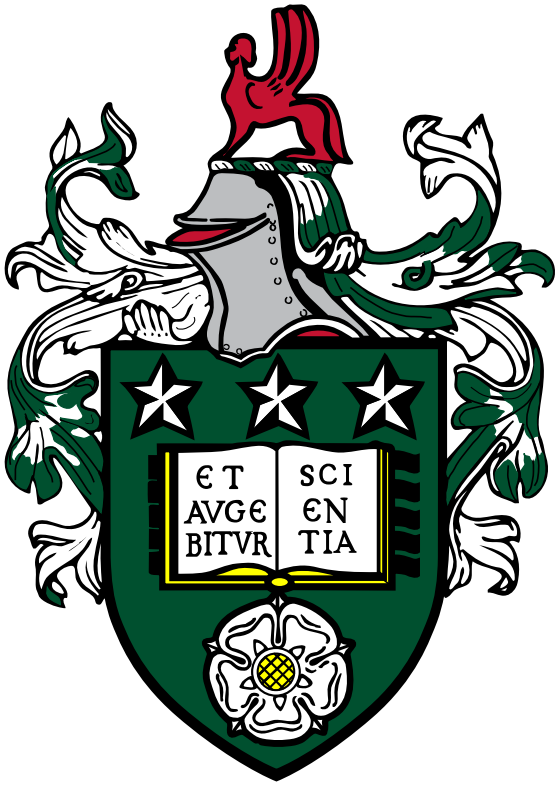
Use of artificial intelligence to understand mountain weather and climate processes



Jonathan David Coney

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

The University of Leeds
School of Earth and Environment

May 2024

Intellectual property and Publication statements

I confirm that the work submitted is my own, except where work which has formed part of jointly authored publications has been included. My contribution and the other authors to this work has been explicitly indicated below. I confirm that appropriate credit has been given within the thesis where reference has been made to the work of others. The right of Jonathan David Coney to be identified as author of this work has been asserted by Jonathan David Coney in accordance with the Copyright, Designs and Patents Act 1988.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

Details of collaboration with others are listed below.

- Sections of Chapter 2, Chapter 3 and much of Chapter 4 contain material from the article: Coney, J., Denby, L., Ross, A. N., Wang, H., Vosper, S., van Niekerk, A., Dunstan, T. and Hindley, N. 2024. Identifying and characterising trapped lee waves using deep learning techniques. *Quarterly Journal of the Royal Meteorological Society.* **150**(758), pp. 213–231. DOI: 10.1002/qj.4592.

  Jonathan Coney was the lead author of this publication, led the work within and wrote the original draft of the article, with discussion and ideas suggested by all co-authors. The author contributions of all the authors are listed below. Leif Denby provided assistance with generating the synthetic wave packets and probing the learned features of the models, and Neil Hindley provided advice and information on the S-transform.

Jonathan Coney:	Conceptualization, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft.
Leif Denby:	Conceptualization, Funding acquisition, Methodology, Software, Supervision, Writing – review & editing.
Andrew N Ross, Simon Vosper:	Conceptualization, Funding acquisition, Methodology, Supervision, Writing – review & editing.
He Wang, Annelize van Niekerk, Tom Dunstan:	Conceptualization, Methodology, Supervision, Writing – review & editing.
Neil Hindley:	Software, Methodology, Writing – review & editing.

Chapter 4 also contains two additional sections, that were not published in the above article. These sections contain:

- Some extra analysis of the model using dimensionality reduction and further exploration of the neural network (Section 4.3.3).
- An exploration of crowdsourced labelling of lee waves using Zooniverse (Section 4.3.4).

- Parts of Chapter 2, Chapter 3 and most of Chapter 5 contain material that has been written as a journal article to be submitted to the *Quarterly Journal of the Royal Meteorological Society*.

  Jonathan Coney will be the lead author of this article, led the work within and wrote the original draft of the article, with discussion and ideas suggested by all co-authors. The author contributions of all the authors are listed below.

Jonathan Coney:	Conceptualization, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft
Andrew N. Ross, Leif Denby, Simon Vosper:	Conceptualization, Funding acquisition, Methodology, Supervision, Writing — review & editing
He Wang, Annelieze van Niekerk, Tom Dunstan:	Conceptualization, Methodology, Supervision, Writing – review & editing

Acknowledgements

Firstly, thank you to my supervisors: Andrew, Leif, He, Simon, Annelize and Tom. I couldn't have asked for a more enthusiastic group of people to help me through and make my PhD as interesting and as enjoyable as it has been. It truly has been a pleasure to work with you all.

Thank you too to the Orography group at the Met Office – currently Steve, Peter and Callum: it has been great to meet with you and talk all things mountain meteorology. Thanks also, to Jonathan Wilkinson and Rob Neal at the Met Office for providing the weather patterns data, and for useful discussion around using them on the lee waves climatology. Thanks too, to Neil Hindley for helping me make sure I got the S-transform right.

I am grateful to NERC for providing funding through the Panorama DTP, and to the Met Office for the CASE award, supervision and access to all their data.

My office mates in 11.06 deserve a huge amount of thanks for tolerating me and my attention span, as well as the people of ICAS for making it a super institute to be in. Thank you to the dynamics group and Chatmosphere for having me along. Special thanks go to Amethyst and Sarah for the moral support throughout. Thanks too to the Scientific Machine Learning group, who have been a fantastic resource and a great bunch of people to chat about machine learning with. The potato club must receive thanks for its stable dinnertime arrangement over the past couple of years despite the Refectory's erratic pricing arrangements – Ben, Ben, Callum, Fran, Iain, Jamie, Roy and Will. Thank you to all my housemates along the way for having me (currently the dynamic duo of Ben and Jamie). Of the many, many, highlights, I think that Productivity Cafe, Only Connect and being chased by cattle in Harewood have been particularly memorable.

Thank you to my mates from North Leeds Fell Runners for keeping me sane, and thank you especially to everyone who gave me lifts to the moor on cold dark evenings, or to races several hours away, not to mention tolerating my love of Pendle. But extra special shout outs to Andy, Jenny and Nick for all the lifts, Ian for getting me round the OMM, and Dave, Will and Josh for our friendly rivalries. Thank you Ben, James and James for getting me onto the fells in the first place. Thank you too, to all my good friends from parkrun, too many to list everyone here: but thank you for all your encouragement along the way, especially Anne, Noel, Frank and Stuart.

Lastly, thank you to my parents Sue and Paul, my brother Richard and the rest of my family for everything you have done for me along the way. This wouldn't have been possible without your support.

Abstract

Trapped lee waves, and resultant turbulent rotors downstream, present a hazard for aviation and land-based transport. While high resolution numerical weather prediction (NWP) models can represent such phenomena, there is currently no simple and reliable automated method for detecting the extent and characteristics of these waves in model output. Spectral transform methods can be used to characterise regions of wave activity in model and observational data, however these methods can be slow and have their limitations. Machine learning techniques offer potentially fruitful methods for tackling this problem.

This thesis presents the development of deep learning models to detect and characterise trapped lee waves from the characteristic patterns made by trapped lee waves in NWP model output, performing well against hand-labels and spectral-derived characteristics. The deep learning models are applied to a large archive of high resolution NWP model data to produce climatology information for both the present-day climate and a future climate projection. The climatology is interrogated, showing that there is no diurnal cycle of lee waves, but there is a seasonal cycle and influence of synoptic weather effects on lee waves and their characteristics. The future climate projections (under Representative Concentration Pathway 8.5) show little headline change in occurrence or characteristics of lee waves, but imply changes to wave occurrence within different weather patterns, and a risk of more high amplitude ($> 3\,\mathrm{m\,s^{-1}}$) waves in the future.

These deep learning models could prove useful for forecasting in the development of a computationally cheap post-processing tool for operational meteorologists, to be able to more easily visualise the effects of lee waves and the potential hazards involved. The climatology information explored in this thesis has informed further understanding about lee waves, such as the weather conditions that result in the

strongest lee wave amplitudes, as well as potential changes to lee wave activity under climate change.

Contents

Inclusion of copyrighted material

- Figure 1.2(b) contains Meteosat Second Generation Imagery and was obtained from the Centre for Environmental Data Analysis archive. Replication for academic purposes is permitted by licence. The image is copyright EUMETSAT/Met Office 2023.
- Figure 2.1 is reproduced from Figure 1 of Teixeira, M. A. 2014. The physics of orographic gravity wave drag. *Frontiers in Physics*. **2**(July), pp. 1–24. DOI: 10.3389/fphy.2014.00043. The journal article is licensed for reuse under CC BY 3.0.
- Figure 2.3 is reproduced from Figure 8 of Hertenstein, R. F. and Kuettner, J. P. 2005. Rotor types associated with steep lee topography: influence of the wind profile. *Tellus A: Dynamic Meteorology and Oceanography*. **57**(2), p. 117. DOI: 10.3402/tellusa.v57i2.14625. The journal article is licensed for reuse under CC BY 4.0.
- Figure 2.4 is reproduced from Figure 11 of Morcrette, C., Brown, K., Bowyer, R., Gill, P. and Suri, D. 2019. Development and Evaluation of In-Flight Icing Index Forecast for Aviation. *Weather and Forecasting*. **34**(3), pp. 731–750. DOI: 10.1175/WAF-D-18-0177.1. The journal article is licensed for reuse under CC BY 4.0.
- Figure 2.5 is reproduced from Figure 9.2 of Mohri, M., Rostamizadeh, A. and Talwalkar, A. 2018. *Foundations of Machine Learning*. 2nd. London: MIT Press. The book is licensed for reuse under CC BY-NC-ND 4.0. No changes were made to the image before its inclusion in this thesis.

List of Abbreviations

3DVOM	Three-Dimensional Velocities Over Mountains, the now retired Met Office lee wave model.
AI	artificial intelligence
AIRS	Atmospheric Infrared Sounder
ANN	artificial neural network
CEDA	Centre for Environmental Data Analysis
CMIP5	Coupled Model Intercomparison Project Phase 5, in some literature "Coupled" is replaced with "Climate".
CNN	convolutional neural network
CPM	convection permitting model
ECMWF	European Centre for Medium-Range Weather Forecasts
ERA	ECMWF Reanalysis
ENDGame	Even Newer Dynamics for General atmospheric modelling of the environment, the dynamical core of the Met Office Unified Model.
FFT	Fast (Discrete) Fourier Transform
GPU	graphics processing unit
IFS	Integrated Forecasting System
IPCC	Intergovernmental Panel on Climate Change
LOLAN	Low Level Aviation, used to describe a type of weather chart produced by the Met Office, also known as a F215.
ML	machine learning
MASS	Managed Archive Storage System
MetUM	Met Office Unified Model
MSG	Meteosat Second Generation
MSLP	mean sea level pressure
NAO	North Atlantic Oscillation
NERC	Natural Environment Research Council
NWP	numerical weather prediction
PCA	Principal Component Analysis
PDF	probability density function
PM	particulate matter
PPE	perturbed parameter ensemble

RCP	Representative Concentration Pathway
RCM	regional climate model
RMSE	root mean squared error
S-transform	Stockwell transform
SHAP	SHapley Additive exPlanation
T-REX	Terrain-Induced Rotor Experiment
TS	time slice
UK	United Kingdom
UKCP18	United Kingdom Climate Projections
UTC	Coordinated Universal Time, sometimes also shortened to “Z”.
VHF	Very High Frequency

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Trapped lee waves

Trapped lee waves are a type of internal gravity wave, which occur in a stratified fluid with gravity as the restoring force. In this case, air that is stably stratified ascending over orography can trigger waves that propagate away from the orography (Durran, 2003). If this energy is trapped within a layer in the atmosphere, these oscillations can propagate downstream on the lee side of the orography for tens of kilometres (Vosper et al., 2018). Lee waves can lead to the generation of rotors, low level regions of strong turbulence and overturning flow associated with gravity waves that form on the lee side of the orography (Doyle and Durran, 2002), as well as strong downslope winds (Colfescu et al., 2021).

Lee waves can sometimes produce lenticular clouds, where air condenses at each wave crest, as seen for example in Figure 1.1 where some lenticular clouds can be seen on the southern (lee) side of Pen-y-Ghent in the Yorkshire Dales. Characteristic striped cloud patterns from lee wave activity can also be visible in satellite imagery (e.g. Figure 1.2 [b]), and can be seen by eye in high resolution numerical weather prediction (NWP) model output, such as within the vertical velocity field.

Some of the earliest studies of lee waves came from a desire to to understand lenticular clouds and what are now known as rotors in the late 19th and early 20th centuries. Scorer (1961) describes how lee waves were exploited by glider pilots in the 1930s, keen to achieve height during flights, who discovered that the winds on

Figure 1.1: Lenticular clouds associated with lee wave activity, downstream from Pen y Ghent on the left (northern) side of the image. The wind was approximately northerly on this day (blowing from left to right).

the lee side of hills were more favourable for "soaring" than on the windward side. Lee waves became recognisable by glider pilots from their characteristic clouds, or winds associated with the formation of clouds on the lee side of the hills. There has been a scientific interest since the late 19th Century in the Helm Wind (Britain's only named wind), a strong downslope wind accompanied by the formation of rotors, which forms over Cross Fell in the Pennines in the north of England (Marriott, 1886; Marriott, 1889; Manley, 1945). The latter commented on the usefulness of the rotor formations to glider pilots, and theorised that a standing wave train is set up in Helm Wind conditions, with a series of stationary evenly spaced bars of cloud. Figure 1.2 (a) shows a photograph depicting a Helm Bar (a characteristic lee wave cloud that forms during a Helm Wind), and the accompanying Meteosat Second Generation (MSG) image (Figure 1.2 [b]) from a similar time over Britain and Ireland shows some of the characteristic lee wave cloud patterns present over Cumbria to the West of Cross Fell.

