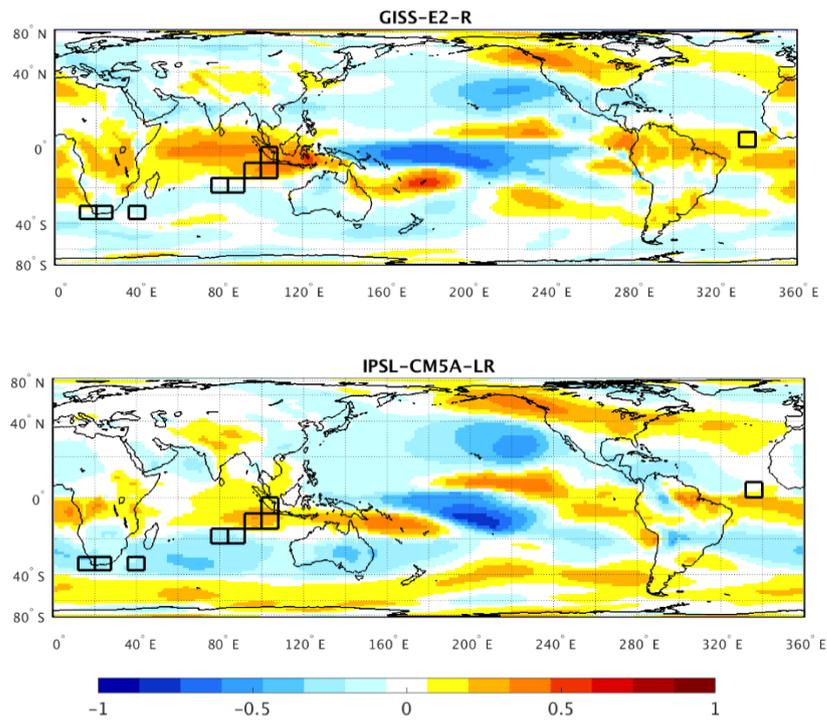


Fig. E6 Pearson's correlation coefficient between the mean DJF zonal wind and the DJF SOI from GISS-E2-R and IPSL-CM5A-LR 1750-1848, covering the logbook-based reconstruction period. All correlations shown. Calculated at $1.5^\circ \times 1.5^\circ$ degree resolution.

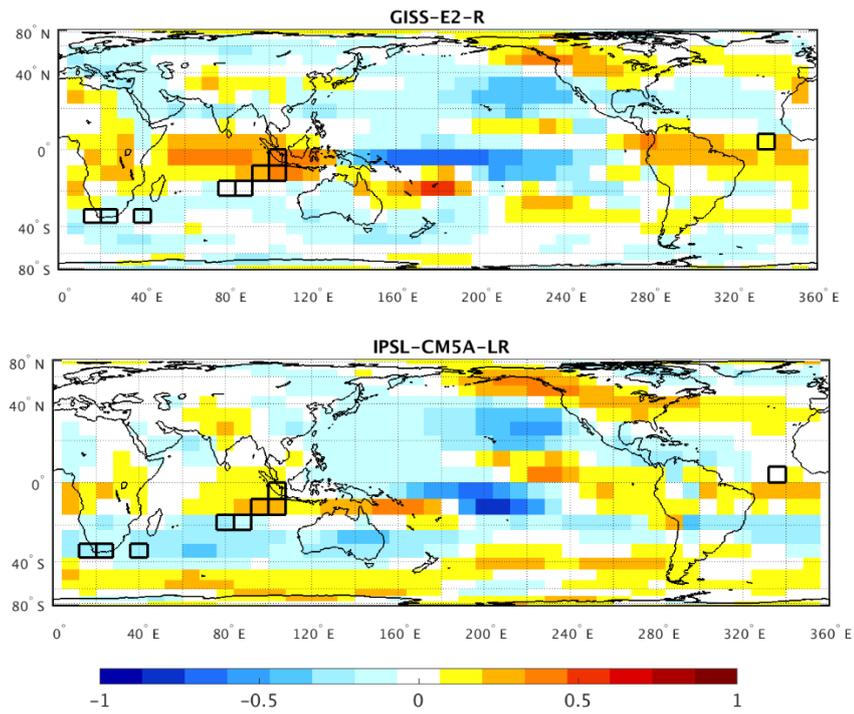


Fig. E7 Pearson's correlation coefficient between the mean DJF zonal wind and the DJF SOI from GISS-E2-R and IPSL-CM5A-LR 1750-1848, covering the logbook-based reconstruction period. All correlations shown. Calculated at $7.5^\circ \times 8^\circ$ degree resolution.