Rotors can be hazardous for air and road transport (for example contributing to strong turbulence experienced by pilots, aeroplane crashes and strong gusts at the surface: Ágústsson and Ólafsson, 2014; International Civil Aviation Organization, 1968; Vosper et al., 2013). Meteorological agencies are interested in forecasting rotors and their effects to be able to give detailed information to (among others) the aviation and freight industries. Developing forecasting tools to identify lee waves better, such as recognising wave activity from high resolution NWP model output,

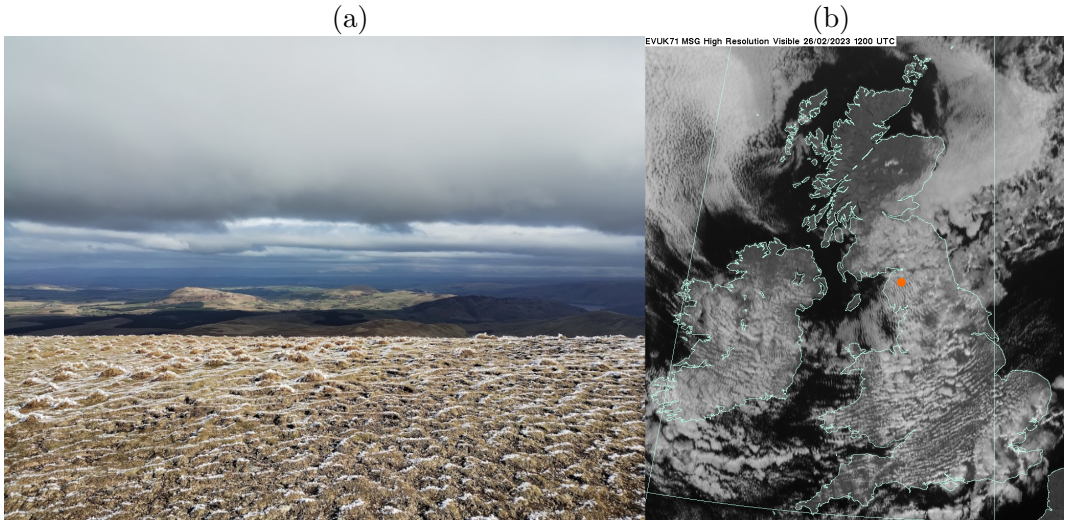


Figure 1.2: (a) A photograph looking east from the Eastern fells of the Lake District (left), towards Cross Fell, showing a Helm bar. Cross Fell is visible on the horizon. The MSG Visible satellite image (b) from 20 minutes before the photograph was taken is shown on the right. An orange dot shows the location of the photographer in (a). A set of lee waves are visible over the Lake District in the satellite image. Satellite image copyright EUMETSAT/Met Office 2023.

means that forecasters could be more informed on precisely where lee wave activity is likely to occur with a domain, and how severe turbulence associated with wave activity could be. By understanding which meteorological variables change when lee waves occur, one can understand the drivers of lee wave conditions: potentially making forecasting lee waves even when they are not explicitly resolved by NWP simpler. One way to develop these tools would be to utilise progress made in the past decade in artificial intelligence and machine learning.

1.2 Artificial intelligence and machine learning

Artificial intelligence (AI) is a broad term used to describe a large group of methods for computers to interpret and learn from data, which are used to perform specific tasks (Haenlein and Kaplan, 2019). Machine learning (ML) refers to a subset of AI tasks, and describes the process of training a model to have skill at a given task (such as classification, regression or pattern discovery) in a data set, and then assessing the skill of the trained model at performing that task on unseen data (Reichstein et al., 2019). One such task is segmentation: classifying each pixel within an input image as belonging to a class. One simple example of this could be distinguishing between the pixels in a photograph: classifying each pixel as either containing part

of a cat or the background.

Deep learning methods describe a set of machine learning techniques that construct large hierarchical structures to learn from and perform a task on a data set. Deep learning typically involves the construction of artificial neural networks (ANNs) and has seen increasing popularity and usage in recent years, thanks to the improvement in computing power, especially the availability of graphics processing units (GPUs) and the development of highly parallelised code and algorithms designed to train models on GPUs, 10 to 20 times faster than had been done previously (Pandey et al., 2022; LeCun et al., 2015). The term "neural network" was originally chosen due to the method's architecture trying to imitate how the brain learns and stores information, but in reality ANNs are combinations of multiplication of tensors (n-dimensional matrices) alongside nonlinear functions (Goh, 1995). Further details about machine learning are given in Section 2.2.

In 1947, Alan Turing gave one of the first public lectures to mention machine learning to the London Mathematical Society in which he proposed a machine which learned from experience (Copeland, 2004, p. 375). Research in machine learning methods was explored in the following decades, but with small progress because the then computing capabilities were unable to process large ANNs, and there was a belief that training would not produce satisfactorily trained models (Haenlein and Kaplan, 2019; LeCun et al., 2015). It was not until the availability of computing power capable of training large neural networks in realistic lengths of time that the method took off, leading to the current interest in deep learning, with applications within the past decade vastly improving on previous efforts at using ML.

Hence, the use of ANNs was small until the mid-2000s onwards, when computing power started to become large enough to produce sufficiently large, well-trained deep learning models. Some early examples of artificial neural networks being used during this resurgence include Krizhevsky et al. (2012), who used ANNs for image classification, beating previous methods considerably. Silver et al. (2016) developed a neural network to play the game Go, which has significantly more possible moves than chess, making it harder for a player (human or computer) to consider the best moves to make in a situation in Go, than in chess. The ANN, AlphaGo, won 99.8% of games against other Go programs and beat the then European champion 5–0: which was thought to be the first time a computer had beaten a professional Go

player.

Machine learning models have been used within atmospheric science, and increasingly so in recent years, typically for performing tasks using model data or observations. Some applications within atmospheric science involved the application of models to supplement observations or NWP, such as the automated the classification of clouds (Denby, 2020) or short-term "nowcasting" of precipitation (Shi et al., 2015). More recently, entire NWP models have been replicated by machine learning models with impressive results, and this is explored further in Section 2.2.

1.3 Motivation and thesis structure

This thesis looks at the use of machine learning tools on a large amount of NWP model output in order to learn about lee waves and draw conclusions about what affects their generation.

Since lee waves are resolved by high resolution NWP models and patterns from lee wave activity can be identified by eye, it makes sense to explore the usage of model output to better inform forecasting personnel, as well as investigating the relationship between other model variables and lee wave activity, including high amplitude cases. Lee waves are a mechanism for momentum transport in the atmosphere and it is important to forecast lee wave and rotor activity well, for aviation and road transport. For example, pilots need to be aware of the potential for lee waves to be generated, so as to avoid rotors and regions of strong turbulence that can be generated by high amplitude lee waves.

Using data-driven methods means that conclusions can be drawn directly from the data. Machine learning methods are typically data-hungry, and there are vast archives of well-labelled and organised meteorological data available in various archives that are ripe for exploration or extracting further information using machine learning methods, which have shown vast improvement in skill within recent years. Machine learning offers the ability to apply a well-performing model to weather data (such as NWP model output, or observations) to perform a task typically only performed in the past by humans. This means it brings an opportunity to develop a computationally cheap post-processing system to convey lee wave warning information

to forecasters. In addition, it allows the opportunity to efficiently process large amounts of historical data to learn about the drivers of lee waves. By applying the same machine learning models to NWP model output forced to a future climate scenario, insights can be made to how lee waves may change under climate change.

Therefore the aims of this thesis are the following:

- Test whether machine learning models can be used to detect and characterise lee waves from high resolution Met Office NWP model output.
- Produce insights into the drivers of lee waves over Britain and Ireland and the effects of different synoptic weather patterns on lee waves, by applying deep learning models to a large high resolution data set of model output of the present-day climate.
- Examine how the occurrence and characteristics of lee waves over Britain and Ireland may change under a high emissions future climate scenario, by using the machine learning models on high resolution model output of a future climate scenario, and comparing this to the data from the present-day climate.

Chapter 2 contains an overview of the current literature about lee waves and machine learning, Chapter 3 introduces the data sources used in this thesis, Chapter 4 contains the results from training ML models to recognise and characterise lee waves, Chapter 5 discusses the results from applying these trained ML models to a climatology data set of Britain and Ireland's present-day climate, and Chapter 6 analyses how lee waves may change over Britain and Ireland under a future climate scenario. Finally, some concluding remarks are presented in Chapter 7.

Chapter 2

Literature Review

This chapter presents an overview of the literature surrounding lee waves, machine learning techniques, synoptic weather patterns, and climate change projections under different emissions scenarios.

2.1 Trapped lee waves

Mountain waves are internal gravity waves produced by the forced ascent of stably stratified air over orography (Durran, 2003). Mountain waves can be subdivided into two main types: vertically propagating waves which can propagate vertically for kilometres above the mountains, and up to the stratosphere (Smith, 2003); and trapped lee waves (lee waves for short), where the gravity waves are trapped in the lower troposphere and propagate horizontally rather than up into the stratosphere or mesosphere. Figure 2.1 shows an idealised mountain and generation of both vertically propagating and trapped lee waves for simple two-dimensional flow, with horizontal wind speed $U(z)$, along with some simple rotors forming under the crest of the lee waves. This thesis looks at the latter of these mountain waves: trapped lee waves.

Depending on the vertical atmospheric profile, lee waves occur over Britain and Ireland on a regular basis, propagating downwind from orography, despite the orography being relatively small in height and width compared to other mountain ranges globally. Lee waves can propagate horizontally in excess of 100 km beyond the moun-

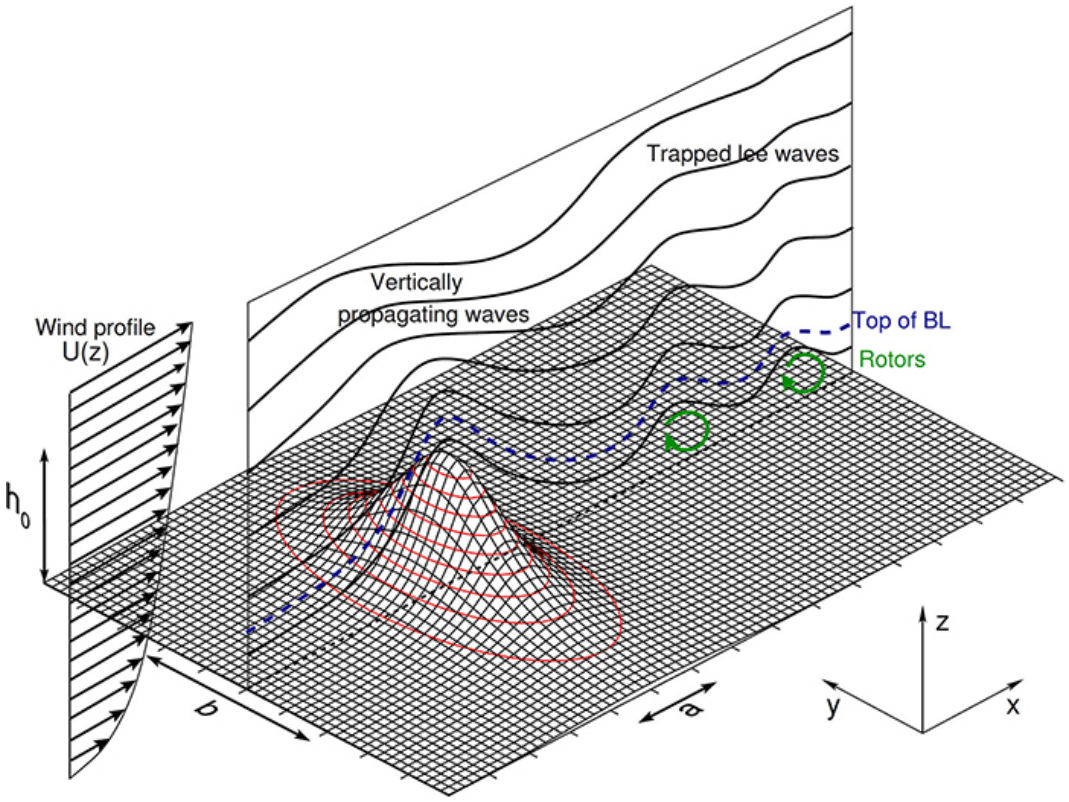


Figure 2.1: A simplified two-dimensional schematic showing the generation of vertically propagating waves and trapped lee waves for a given horizontal wind profile for a theoretical isolated mountain, taken from Figure 1 of Teixeira (2014).

tain with horizontal wavelengths typically between 5 and 35 km (Ralph et al., 1997; American Meteorological Society, 2012). The Scorer parameter (Scorer, 1949) is often used to diagnose conditions conducive to mountain wave activity, either vertically propagating waves or trapped lee waves, as considered here. The Scorer parameter at a height z depends on the horizontal wind speed $U(z)$ and the Brunt-Väisälä frequency $N(z)$. The Brunt-Väisälä frequency at some height z is a measure of atmospheric stability, defined as (e.g. Durran, 2003):

$$N = \sqrt{\frac{g}{\theta}\frac{\partial\theta}{\partial z}},$$

where θ is the potential temperature and g the acceleration due to gravity. The Scorer parameter can be derived from linear wave theory in a dry inviscid atmosphere, using a modified Taylor-Goldstein equation (e.g. Markowski and Richardson, 2010; Vosper, 2003). The Scorer parameter l at a height z is defined as (e.g. Bretherton, 1969):

$$l^2(z) = \frac{N^2}{U^2} - \frac{1}{U}\frac{\mathrm{d}^2 U}{\mathrm{d}z^2}.$$

In general, the first term in the Scorer parameter is the more important and the second (shear) term is only significant when the wind shear is large (Blockley and Lyons, 1994). Gravity waves can propagate vertically if their horizontal wavenumber is greater than l.

While the full Scorer parameter takes into account the stability, wind speed and wind shear, this second shear term $\frac{1}{U}\frac{\mathrm{d}^2 U}{\mathrm{d}z^2}$ is often small, and is 0 in the case of constant U (Nance and Durran, 1998). Blockley and Lyons (1994) note that the second term is only significant when wind shear is large. In this thesis, the Scorer parameter is taken as $l = \sqrt{\frac{N^2}{U^2}}$, since the shear term is often small, and the model data used in this thesis is only available on limited pressure surfaces as it is. Calculating the second derivative of U with respect to z would only limit values of the Scorer parameter available for analysis, since more data would be required to approximate $\frac{\mathrm{d}^2 U}{\mathrm{d}z^2}$ on a given pressure surface. In addition, several other studies use this simplified Scorer parameter, such as Udina et al. (2020), Bramberger et al. (2018) and Vosper et al. (2013). Blockley and Lyons (1994) note that "most authors" neglect the shear term, however this study is now over 30 years old, but does demonstrate precedent for neglecting the shear term when calculating the Scorer parameter, along with more recent studies cited above which show that this simplification is still not uncommon.

If the Scorer parameter decreases with height, then these are favourable conditions for trapped lee waves (Durran, 2003). That is, lee wave generation in general relies on strong horizontal winds and decreasing stability with height. High amplitude lee waves lead to a higher likelihood of rotors, regions of strong turbulence downstream of the orography, which can be hazardous for air and road transport (Vosper et al., 2013). Therefore, forecasting lee wave activity is important for meteorological agencies, for example being able to advise aviation personnel on the likelihood of rotors around airfields and flight paths. In addition, lee waves are a source of horizontal momentum transport in the atmosphere, as well as wave drag (Bretherton, 1969; Shutts, 1992). Lee waves have been resolved in high resolution simulations of the Met Office Unified Model (MetUM) since an upgrade of the dynamical core in 2015, and are forecasted well in line with observations (Sheridan et al., 2017,

expanded in Section 2.1.4).

2.1.1 Controls on trapped lee waves

The Taylor-Goldstein equation describes gravity waves in the atmosphere (Nappo, 2002). For a simplified case with constant density with height (Markowski and Richardson, 2010),

$$\frac{\mathrm{d}^2\hat{w}}{\mathrm{d}z^2} + \left[\frac{N^2}{(c_x - u)^2} + \frac{1}{(c_x - u)}\frac{\mathrm{d}^2 u}{\mathrm{d}z^2} - k^2\right]\hat{w} = 0$$

where $\hat{w}$ is the complex wave amplitude, u the wind speed in the x (horizontal) direction, c_x is the phase speed of the waves in the x direction with wavenumber k (assuming a steady state ($c_x = 0$), the Scorer parameter can be substituted into the square brackets). This equation has solution of the form $\hat{w} = Ae^{imz} + Be^{-imz}$, where A and B are complex wave amplitudes. If m is real, then vertically propagating waves are generated. If m is imaginary, then the generated waves will be trapped (Holton and Hakim, 2013).

When the phase speed c_x is the same as the horizontal wind speed u in the Taylor-Goldstein equation, this is a singularity known as a critical level (Markowski and Richardson, 2010). At critical levels, the waves can become unstable, overturn and break, causing turbulence (Grubišić, 1997; Wurtele et al., 1996). There have been few observational studies of wave breaking: Strauss et al. (2015) summarised known observations of mid-troposphere wave-breaking, from Europe, the USA and Antarctica, with reported turbulent kinetic energies of these wave-breaking events varying from $4.6\,\mathrm{m^2 s^{-2}}$ to $150\,\mathrm{m^2 s^{-2}}$.

Idealised modelling of trapped lee waves typically use simplified conditions and model setup, such as simple flows such as linear ridges (Vosper et al., 2006) or the Witch of Agnesi (Bell, 1975), which although useful for developing theory when processes are mostly linear, do not accurately reflect the complex non-linear interactions in the atmosphere. Vosper (2003) developed a lee wave forecasting system (more in Section 2.1.4) to produce numerical solutions to the Taylor-Goldstein equation. They used radiosonde observations for the background flow and used the fluctuations in vertical ascent rate of the radiosondes to represent gravity wave motion,

finding that the model produced vertical velocities within $0.6\,\mathrm{m\,s^{-1}}$ of the radiosonde ascent velocity. Colfescu et al. (2021) observed a lee wave event in Iceland using an aircraft, and used both linear theory and a NWP model to simulate the lee wave generation, along with the aircraft observations to verify the modelled waves. However, linear theory has limits. Vosper (2004) note that amplitude of waves is governed by non-linear processes for more moderately sized hills, and lee wave amplitude underestimated by linear wave theory when the wavelength is significantly less than the width of the hill.

2.1.2 Lee wave occurrence and characteristics

The prevalence and characteristics of lee waves over Britain have been investigated before. Vosper et al. (2013) presented results from a three-year climatology of trapped lee waves, using the output from the (then operational at the Met Office) Three-Dimensional Velocities Over Mountains (3DVOM) model. 3DVOM was operational at the Met Office from May 2006 until the early 2020s, and is an idealised linearised model driven by a single profile from the global Met Office model configuration, with realistic orography (Vosper, 2003). They found that lee waves were more likely to occur in winter months than in summer months, and large amplitude lee waves (waves with vertical velocities $> 3\,\mathrm{m\,s^{-1}}$) occur more frequently in the Scottish Highlands than in North Wales or the Pennines. Worthington (2006) explored the possibility of a diurnal cycle of lee waves over Britain using Very High Frequency (VHF) radar measurements (from one radar station near Aberystwyth in Wales), and satellite imagery on a region covering Wales, Ireland and much of the Midlands and Northern England from 1990–2006. They found no evidence of a diurnal cycle in the orientation or amplitude of mountain waves, but there was a seasonal cycle in the amplitude of mountain waves over their area of observations.This previous work did not consider different synoptic weather conditions over Britain and Ireland such as weather patterns, and only considered regions of known wave activity over Britain and Ireland, rather than an analysis of lee wave activity within the entirety of the islands.

However, there may be a diurnal cycle of mountain waves in other parts of the world, despite there being no evidence of such a cycle in the studies completed in Britain, albeit with their limitations. Ruff and Ólafsson (2019) attribute a small

diurnal cycle in downslope windstorms in Iceland to the changing prevalence of gravity waves throughout the day. During the Terrain-Induced Rotor Experiment (T-REX) campaign in the Sierra Nevada, California, Jiang and Doyle (2008) used observations and numerical models to diagnose the impact of a diurnal cycle on the flow within the valley, and on downslope winds. This diurnal cycle was only present for more moderate downslope winds, but not stronger cases of downslope winds where dynamic forcing was the main driver (Sheridan and Vosper, 2012).

There have been several climatologies of mountain waves produced, in different parts of the world, in addition to those discussed above which focused on Britain and Ireland (Worthington, 2006; Vosper et al., 2013). These have focused on areas of large orography, such as the Alps and North America. Wilson et al. (1991) produced a climatology of vertically propagating waves between 30 km and 75 km in altitude, using lidar observations. They found a seasonal cycle of wave activity, with maxima during the winter months.

Grubišić and Billings (2008) produced a climatology of mountain wave clouds over the Sierra Nevada from visual inspection of visual satellite imagery, from a two-year period (1999 to 2001 inclusive). They classified mountain waves into single wave clouds or a lee wave "train" of clouds forming the characteristic striped pattern. They stated that the contemporary pattern identification methods from data (a small ANN: Peak and Tag, 1992) were insufficient at identifying mountain wave clouds from satellite imagery compared to the human eye, something that more sophisticated models from developments within machine learning since the 1990s should be able to supplant. Grubišić and Billings (2008) found that the most trapped lee wave events around the Sierra Nevada occurred in April, with large variability in January and April. However, this study only used two years worth of data, and by using observations of wave clouds, limited themselves to cases where there was sufficient moisture for cloud to be present. Lester (1978) also used clouds to produce a climatology of lee waves, for Pincher Creek in Alberta, an area of relatively high topography in Canada. They used hourly daylight observations made by staff at the Pincher Creek observatory covering most of January 1964 to April 1967 and acknowledged the subjectivity in deciding whether a cloud was a wave cloud by a human observer. The data showed more lee wave clouds occurred during cooler seasons of the year, and that there was a diurnal cycle of lee wave cloud formations,

with minima at sunrise and sunset, and maxima during the mid to late afternoon.

By producing a climatology of lee waves over Britain and Ireland using data from more sophisticated models, over a longer time period than those considered above, and by taking into account more than amplitude, evidence either way will be clear as to if there exists a diurnal cycle in lee waves in Britain and Ireland as has been observed elsewhere (e.g. Lester, 1978). Since there is a diurnal cycle in the inversion in the boundary layer, it is feasible to suggest that this could affect the trapping of lee waves, and perhaps the amplitude of lee waves at different times of the day.

Vicari et al. (2024) analysed trapped gravity waves over the eastern Pacific (concluding that the gravity waves were more likely to be generated by convection rather than orography as is more often the case in Britain and Ireland), using a Gaussian filter on water-vapour measuring satellite brightness temperatures. They found that trapping was linked to "increased upper-tropospheric wind shear", and occurred more often in December, January and February.

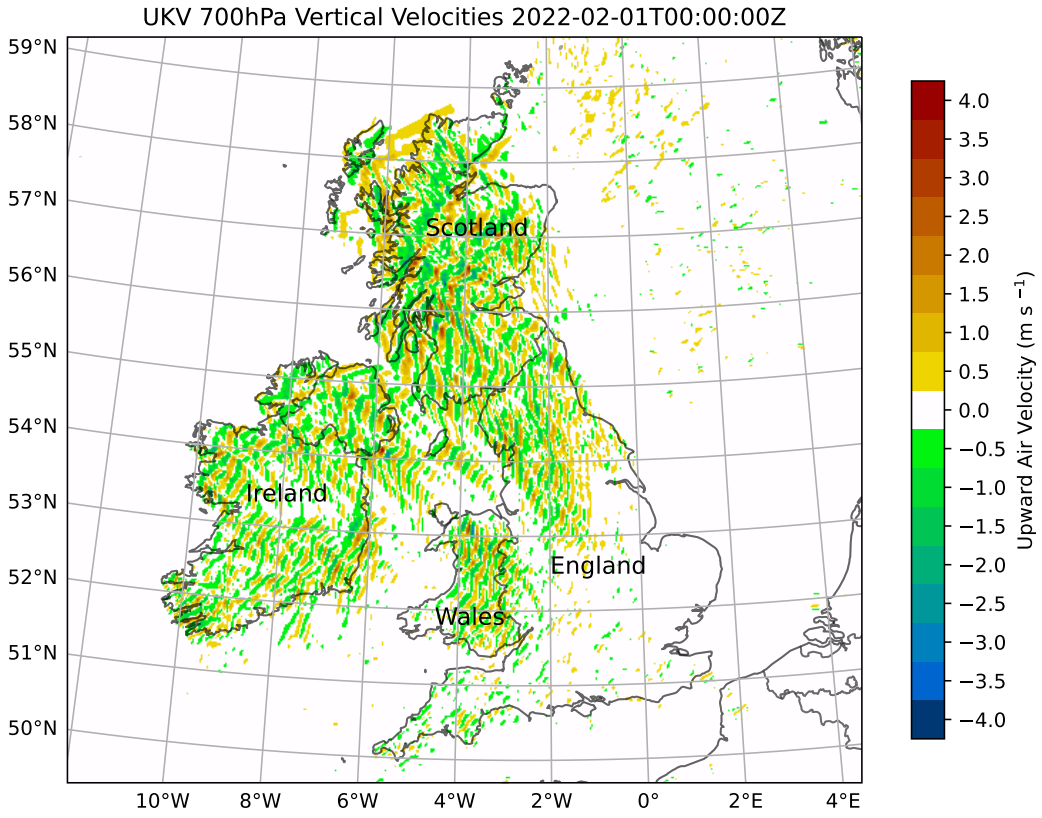


Figure 2.2: Example of the UKV 700 hPa vertical velocity analysis data, at 00Z on 1 February 2022. Several lee wave patterns over Ireland, Wales, northern England and Scotland can be discerned by eye. These regions are labelled to aid the reader unfamiliar with the geography.

2.1.3 Lee wave hazards

Lee waves with a sufficiently large amplitude can induce strong downslope winds from topography and flow separation, where air close to the surface is transported aloft (Vosper et al., 2006). The flow separation causes low-level turbulent vortices near the surface. These regions of strong turbulence, where the flow at low levels may be reversed compared to the background flow, are called rotors (Doyle and Durran, 2002), where the boundary layer separation is dependent on the adverse pressure gradients induced by the lee waves (Kühnlein et al., 2013).