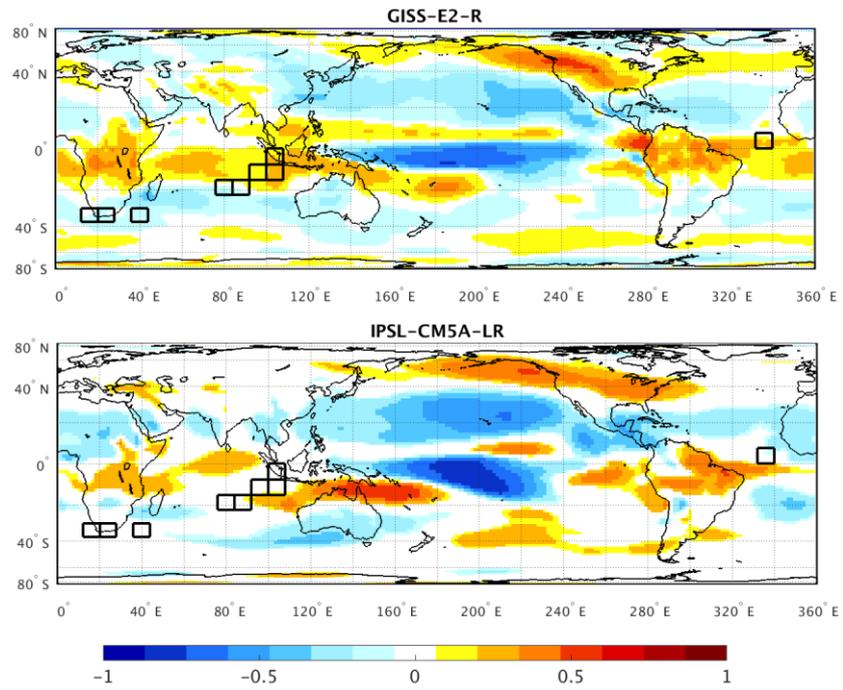


Fig. E8 Pearson's correlation coefficient between the mean DJF zonal wind and the DJF SOI from GISS-E2-R and IPSL-CM5A-LR for the 1050-1149 of the LM simulations covering the logbook-based reconstruction period. All correlations shown. 1.5° x 1.5° degree resolution.