In theory, rotors are classified into two types: type 1 rotors which are regions of turbulence that form under the crests of lee waves (a simple version is shown in Figure 2.1); and type 2 rotors which resemble hydraulic jumps (Hertenstein and Kuettner, 2005). Figure 2.3 shows schematics of streamlines for these two types of rotors. The type 1 rotors in (a) that have formed under the crest of the wave are much more ordered than the complex pattern of turbulent eddies that have formed in the type 2 rotors in (b). Observations of rotor activity in the Falkland Islands by Mobbs et al. (2005) suggested that a strong temperature inversion near the summit of the mountains was necessary for the formation of rotors in the Falklands, but this is not a requirement for rotor activity in Britain and Ireland (Sheridan et al., 2007).

Udina et al. (2020) observed a rotor event in the Pyrenees during a field campaign using a Lidar and UHF wind profiler. They noted that the strongest turbulence was in a region on the upwind side of a lee wave crest. In addition, they noted that rotors are characterised by complex 3D flow interactions, and the 2D structures seen in literature were most likely over-simplistic. The T-REX campaign came to similar conclusions: Strauss et al. (2016) looked at observations of lee waves and rotor-like events in a deep valley in California. They found that pressure perturbations from high amplitude lee waves were only one factor in lee wave formation, with other effects such as dynamical and thermal processes also governing rotor and turbulence formation within the valley. Also as part of T-REX, Cohn et al. (2011) used wind profilers to observe rotor effects in the east of the Sierra Nevada in California. They observed small regions of strong turbulence, which they called subrotors, around the crest of lee waves, in agreement with earlier modelling studies (such as Doyle and Durran, 2002).

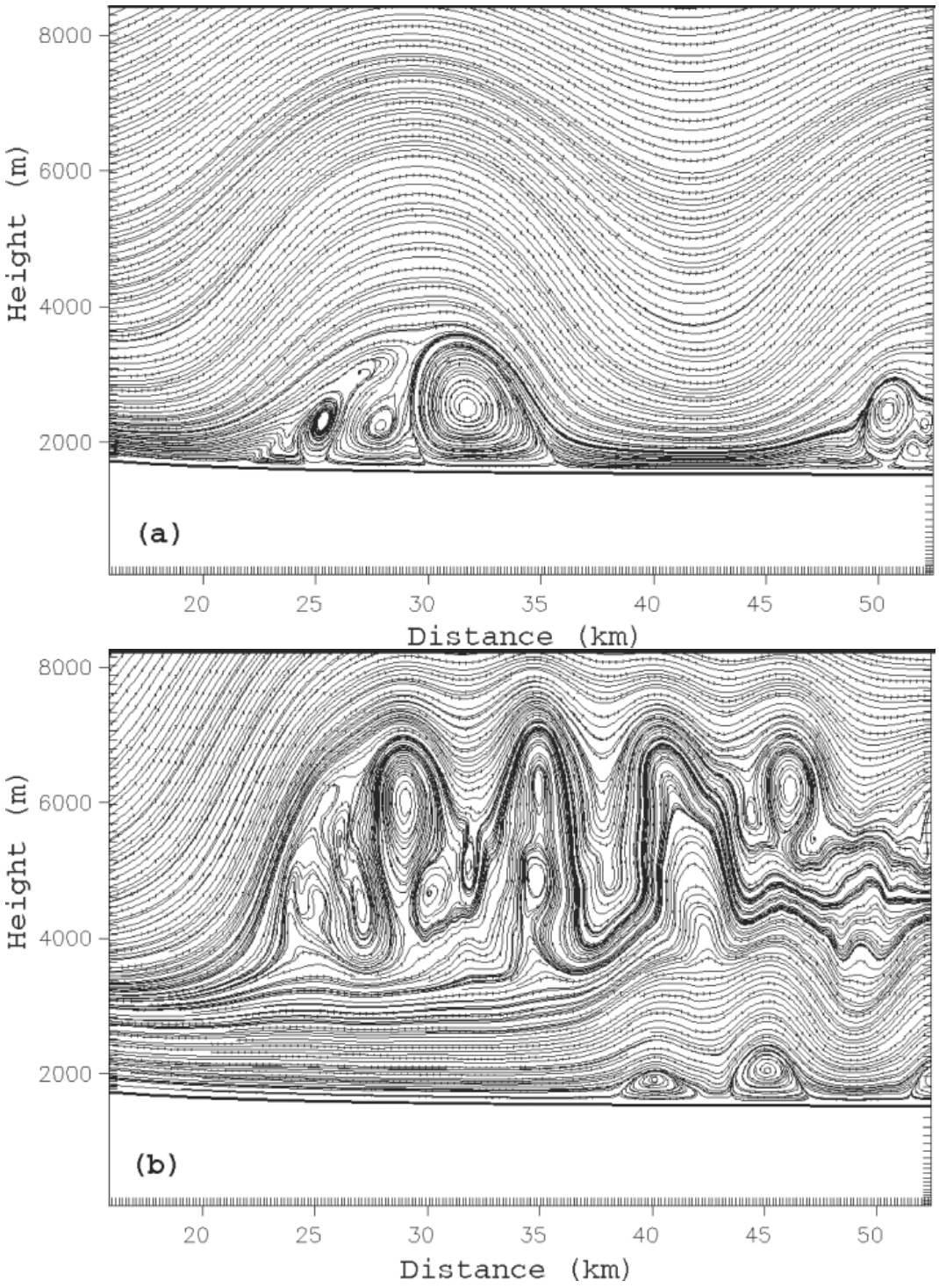


Figure 2.3: Schematic, reproduced from Figure 8 of Hertenstein and Kuettner (2005), showing streamlines for type 1 (a) and type 2 rotors (b).

Gusts associated with lee waves can be problematic for road transport, and Vosper et al. (2013) noted that a number of wind-related incidents to high-sided vehicles have occurred on the lee side of the Pennine hills in Northern England. The turbulence that caused the crash of a commercial aeroplane over Mount Fuji in 1966 was attributed to mountain waves (International Civil Aviation Organization, 1968). Turbulence and large wind shears associated with rotors on the lee side of a mountain are particularly hazardous for aircraft, especially during take off and landing (Darby and Poulos, 2006). During an investigation of lee waves over the Sierra Nevada, Holmboe and Klieforth (1957) discuss the turbulence experienced by gliders during a field campaign: one glider observed horizontal wind gusts in excess of $90\,\mathrm{ft\,s^{-1}}$ ($\sim 27\,\mathrm{m\,s^{-1}}$), while another was destroyed by gusts estimated at $120\,\mathrm{ft\,s^{-1}}$ ($\sim 34\,\mathrm{m\,s^{-1}}$). Pilots need to be aware of lee waves and rotors: Ágústsson and Ólafsson (2014) discussed the severe turbulence experienced by an aircraft near Iceland in 2008 and the occurrence of strong turbulence hazardous to air traffic at Mount Pleasant Airfield in the Falkland Islands motivated the rotor study by Mobbs et al. (2005), as well as being relevant to airfields in Britain, such as those on the lee side of the Pennines. Hence, forecasting centres are interested in identifying and communicating these risks accurately.

2.1.4 Forecasting lee waves and their effects

Lee waves can sometimes be observed during daylight hours in visual satellite imagery, with characteristic striped cloud patterns, caused by condensation of water vapour in air that has risen and cooled in the peak of the wave. When clouds are not present, or at night, visual satellite imagery cannot be used to observe lee waves. Water-vapour sensitive satellite data have also been used to detect gravity waves (for example by Vicari et al., 2024). In addition to using satellite imagery to detect the presence of gravity waves, lee waves can also be identified in the output data from NWP models run at a sufficiently high spatial resolution to resolve the waves.

Weather and climate phenomena are linked by the same physical processes, and the MetUM suite of models run by the United Kingdom (UK) Met Office (herein the "Met Office") consist of a nested suite of model configurations to forecast weather and climate regionally and globally to account for processes on different scales (Brown et al., 2012). This way, the physics of the different models, although they

contain different configurations and resolutions, are kept similar to each other.

The high resolution limited area deterministic NWP forecasting model configuration, nested within the global model in the MetUM is UKV, which has a spatial resolution of 1.5 km in mid-latitudes over Britain and Ireland (Tang et al., 2013; Bush et al., 2023). The UKV model configuration has resolved lee waves since the dynamical core was upgraded in 2015. The previous dynamical core used a greater amount of off-centring, which meant that over the operational time step, gravity wave motion would be damped by the time integration scheme (Shutts and Vosper, 2011). The current dynamical core, Even Newer Dynamics for General atmospheric modelling of the environment (ENDGame) has better numerical stability than its predecessor, meaning that gravity wave motions are no longer damped in simulations with operational time steps (Elvidge et al., 2017). The ability of ENDGame to resolve orographic gravity waves was tested by Wood et al. (2014), who found that the simulated gravity waves broadly agreed with other modelling groups within the Dynamical Core Model Intercomparison Project. More details about model setup are given in Section 3.1.

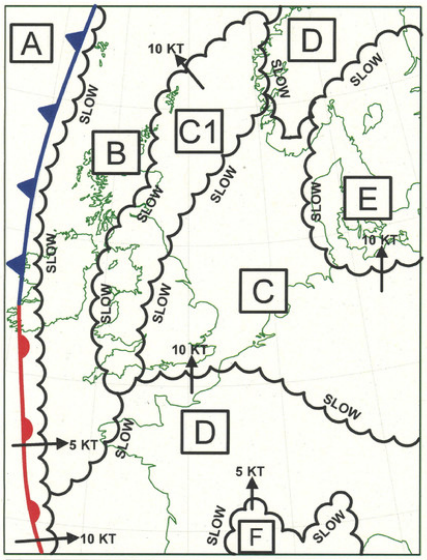
Lee waves are visible in the UKV model output from a range of fields, however vertical velocities just above the height of the orography are particularly helpful as the background values of the vertical velocities of atmospheric motion not associated with gravity waves are typically small and so the wave signal is clearer. Figure 2.2 shows an example of UKV vertical velocity model output on the 700 hPa surface over Britain and Ireland (well above the height of the orography) where lee waves have been resolved. The striped vertical velocity pattern associated with lee waves, and their dominant orientation, amplitude and wavelength can be seen by eye, but there is no operational method to automatically retrieve these characteristics from the NWP output for further use (such as generating risk maps for operational meteorologists based on modelled lee wave amplitudes).

The UKV model configuration resolves lee waves in good agreement with observations from aircraft campaigns and surface observations. Sheridan et al. (2017) note that lee waves predicted by the UKV model agree “fairly closely” with those observed by an aircraft over the Grampian mountains in Scotland, and surface observations of 10 m horizontal wind variability was observed on similar timescales and amplitudes to the UKV model fields. However, the authors note that the sparse

nature of surface observations make validating modelled features difficult, and that other complex interactions not resolved by the NWP model could have been occurring at the surface observation sites, as the strongest effects of lee wave activity do not typically occur at the surface (and so are not observed by surface meteorological observation stations), but aloft.

Forecasting of lee wave effects at the Met Office is mainly concerned with the likelihood of rotor activity, and diagnostics are used to convey this information to forecasters, who may be, for example, advising aviation personnel on the safety of flying. Output from UKV is now used for operational lee wave forecasting at the Met Office, combining UKV output with the diagnostic tools from its predecessor, the 3DVOM model (Vosper, 2003; Vosper et al., 2013). As noted above, UKV resolves lee wave activity well, and more realistically than 3DVOM did (Sheridan et al., 2017). 3DVOM was used to produce lee wave forecasts for multiple mountainous areas (e.g. Scotland, the north of England, Wales, Dartmoor and northern Ireland) by solving a dry, linear approximation for the equations of motion forced by simple boundary conditions. Since UKV resolves lee waves realistically, model data from UKV is suitable for an investigation into lee wave activity over Britain and Ireland. In addition to forecasting rotor activity, gravity waves are of interest to forecasting agencies for momentum transport, and gravity wave drag (Vosper et al., 2020). Since gravity waves are typically too fine-scale to be captured at the horizontal spatial resolution of typical global NWP models, their effects need to be parametrised within global models. A climatology of lee waves could help parametrise the effects of lee waves in global models by improving the understanding the relationship between lee wave events, particularly high amplitude waves, and other meteorological variables.