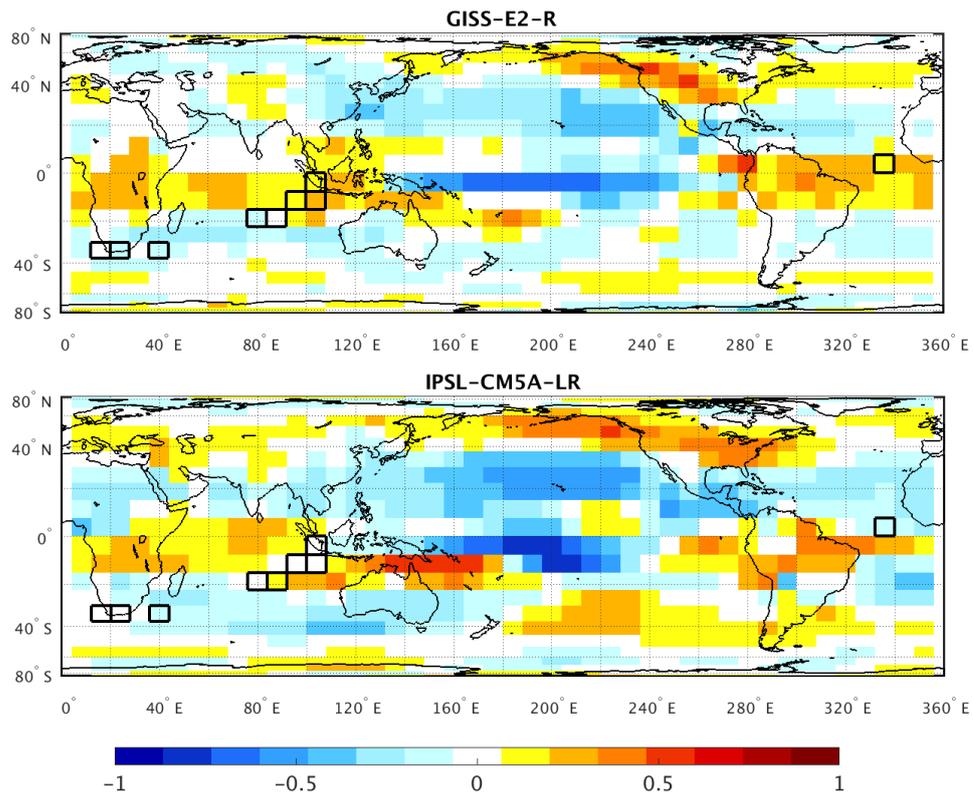


Fig. E9 Pearson's correlation coefficient between the mean DJF zonal wind and the DJF SOI from GISS-E2-R and IPSL-CM5A-LR for the 1050-1149 of the LM simulations covering the logbook-based reconstruction period. All correlations shown. $7.5^\circ \times 8^\circ$ degree resolution.

Table E2 Number of EP and CP El Niño events during 100-year periods from the Last Millennium simulations (based on one standard deviation from mean of 100-year period).

	GISS-E2-R			IPSL-CM5A-LR		
	EP El Niño	CP El Niño	CP-EP ratio	EP El Niño	CP El Niño	CP-EP ratio
	850-949	16	5	0.3	7	6
950-1049	11	9	0.8	9	12	1.3
1050-1149	13	7	0.5	12	6	0.5
1150-1249	14	11	0.8	13	12	0.9
1250-1349	11	6	0.5	15	9	0.6
1350-1449	15	8	0.5	8	10	1.3
1450-1549	14	10	0.7	12	7	0.6
1550-1649	14	8	0.6	8	11	1.4
1650-1749	14	7	0.5	6	8	1.3
1750-1849	12	5	0.4	8	10	1.3

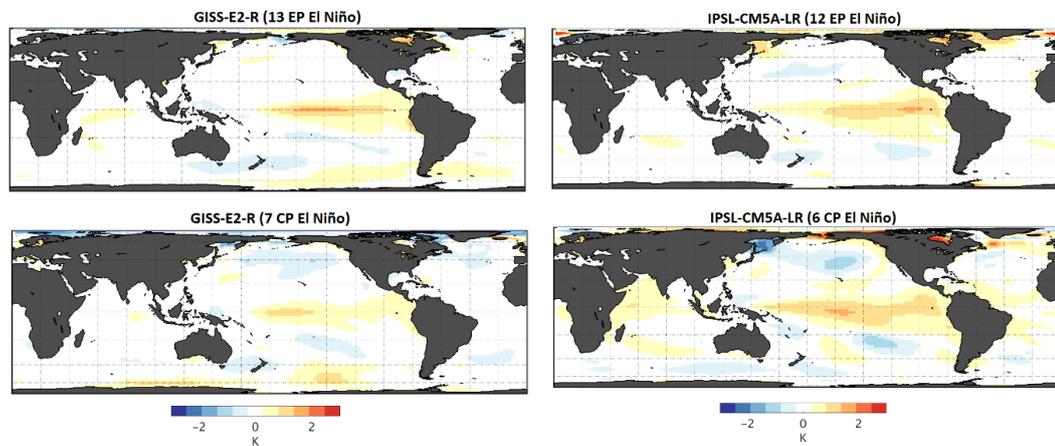


Fig. E10 EP and CP El Niño surface temperature (K) anomaly composites defined by one standard deviation of NINO method from the LM simulations, 1050-1149 from GISS-E2-R and IPSL-CM5A-LR. Numbers in brackets indicate the number of events used to calculate each composite plot.