Before numerical models that can resolve lee waves were used operationally, there were methods used to identify the risk of potentially hazardous effects associated with lee wave generation, mentioned by Shutts (1997) and described in detail in the then Forecasters' Reference Book (Met Office, 1997). For example, one earlier method describes unidirectional winds which get stronger with increasing height and a higher static stability in the bottom 3 km of the atmosphere. These could be diagnosed from data from radiosonde ascents, or the output of numerical models. These rough rules of thumb above have been superseded: firstly, by numerical linear models such as 3DVOM, and now by lee wave activity being resolved within the



Forecast Weather below 10000 FT

Met Office Valid 260800 to 261700 Z JAN 17 Fronts/zones valid at 261200 Z

AREA	SURFACE VIS AND WX	CLOUD	0 C
A	15 KM NIL/-RA OCNL 7 KM RA/-RADZ ISOL 3000 M RADZ/+RA MTW MAX VSP 800 FPM AT 090 ... OCNL HILL FG	BKN/OVC AC AS ... 080 / XXX WDSPR BKN/OVC CU SC ... ... 015-025 / 060-080 WDSPR SCT/BKN ST ... ... 008-012 / 015 (LCA BASE 004 +RA)	050-070
B	20 KM NIL ISOL 3000 M DZ S MAINLY UPSLOPES ISOL 2000 M -SN MON IRELAND MTW MAX VSP 600 FPM AT 050 ... (OCNL ...) ... ISOL HILL FG	SCT/BKN (LCA FEW LEE MON) CU SC ... ... 015-025 / 030-040 ISOL SCT/BKN (NIL LEE MON) ST ... ... 008-012 / 015	060-080 OCNL ... ... SUB- ... ... ZERO 015-040
C	12 KM NIL WDSPR 6 KM HZ ISOL 3000 M SG/-DZ MAINLY UPSLOPES ISOL (OCNL NE ENGLAND TL 12 Z) ... ... 2000 M BR ISOL 200 M FG/FZFG CONTINENT ... ... (AND NE ENGLAND TL 11 Z) MTW MAX VSP 500 FPM AT 040 **C1** ... (ISOL ... N) **C1** ... OCNL HILL FG	BKN/OVC (LCA SCT LEE MON) ST SC ... ... 010-020 / 025-035 OCNL (ISOL LEE MON) SCT/BKN ST ... ... 004-008 / 010 (BASE 000 FG/FZFG)	060-080 SUB-ZERO ... ... 010-035 ... ... BUT ... ... 000-035 ... ... LAN
D	15 KM NIL OCNL (WDSPR S) 6 KM HZ ISOL 200 M FZFG CONTINENT ISOL ... SW ISOL HILL FG	ISOL (OCNL SEA COT) SCT/BKN CU SC ... ... 015 / 025-030 ISOL SCT/BKN ST 008-012 / 015 ... ... (BASE 000 FZFG)	030-050 S 060-080 N OCNL ... ... SUB-ZERO... ... 000-030
E	10 KM NIL WDSPR (OCNL FM 12 Z) 3000 M BR ISOL 2000 M -DZ/-FZDZ/SG AREAS (ISOL FM 12 Z) 200 M FG/FZFG WDSPR HILL FG	SCT/BKN SC AC ... 060 / XXX FAR N BKN/OVC ST SC ... 002-006 / ... ... 020-040 (BASE 000 FG/FZFG)	060-080 SUB-ZERO 000-030
F	20 KM NIL OCNL 4000 M HZ/SHRA ISOL 2000 M +TSRA W ISOL 200 M +SHSN MON OCNL ... (ISOL ... FM 15 Z) W ISOL HILL FG	BKN/OVC CU AC ... 020-040 / XXX ISOL EMBD CB 015-025 / XXX W ISOL SCT/BKN ST 005-010 / 015	020-040

Figure 2.4: Example LOLAN chart from the Met Office (reproduced from Figure 11 of Morcrette et al., 2019), describing the weather relevant to aviation over North-western Europe. Mountain wave activity is shown as MTW, with maximum vertical velocities shown in feet per minute.

high resolution convection-permitting non-linear NWP model configurations.

Operational meteorologists at the Met Office produce a Low Level Aviation (LOLAN) chart describing the weather forecast relevant to aviation, for Britain and Ireland below 10 000 ft ($\sim$ 3000 m) in altitude. These charts show regions of forecasted mountain wave activity and describe the amount of turbulence expected, among other features. An example LOLAN chart can be seen in Figure 11 of Morcrette et al. (2019), and reproduced here as Figure 2.4. These charts are produced manually by operational meteorologists and synthesise a lot of information into one document. For example, in region A in Figure 2.4, the mountain waves (MTW) are forecasted to have a maximum vertical velocity (MAX VSP) of 800 ft min^{-1} ($\sim$ 4 m s^{-1}) at 9000 ft (090 on the chart; $\sim$ 2700 m) above sea level, in addition to the other information regarding visibility, cloud, weather and freezing level.

2.1.5 Spectral techniques for identifying gravity waves

Typically, spectral techniques are used to detect and measure gravity wave characteristics, such as the Fourier transform, wavelet transform (e.g. Hindley et al., 2015) or the Stockwell transform (S-transform, e.g. Stockwell et al., 1996; Hindley et al., 2019). The S-transform is a spectral analysis technique used to provide localised, spatial, spectral information of waves, given input data (so time-frequency or distance-wavelength, depending on the dimensions of the input data), and has been used extensively for retrieving gravity wave characteristics (though not necessarily lee waves, until this work). One such application was an analysis of stratospheric gravity waves over South Georgia using Atmospheric Infrared Sounder (AIRS) satellite observations and MetUM models (Hindley et al., 2021). Another was the study of stratospheric gravity waves in the Andes, using the S-transform to find gravity waves with large momentum fluxes (Wright et al., 2017). The S-transform was used by Alexander et al. (2008) to produce global gravity wave momentum flux information from satellite observed temperature perturbations.

The S-transform is similar to a continuous wavelet transform, but extends it because the magnitudes of the complex coefficients of the S-transform are related to the magnitudes of the underlying amplitudes of the gravity waves (Alexander et al., 2008). This is not the case for amplitudes derived using the continuous wavelet transform (Hindley et al., 2016). The two-dimensional S-transform used for the characterisation of stratospheric gravity waves in AIRS satellite data described by Hindley et al. (2016) was later extended to an n-dimensional S-transform (Hindley et al., 2019). More details on the S-transform implemented in this thesis are given in Section 4.2.6.

2.2 Machine learning

Machine learning (ML) is a blanket term that refers to computer programs that learn from data and apply that learning to complete tasks, and maximising the ability of the programs to complete these tasks using statistics or other means. The development of machine learning techniques means that automated analysis of large data sets (such as NWP model output) can be performed efficiently. One example

of ML is deep learning, such as the use of deep (i.e. the networks contain many layers) convolutional neural networks which extract patterns from large quantities of data through learning feature representation of data via convolutions and non-linear transforms (Gu et al., 2018). This is particularly useful for two-dimensional data such as images, or NWP model output on a pressure surface.

ML models are increasingly being used for forecasting and investigating climatological trends, as well as in environmental science more generally. For example, Böhm et al. (2021) produced a climatology of fog occurrence over the Atacama desert using a neural network trained on satellite brightness temperatures. Weather fronts can be identified using machine learning techniques and their climatology and impact on precipitation events can be detected and evaluated (e.g. Niebler et al., 2022; Justin et al., 2023).

Nowcasting is forecasting over a short length of time, and is commonly used to forecast precipitation on short timescales (< 6 hours), by predicting where precipitation will occur based on the current state. Typically, optical flow approaches are used which approximate the motion of features within a series of consecutive images (for example radar retrievals as done by Bowler et al., 2004). Optical flow methods assume that the patterns in the data do not change in intensity over time, and only account for motion (Fortun et al., 2015). Machine learning methods have also been used for nowcasting, such as by Shi et al. (2015), who used ANNs to nowcast precipitation from radar retrievals in Hong Kong. Machine learning techniques have also been used to nowcast lightning occurrence from air pressure, temperature, humdity and wind speed, with skill for 30 minute lead times (Mostajabi et al., 2019). Entire NWP models have been replicated using machine learning models, and this is detailed further in Section 2.2.3.

There are several methods of machine learning: supervised, unsupervised and reinforcement learning. In supervised learning, a model is trained to extract patterns from a data set of labelled data and complete some task (for example classifying an image as a dog or a cat). Then, the model's ability to perform this task is evaluated on an unseen (by the model) test data set (Reichstein et al., 2019). This ensures that the trained model has been generalised to the population data rather than overfitting to the sample that it was trained on. Unsupervised learning does not require labelled data and is more focused on pattern extraction: for example, clustering a

data set (e.g. Section 2.2.1) into classes based on similarities between the data. In reinforcement learning, models must discover optimal outputs from being rewarded and penalised by experimentation, rather than being given explicit labelled training data (Bishop, 2006).

Three different categories of ML are explored in this thesis, with most of the focus being on deep learning, an example of which is using convolutional neural networks (CNNs: Section 2.2.3). The other types of machine learning models that are explored here are clustering and random forests.

2.2.1 Clustering

Clustering techniques involve grouping similar data based on their features. This is useful when there is unlabelled data. k-means clustering (e.g. Steinley, 2006) involves clustering data into k groups of similar features, where the Euclidean distance between a member of a cluster and the centre of its cluster is at least as small as the distance between that member and the centre of another cluster. The `sklearn` Python library (Pedregosa et al., 2012) employs several algorithms to perform k-means clustering, as well as other options to approximate k-means clustering in a quicker time, involving taking different sized batches of data. Clustering is used in atmospheric science, typically for grouping different weather conditions into regimes. The weather patterns produced by Neal et al. (2016), introduced later in Section 2.4 and Section 3.3, used k-means clustering to group patterns of mean sea level pressure (MSLP) into similar synoptic regimes. k-means clustering has also been used for analysing pollution levels, as reviewed by Govender and Sivakumar (2020).

2.2.2 Random forests

Random forests are an ensemble of decision trees used to characterise data based on features within the data (Breiman, 2001). Decision trees can be used as data classification or regression methods where some input data are classified based on a sequence of rules, where each node within each tree is split recursively until the members within each node are given the same value, whether that be a class or a value on a continuous interval (Kotsiantis, 2013). How these rules are created depends on the implementation, with many algorithms proposed on how best to

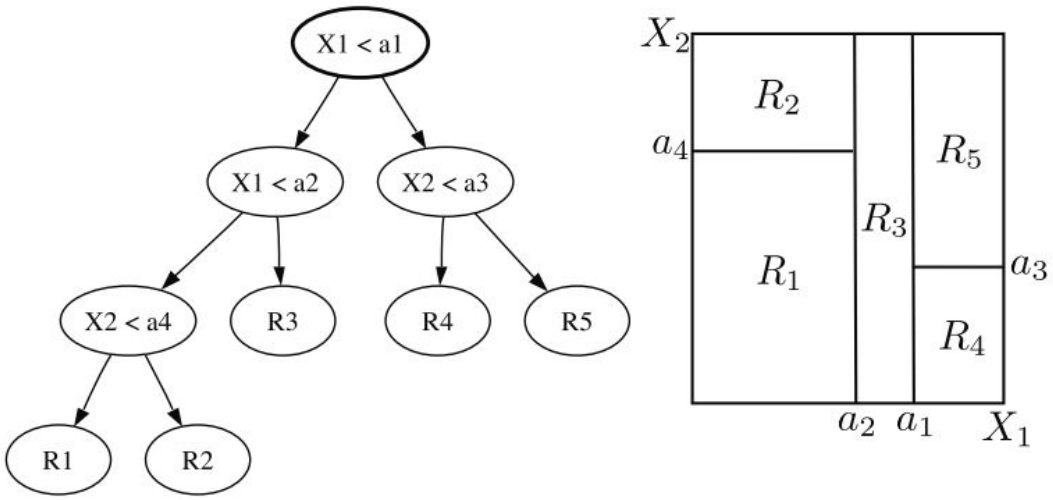


Figure 2.5: Example decision tree for classification (left), taken from Figure 9.2 of Mohri et al. (2018). At each decision node, whether the statement in the node is true (arrows pointing left) or false (arrows pointing right) dictates the flow within the tree. The data set described by X_1 and X_2 has been characterised into five classes $R_1, \ldots, R_5$ (right) by this tree.

split data at nodes within decision trees. Breiman (2001) used random splitting within trees as part of random forest creation, while another technique involves growing trees by splitting each node based on the best classification, from testing each input variable in turn (Bishop, 2006). Figure 2.5 shows an example decision tree for a simple data set, where data described by X_1 and X_2 has been characterised into five classes, $R_1, \ldots, R_5$ by the rules within the decision tree.

Random forests train multiple decision trees on a random subset of the training data, and the output of each decision tree is aggregated into the output of the random forest (typically, this is averaging each tree's probabilistic output as done in Pedregosa et al., 2012). A schematic of a simple random forest is shown in Figure 2.6. Random forests combine multiple (often 100 or more) decision trees, each trained on a subset of the overall training data (bagging), and the combination of the decorrelated trees to form the forest results in a powerful predictive model (Grange et al., 2018). Random forests have fewer parameters to tune than deep learning models and perform well on data with small sample sizes (Biau and Scornet, 2016).

Hill et al. (2020) used a random forest to predict severe weather across the United States of America. The statistical relationships learned by the random forest was similar to known physical relationships used to predict severe weather. For example,

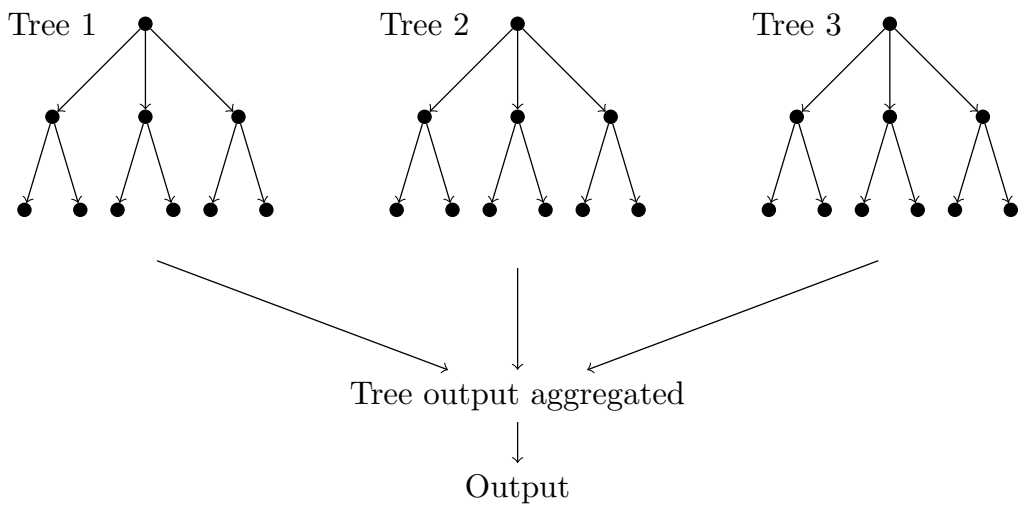


Figure 2.6: A schematic of a simple random forest, made up of three trees. The output of each tree is aggregated to make the output of the forest.

the random forest identified the convective available potential energy, convective inhibition and wind shear as being important for predicting severe weather, and these are all used operationally to diagnose severe weather (Johns and Doswell, 1992). Random forests were used by Grange et al. (2018) to explain changes in particulate matter (PM) levels at Swiss monitoring sites from meteorological conditions and time, performing well at predicting the concentration of PM_{10} at the monitoring sites in the test data. The random forest was then used to produce a time series of PM_{10}, which was investigated using partial dependence plots, finding a significant decrease in the concentration of PM_{10} at the monitoring sites.

ML techniques can be combined too: for example, Grazzini et al. (2020) used k-means clustering combined with a random forest to classify extreme precipitation events over Italy, along with a subjective method using thresholds of environmental variables. They found that the approach using k-means clustering and a random forest was more robust than the subjective method.

2.2.3 Deep learning

Deep learning, a type of machine learning, is the discipline of training deep (many layered) artificial neural networks (ANNs) to autonomously extract nonlinear relationships between large quantities of data to produce a given prediction (Reichstein et al., 2019). Deep learning models learn representations of data through matrix

multiplication and non-linear transforms, by multiplying an input by various layers (LeCun et al., 2015). The optimal weights of these layers are learned through successive passes through the model and updated using backpropagation (more in Section 2.2.4).

An ANN is an example of a deep learning model. Artificial neural networks are a simplification by a computer program of how the human brain's neurons are wired for learning patterns and were first proposed for hand-written digit recognition (LeCun et al., 1989). Since then, ANNs have been used for signal and image processing in a wide variety of fields and applications (e.g. Gu et al., 2018). A convolutional neural network (CNN) is a type of ANN. CNNs combine convolutional layers (which feature convolutions) with non-linear scaling (or "activation") functions. Convolutions combine features in a neighbourhood around a point, ensuring that a set of local features are passed down the network, leading to the model being able to learn representations of features such as edges and corners (LeCun et al., 1999). The concept of the neighbourhood around a point being able to influence the value of that point is the idea of a receptive field: with too small a receptive field, CNNs would not be able to learn features.

Figure 2.7 shows a schematic of a CNN, used for image classification (such as classifying a hand-written digit as an integer from 0–9). It shows the transformation of a $3 \times 128 \times 128$ image (on the left) through intermediate feature maps before producing a 1×10 vector. It shows how the intermediate feature maps are produced, and the neighbourhoods used on the previous feature maps to inform the next. Within this simple CNN, there are maximum-pooling layers (where the maximum value in a sliding window is taken, reducing the complexity of the feature map), along with convolutional layers and dense layers (also called fully connected layers), which perform a linear transformation of the previous feature map (O'Shea and Nash, 2015). Finally, the output can be interpreted: in this case, the model will produce a prediction of how closely the input resembles each of the integers 0–9.

One such task that deep learning can be applied to is segmentation: an image classification problem where each pixel within an image is classified as belonging to a specific predefined class. An example of a deep learning architecture for such segmentation is a U-Net (called because of their architecture shaped like the letter "U"), a type of CNN first described for the segmentation of medical imagery, for cell

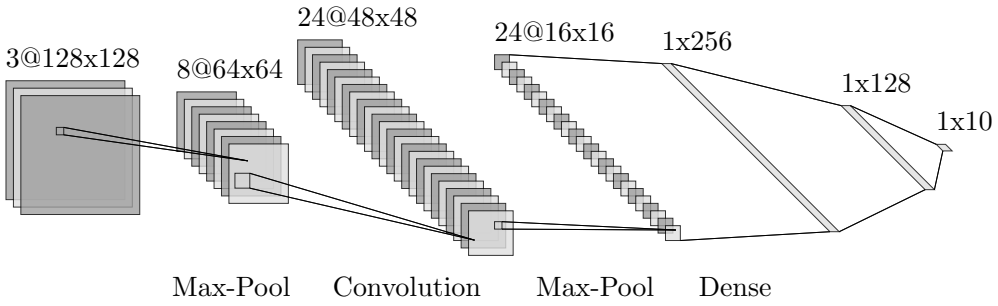


Figure 2.7: Neural network schematic for a simple CNN, created using LeNail (2019)

detection and shape measurement (Ronneberger et al., 2015). U-Nets are used for pixel-wise predictions that can take the form of segmentation masks (boolean 0/1), classification (discrete) and regression (continuous) problems. U-Nets are a type of CNN designed for pixel-wise classification since the output of the neural network can be the same height and width as the input data. This is useful for segmentation tasks such as boundary detection.

Since its introduction, the U-Net architecture has been used for image segmentation problems in a wide range of fields including within Earth Sciences. Examples include a land-cover classification from high-resolution satellite imagery over Beijing, identifying regions of buildings, water, roads, vegetation, and separate classifications for shadows, and "other" (Zhang et al., 2018), recognition of regions of clouds within photographs of the sky (Dev et al., 2019) and estimation of gravity wave momentum fluxes at 100 hPa from low-resolution winds, temperature and specific humidity data at lower levels in the atmosphere using a 29-year reanalysis data set (Matsuoka et al., 2020).

By using deep learning methods to automate tasks traditionally only capable of being undertaken by humans or complex algorithms, machine learning models can be used on unseen data to analyse and identify features, such as hazardous weather phenomena (e.g. hailstorms: Gagne et al., 2019; thunderstorms: Guastavino et al., 2022). In addition, the application of deep learning models to a large, in both size and dimensionality, data set (such as an archive of NWP model output) provides the opportunity to improve the understanding of weather phenomena by being able to analyse large data sets for cases of interest, a task which would be unfeasible by hand. Meteorological data, of which there are often large volumes on long timescales,

offer a prime data source for typically data-hungry machine learning techniques.

While this thesis concerns itself with the application of machine learning techniques to data produced by NWP models, during the past couple of years, neural networks have been used to emulate entire NWP models and it would be remiss to not mention the swathes of work occurring here. Keisler (2022) used graph neural networks to produce skilful weather forecasts. Graph neural networks are suited to performing on any data that can be constructed as a graph (a set of vertices and edges that connect them), which means that operations such as convolutions that perform well on image data are hard to construct, and do not take into account features that are connected as may happen in a graph (Wu et al., 2021). Since then, networks part-funded by large machine learning groups have also released their own ANNs for weather forecasting, for example: FourCastNet (NVIDIA: Pathak et al., 2022), PanguWeather (Huawei: Bi et al., 2023) and GraphCast (Google: Lam et al., 2023). The European Centre for Medium-Range Weather Forecasts (ECMWF) have also released their own AI version of their Integrated Forecasting System (IFS), dubbed AIFS (Chantry et al., 2024). These achieved impressive results on a number of metrics against the operational ECMWF global NWP model in a fraction of the time. For example, Google (Lam et al., 2023) claim that GraphCast "significantly outperforms" the IFS on 90% of its verification targets, and Huawei (Bi et al., 2023) claim that PanguWeather outperforms the IFS on "all tested variables". Output from all these ML weather models can be viewed on the ECMWF website (ECMWF, 2024).

Ben-Bouallegue et al. (2023) compared the IFS to the PanguWeather model, finding that the ML model performed well at forecasting 850 hPa temperature and 500 hPa geopotential height against NWP analysis, and skilful at 2m temperature compared to observations. However, they note the relative lack of skill of PanguWeather at forecasting tropical cyclone intensity and producing overly-smooth forecasts. Charlton-Perez et al. (2024) compared the skill of GraphCast, PanguWeather, FourCastNet, and the IFS, at forecasting the effects of Storm Ciarán, a high impact extratropical cyclone over the UK in November 2023. This was an extreme event: the Met Office observed the lowest mean sea level pressure on record for England in November during Storm Ciarán. While the machine learning models forecasted the storm track and MSLP contours well, they performed less

impressively than the conventional NWP at forecasting the 10 m wind speeds, and the frontal structure of the storm. Weather forecasting is arguably most important during extreme events, as people need to know what hazards may be coming their way.

While machine learning based models have been shown to perform at least on par with conventional NWP at weather forecasting, their trust is currently less than that of conventional NWP. Case studies such as the study by Charlton-Perez et al. (2024) are important to demonstrate where machine learning based forecasting can emulate extreme events, and where it cannot replicate the physics of current NWP (for example, forecasting the MSLP well but not linking this to the geostrophic winds). Trust in weather forecasting depends on having explainable predictions (for example by documentation and testing of model physics and parametrisation schemes), and the explainability of machine learning models is not currently on par with the documentation about model physics and parametrisation schemes used in conventional NWP models. However, this does not discount that machine learning models have been shown to perform well at weather forecasting.

2.2.4 Learning in artificial neural networks

An artificial neural network learns by using predictions made by the network, and updating weights in the model so that performance improves. In supervised learning (where the data are labelled), ANNs have a loss function relating the prediction made by the model and the labels (the root mean squared error between a model prediction and the labels being one possible loss function). Training aims to minimise this loss function, by updating the weights of each layer in the model through many passes through the network. Construction of the training data necessitates that the model not only performs well on the seen training data but also is generalisable, meaning that the trained model also performs well on unseen data too (Goh, 1995).

During training, after a pass through the network, the weights within each layer are updated using backpropagation (multiple applications of the Chain Rule), where the gradients between a layer's input and its weights are calculated. Starting at the output and working backwards, layer by layer until the input layer, the weight in each layer is updated, by using the calculated gradient of the loss with respect to

that weight to adjust the weights in the direction of decreasing loss (Elliott, 2001; Parr and Howard, 2018). This is to have a smaller loss by the next forward pass through the network.

2.2.5 Fitting models to data

Typically, during training, a portion of the training data is held back to check that the trained model can generalise to, and perform similarly to the main training set on, unseen data. Training can continue for a fixed number of epochs (one epoch being one pass through the network, followed by backpropagation and updating weights). However, a technique such as early stopping can be used to stop training early, aiming to produce a sufficiently trained model on the training data that also generalises well to unseen data (Prechelt, 1998). Under early stopping, training continues until the loss for the main training set starts to diverge from the loss from the validation set. At this point, the model may start to over-fit to the training set, and therefore be less good at its task on some unseen data. Determining when this divergence is occurring is not obvious all of the time since the training and validation loss do not tend to decrease linearly. One method of early stopping involves stopping training only when the validation loss has not decreased beyond the previous minimum for n successive epochs ($n = 5$ was used in the models trained in this thesis).

The learning rate specifies how far to follow the gradient of the loss function with respect to a weight, when updating a weight during training. A learning rate that is too small results in a network that takes a long time to train, and too large a learning rate may mean that the weights are altered too quickly, missing out on a possible optimal selection of weights (Yu and Chen, 1997).

It is useful to imagine the loss function as a surface of the weights, with many areas of peaks and troughs of how the loss changes in response to different weights (as in Li et al., 2017). At each point on the loss function surface, the aim is to follow the gradient of the surface down to minimise the loss with respect to that weight. Figure 2.8 shows one of these loss landscapes, where the loss with respect to the higher dimensional set of weights is displayed. Various peaks and troughs within the landscape can be seen.

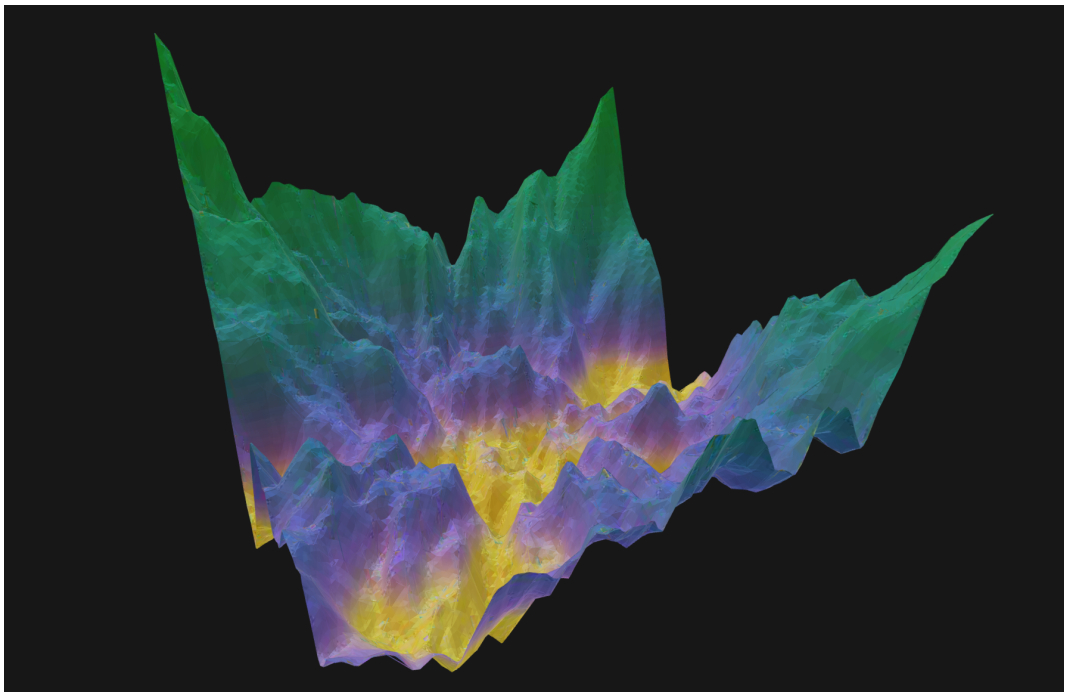


Figure 2.8: Example loss landscape for a convolutional neural network, created using Ideami (2024). The surface with various local minima and maxima is visible, and training aims to not converge on local minima.

In cases where there is limited data availability, there exist techniques to apply deep learning models to these smaller than desired data sets. Transfer learning is the process of taking a model trained on one data set, and re-training it, or part of it, on new data or to perform a different task. It uses the features learned by the model on the previous task as a starting point in the new task, which can be further refined to the new task by training (Pan and Yang, 2010), either the whole model, or certain layers – so (for example), an already well-trained feature extractor does not need to be retrained. This is useful in the case where training data for the desired task may be limited or difficult to collect, and means that models typically take less time to train because they have already learned useful feature extraction (Weiss et al., 2016).

2.2.6 Probing machine learning models

Machine learning models have a reputation as a "black-box", where probing why a model has predicted a certain outcome is not necessarily clear. Explainability is important in science, particularly if machine learning techniques are used to inform decision-making (such as issuing weather warnings, for instance). Hence, vari-

ous methods have been proposed to explain the predictions made by a ML model. "Black-box" is not necessarily a fair term to describe a whole suite of modelling techniques which have a range of interpretability depending on the technique. For example, the individual decision trees within a random forest are obtainable, so the output from a random forest can be probed, but exploring precisely why a neural network produced a certain result is less clear.

Principal Component Analysis (PCA) is one such method for exploring intermediate output from an ANN and provides a way for approximating high dimensional data in terms of lower dimensional data, while still retaining much of the information in the higher dimensional data (Wold et al., 1987). By taking n principal components of a large data set (such as some mid-level output of a machine learning model), the gist of this data can be found by reducing its dimensionality to 1 or 2 dimensions (easily visualised by eye) while minimising the amount of information lost. Then clustering techniques can be used to visualise the effects on the intermediate output from a model by changing the input.

SHapley Additive exPlanation (SHAP) is a technique that explains each input feature's contribution to the model output, by extending the concept of Shapley values (Lundberg and Lee, 2017; Shapley, 1953). Shapley values are used in game theory to estimate the contribution of different variables to a particular outcome, and SHAP expands this to use the idea for explainability in machine learning models, by approximating Shapley values for each of the inputs to a machine learning model and its contribution to the model output.

Shapley values in game theory describe coalition games played by a group of people, and how the outcome of a game is affected by whether each person participates or not (Chen et al., 2023). For a game G with n players $P = \{p_1, \ldots, p_n\}$, the power set $\mathcal{P}(P)$ (the set of all possible subsets) of the players is taken, and the outcome of the game for each element of the power set is calculated. For the player p_i, let $Q \subset \mathcal{P}(P)$ be the set of all games without p_i participating, and Q_k be the kth element of Q. Then the Shapley value S for the player p_i is calculated as (Chen et al., 2023):

$$S(p_i) = \sum_{k=1}^{|Q|} \frac{|Q_k|! \times (|P| - |Q_k| - 1)}{|P|!} \times \Big(G(Q_k \cup \{p_i\}) - G(Q_k) \Big).$$

As a simple example, for a two-player game with players a and b with outcomes 0 when nobody plays, 2 when only a plays, 1 when only b plays and 4 when both a and b play, the Shapley value for player a would be calculated as follows. $\mathcal{P}(\{a,b\}) = \Big\{\emptyset, \{a\}, \{b\}, \{a,b\}\Big\}$, so $Q = \{\emptyset, \{b\}\}$. Then

$$\begin{aligned} S_a &= \frac{|\emptyset|! \times (|P| - |\emptyset| - 1)!}{|P|!} \times \Big(G(\{a\}) - G(\emptyset)\Big) \\ &+ \frac{|\{b\}|! \times (|P| - |\{b\}| - 1)!}{|P|!} \times \Big(G(\{a,b\}) - G(\{b\})\Big) \\ &= \frac{0! \times (2 - 0 - 1)!}{2!} \times (2 - 0) + \frac{1! \times (2 - 1 - 1)!}{2!} \times (4 - 1) \\ &= \frac{1}{2} \times 2 + \frac{1}{2} \times 3 = \frac{5}{2}. \end{aligned}$$

Similarly, $S_b = \frac{3}{2}$. However, machine learning models are not games with inputs that are not necessarily binary as in this example. The method of finding SHAP values for tree-based models (which is used in this thesis), TreeSHAP, transforms the problem of a tree-based ML model and approximates the Shapley values by collapsing the sum over all subsets of features (players in the game theory example above) into calculations at each leaf node in the tree, and using the distribution of the data set used during training (or otherwise supplied) (Lundberg et al., 2020).

Using SHAP values to visualise the importance of NWP variables to the prediction, instead of other importance methods such as the permutation feature importance, means that the correlation between variables in the data is accounted for in the game theoretic approach, but does mean that SHAP values can still be spread between features that are correlated (Chen et al., 2023).

2.2.7 Alternatives and supplements to machine learning

There have been several well-publicised crowdsourced labelling campaigns of meteorological data, where members of the public ("citizen scientists") labelled records that would have been infeasible for a handful of researchers to do themselves. For example, Hawkins et al. (2019) describes a project to digitise meteorological observations made at the Ben Nevis summit observatory between 1883 and 1904. 3500 members of the public took part in the project, and it was completed in 3 months.

Several platforms exist to crowdsource labels, such as Zooniverse used for the Ben Nevis observations. However, getting a sufficient number of participants requires the project to be well-publicised, and with a simple interface and sufficiently well explained for participants to understand the task well. There are also issues with data quality – a simpler task when all the humans need to do is transcribe numbers as was the case for digitising hand-written observations, but harder if the volunteers need to make a more subjective decision (such as the segmentation of an image into classes).

Tools such as Zooniverse can also be used to generate training data for use in machine learning models – which is useful if there is a large volume of data and a limited supply of volunteers. It can also be used to develop an understanding of subjectivity and to investigate more probabilistic predictions made by machine learning models by having multiple labels for the same data.

2.3 Climate projections

As the Earth's climate changes, various scenarios have been produced which reflects the different mitigations that humans may choose to do, in order to limit (or not) climate change. The Intergovernmental Panel on Climate Change (IPCC) explain that the scenarios are based on a range of assumptions, and are not "predictions or forecasts", but a quantitative projection of how the Earth's climate may change under a range of different changes in emissions (IPCC, 2023). These different Representative Concentration Pathway (RCP) scenarios cover a broad range of possible climate change outcomes, representing a range of radiative forcings experienced in the year 2100, from a mitigation scenario: a peak of $3\,\mathrm{Wm}^{-2}$ decreasing to a forcing of $2.6\,\mathrm{Wm}^{-2}$ in 2100 (RCP 2.6), to a high emissions scenario: a forcing of $8.5\,\mathrm{Wm}^{-2}$ in 2100 (RCP 8.5: van Vuuren et al., 2011). RCP 8.5 is a scenario with little mitigation against climate change, and high levels of greenhouse gas emissions (Andrade et al., 2021), with projected global warming since 1850 in excess of 4 K and ∼1370 ppm CO_2 by the year 2100 (IPCC, 2023; Riahi et al., 2011).

Climate models are challenging to produce because the climate system consists of many interacting processes and feedbacks on many different spatial and temporal scales, which need to be captured well within the models (Laprise, 2008). Hence,

there are many uncertainties within climate models because it is not fully certain how quickly the Earth's climate will change, and is changing in response to greenhouse gas emissions (IPCC, 2023; Lemos and Rood, 2010). Multi-model projects such as the Coupled Model Intercomparison Project (CMIP) aim to understand climate changes (both past and future) through the evaluation of climate models based on performance on past data, and quantifying the factors contributing to spread in future simulations (Taylor et al., 2012; WCRP, 2024).

The United Kingdom Climate Projections (UKCP18) data set are a set of global climate projections run by the Met Office with a spatial resolution of 60 km, with nested regions of higher resolution (Lowe et al., 2018) (with the Local projections having a spatial resolution of 2.2 km). They span a 100 year period covering past and future climate, from ∼1980 to 2080. Data from the UKCP18 projections have been used to analyse potential changes due to climate change, such as an increase in winter precipitation over the UK (Kendon et al., 2020), and changes in urban heat island effects under climate change in the UK (Doger de Speville et al., 2023). The UKCP18 projections of heat extremes under climate change were compared to Coupled Model Intercomparison Project Phase 5 (CMIP5) models by Kennedy-Asser et al. (2021), who found that the UKCP18 simulations performed comparably to CMIP5, and they found that projected summer air temperatures in the UK increase quicker than the global mean.

2.4 Weather patterns

The mean sea level pressure over Western Europe have been clustered into similar weather patterns by Neal et al. (2016), by grouping similar synoptic weather conditions into 30 patterns and further into 8, more broad, patterns.

The weather patterns have been used to explore the relationship between synoptic weather and observed lightning activity by Wilkinson and Neal (2021), who found that there were seven (of 30) patterns where thunderstorms were observed more than 50% of the time during the summer, with the three most likely to result in thunderstorms associated with either slack low pressure conditions, or "Spanish plume" southerly flow conditions over Britain and Ireland, known to be associated with strong thunderstorm activity (Morris, 1986). The weather patterns have also been

used by Hendry et al. (2019) to investigate coastal flooding impacts (involving surge, river flow or both flooding types) under different weather patterns. They found that the stormiest conditions (involving strong cyclonic westerly or southwesterly flow over the UK) resulted in surge flooding events over the west coast of Britain, and cyclonic or low pressure conditions over the UK were related to high river flooding events. Richardson et al. (2018) created a climatology of precipitation and drought conditions over Britain and Ireland using the weather patterns, finding that, on the full set of 30 patterns, 6 of the patterns were related to nationwide drought, others were related to regional drought. The smaller set of 8 weather patterns were not sufficient for drought analysis as some of the patterns within the 8 were clustered from others with drastically different precipitation anomalies.

In addition to being used on current or past meteorological data, the weather patterns have been used on UKCP18 future climate data. Pope et al. (2022) used the weather patterns on the global UKCP18 data, finding that in winter, weather patterns favouring cyclonic, stormy conditions are projected to increase in occurrence over Britain and Ireland, and in summer a move to drier, less windy weather patterns than the current climate. Cotterill et al. (2023) agreed and noted that the shift to drier weather patterns in the summer, and fewer storms, could occur within the 2020s. However, they found that the signal strength differed between different models within the UKCP18 projections. Perks et al. (2023) investigated future flood risk under climate change using the UKCP18 data and the 30 weather patterns from Neal et al. (2016). Under RCP 2.6, the change in patterns from the current climatology to 2079-90 did not exceed 20%. Under RCP 8.5, the pattern was not clear: some of the weather patterns associated with storm surges were modelled to increase in frequency of occurrence but others were modelled to decrease.

Chapter 3

Data sets

This chapter describes the Met Office NWP model output data sets used in the thesis. These are:

- Output from the Met Office UKV model configuration, used for operational forecasting in the UK (further details in Section 3.1).
- Output from the Met Office UKCP18 Local data: high resolution current and future climate projections over Britain and Ireland (further details in Section 3.2).

Figure 3.1 shows the domains of the NWP model data, and some subsets of the data, used within the thesis.

3.1 UKV

The MetUM is the Met Office's current operational NWP and climate model (Met Office, 2019c). The global deterministic configuration has a horizontal spatial resolution of ∼10 km in midlatitudes, and 70 vertical levels, with a run time of up to 6 days (Met Office, 2019a). The current dynamical core of the MetUM is called Even Newer Dynamics for General atmospheric modelling of the environment (ENDGame): full details are provided by Wood et al. (2014). ENDGame solves the non-hydrostatic (that is, the hydrostatic approximation $\frac{\partial p}{\partial z} = -\rho g$ is not

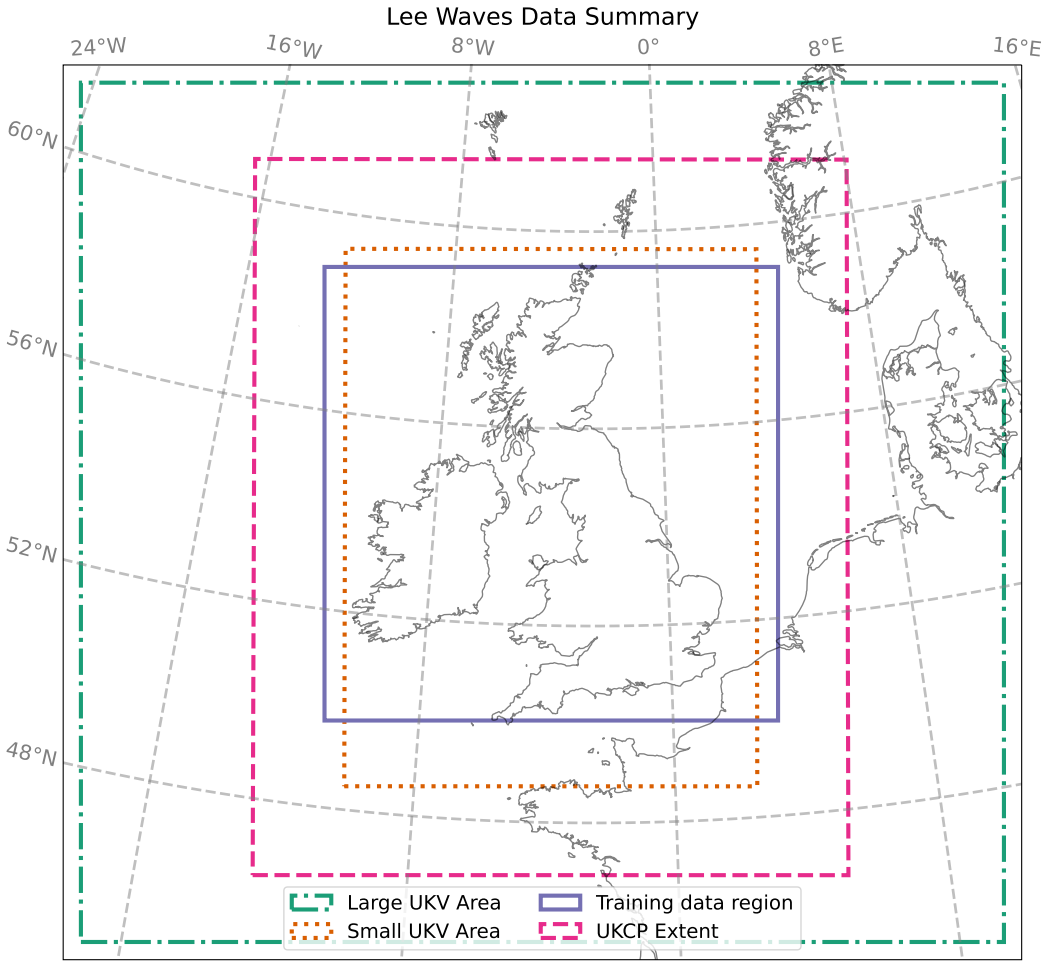


Figure 3.1: The model domains used in this study. The UKV areas shows the full extent of the UKV area and the section with a spatial resolution of 1.5 km, and the training data region that the data were cropped to when generating the lee wave training data. The UKCP18 area shows the extent of the Local simulations used for the lee wave climatology.

made: Markowski and Richardson, 2010) fully compressible deep-atmosphere equations of motion, using a semi-implicit, semi-Lagrangian formulation (Walters et al., 2019). Semi-implicit, semi-Langrangian schemes aim to provide stable models with long time steps, through re-mapping Langrangian (parcel-following) advected fields onto a regular Eulerian grid and treating the linear terms for the propagation of gravitational oscillations implicitly (Diamantakis, 2013; Staniforth and Côté, 1991; Wood et al., 2014).

Nested within the global model is the high resolution convective scale UKV configuration. UKV has a horizontal grid spacing of ∼1.5 km in the centre of the domain, and ∼4 km on the outer domain (shown as the "Small UKV area" and "Large UKV area" in Figure 3.1 respectively), with a transition zone of varying grid spacing between the two areas (Tang et al., 2013). UKV has 70 vertical levels, and is run every hour for up to 120 hours into the future, using boundary conditions from the global model (Met Office, 2019a). UKV uses the second Regional Atmosphere and Land (RAL2) configuration of physics schemes and dynamics of the atmosphere, consisting of the ENDGame dynamical core and physics schemes, such as those for radiation, microphysics and atmospheric boundary layer turbulence (Bush et al., 2023). Convection is explicitly modelled in UKV, unlike in the global model where it is parametrised (Tang et al., 2013).

UKV has permitted the generation of lee waves since the dynamical core was upgraded in 2015. The improved numerical stability in ENDGame compared to the previous dynamical core (New Dynamics) allows the use of reduced off-centring in the temporal discretization, and the numerics of the transport scheme is more accurate compared to New Dynamics (Elvidge et al., 2017). Hence, UKV more readily supports short wavelength gravity waves, without the need to use shorter time steps for wave motion to not be damped by the semi-implicit scheme (Sheridan et al., 2017; Vosper et al., 2013). Under New Dynamics, the off-centring meant that over the operational time step (15 minutes), gravity wave motion was damped unless a significantly shorter time step was used (1–2 min were used by Shutts and Vosper, 2011). The scheme used by Sheridan and Vosper (2012) using New Dynamics for simulations of lee waves in the Sierra Nevada during the T-REX campaign required a short time step (30 s or less) to maintain numerical stability so that wave motion was resolved within the high resolution model configuration.

Figure 3.1 shows the domain of UKV, and the parts of the domain with a horizontal resolution of ~ 1.5 km, and ~ 4 km (Met Office, 2019a). The data used for training the lee wave model is also displayed, mostly within the domain of the 1.5 km area. The UKV model data used in this thesis were regridded from their native variable resolution grid (1.5 km on the inner domain and 4 km on the boundaries, details in Tang et al., 2013) to a 2 km regular grid prior to archival. Although some detail may have been lost in the regridding from 1.5 km to 2 km, characteristics such as wavelength were more easily diagnosed on a regular fixed resolution grid than on a variable resolution grid. The concept of effective resolution suggests that at least 6 grid points are needed in order to represent a wave and so waves with wavelength over 12 km should be detectable (Sheridan et al., 2017). The UKV data used in this thesis is all "analysis" data, valid at $T + 0$.

While this training data region does extend slightly beyond the borders of the small (inner) UKV domain in Figure 3.1 to the east and west of Britain and Ireland, this should not affect the identification of lee waves, since the majority of lee wave activity was expected to be over the land than out to sea. Therefore, this compromise allows square training data over the entirety of Britain and Ireland, to the cost of smaller wavelength lee waves being less likely to be resolved on the boundaries of the model. However, these cases will be rare due to the edge of the domain being over the sea, well away from land. Having a smaller training data domain to fit entirely within the small UKV area would have meant either cropping out part of Britain and Ireland, which was rejected due to Scotland in the north and Dartmoor in the south both being locations of wave activity, or having non-square training data, which makes pre-proccessing the data more difficult before training a model. Hence, this resulted in the compromise of including some data that had been regridded from 4 km to 2 km on the edge of the domain within the training data set.

3.2 United Kingdom Climate Projections (UKCP18)

As introduced in Section 2.3, the United Kingdom Climate Projections (UKCP18) data set were used to investigate the prevalence and characteristics of trapped lee waves for both the present-day climate, and a future climate scenario.

Within UKCP18, the global climate model (HadGEM3-GC-3.05-PPE: Murphy et

al., 2019) provides the boundary conditions for a nested regional climate model (RCM) over Europe, which in turn provides the boundary conditions for a nested convection permitting model (CPM) using the MetUM version 10.6. The setup of the CPM is similar to, and based on, the 1.5 km resolution UKV configuration of the MetUM used operationally by the Met Office (Met Office, 2019b). The data from the CPM (herein UKCP18 Local) has a horizontal spatial resolution of 2.2 km, and 70 vertical levels (Manning et al., 2023). The spatial extent of the UKCP18 Local data is shown in Figure 3.1.

There was some evaluation of the RCM output in Murphy et al. (2019). For example, the RCM output generally agrees with observed summer precipitation, but for a case study of a heavy precipitation event, the amount of precipitation modelled was underestimated compared to National Climate Information Centre observations. Manning et al. (2023) evaluated the performance of UKCP18 data for windstorms and strong wind gust events. The UKCP18 Local data, while underestimating the frequency of occurrence of the strongest wind gusts ($> 32\,\mathrm{m\,s^{-1}}$) compared to point-based observations, performed better than the ECMWF Reanalysis (ERA)-Interim reanalysis at representing strong wind gusts.

The UKCP18 Local data is split into three time slices (TSs): TS1, covering the present-day climate from 1982–2012; TS2, covering the future climate from 2021–2041, and TS3 also covering the future climate, from 2060–2080. The simulation for TS1 used in this thesis was driven by ERA-Interim data (Dee et al., 2011), and the lee waves climatology analysis from TS1 is presented in Chapter 5. Chapter 6 looks at how lee waves might change under the high emissions future climate scenario RCP 8.5, and uses data from the UKCP18 simulations driven by a perturbed parameter ensemble (PPE) for TS2 and TS3. ERA-Interim has been superseded by reanalysis products using more observations, and with better spatial resolution (such as ERA5), but has still been used to drive analysis of small-scale phenomena. For example, Lorenz and Barstad (2016) downscaled ERA-Interim reanalysis to a 3 km grid for use in wind energy applications, finding that the additional features captured within the downscaled model were a benefit, being closer to observations than the reanalysis.

3.2.1 UKCP18 Local: Present-day climatology

The UKCP18 Local data for the present-day climatology (TS1) were hindcasts, with the global model driven by ERA-Interim data (Dee et al., 2011; Kendon et al., 2021). Using the UKCP18 Local data instead of the archive of operational UKV data meant that there was a longer (31 years vs ∼8 years) data set, and that the data were all produced with a consistent MetUM version (10.6), whereas the archive of UKV output will be produced from different MetUM versions because it is the archive of the operational model output. Therefore, using UKCP18 Local data were more appropriate for developing a climatology of lee waves than UKV, despite UKV being available on a slightly higher spatial resolution (2.2 km vs 2 km). Using UKCP18 Local data also enables comparison with the future climate scenario discussed in Section 3.2.2.

The climatology meteorological data consisted of Met Office UKCP18 Local vertical velocity data at 700 hPa. The data were available at three hourly intervals at a 2.2 km horizontal resolution from 1 January 1982 03 Coordinated Universal Time (UTC) until 30 December 2012 21 UTC inclusive (Kendon et al., 2021). The data were available through the Met Office Managed Archive Storage System (MASS), and accessed via the Natural Environment Research Council (NERC) computing facility JASMIN. Vertical velocity slices on the 700 hPa pressure surface were acquired from MASS for the full period 1982–2012.

In addition to the vertical velocity slices used by the deep learning models, other variables from the UKCP18 Local data were used to understand which conditions are important for the generation of lee waves. These variables included the horizontal wind speed and direction, virtual potential temperature and height on pressure surfaces from 925 hPa to 200 hPa. The model orography was also obtained, and used to calculate a measure of local orography variability by calculating a 2D standard deviation of the orography (with a kernel of 5×5 pixels). Data from these other variables were a subset of the available data, covering 1982–1983. More data would have taken a disproportionately long time to acquire from MASS (on the order of months), taken a long time to train models with and would be difficult to store: the data that was acquired for just these two years took up more than 500 GB of storage space.

3.2.2 UKCP18 Local: Future climatology

The UKCP18 data for the two future time slices (TS2 and TS3) provide model output for a future climate, under the high-emissions RCP 8.5 scenario (Kendon et al., 2021). The global model in the two future time slices consists of a 15 (originally 25, see below) member PPE, in addition to 13 models from the CMIP5 to make 28 members in total (Met Office, 2019b). Within the PPE, combinations of the model parameters were varied to capture as much of the parameter space as possible, while also remaining realistic: 25 combinations of 47 parameters were perturbed (Sexton et al., 2021).

The PPE was assessed in Yamazaki et al. (2021): ten of the original 25 ensemble members in the PPE were dropped because the simulated climate was too cool (due to the Atlantic meridional overturning circulation being too weak), but the remaining members were plausible simulations, despite simulating a narrower range of surface temperature changes than the authors hoped for.

12 of the remaining 15 ensemble members were then used as boundary conditions for the RCM over Europe, which in turn provides the boundary conditions for the CPM. Each of the 12 ensemble members is shown in Table 3.1. All the ensemble members used to produce the Local data were ensemble members of the HadGEM3-GC3.05 model, rather than from CMIP5 (Murphy et al., 2019).

Similarly to the data from TS1, the future climate data from UKCP18 for TS2 and TS3 were available at three hourly intervals on a 2.2 km grid. Vertical velocity slices on the 700 hPa pressure surface were obtained from MASS for each ensemble member in both time slices. In addition, the MSLP for each day were obtained for each ensemble member from the Centre for Environmental Data Analysis (CEDA) archive, as was some 750 hPa wind speeds for two weather patterns from MASS.

Due to the size of the data (12 ensemble members for 40 years, so approximately 16 times the size of the present-day climatology data), it was not feasible to retrieve and store the data to replicate the analysis in the present-day climate (TS1) for the two future-climate time slices. One year's worth of horizontal wind speed data is 28 GB in size, so the full data (40 years, and 12 ensemble members) would take up over 13 TB in storage, which was not feasible with the storage available. In

Table 3.1: Description of each UKCP18 Local ensemble member, adapted from Table D1 in Met Office (2018).

Climate Model	Global Climate Model Name	Member ID	Perturbed-Physics ID	Suite ID (2021-40)	Suite ID (2061-80)
Met Office Hadley Centre climate model	HadGEM3-GC3.05	01	r001i1p00000	mi-bb188	mi-bb189
		02			
		03			
		04	r001i1p01113	mi-bb191	mi-bb192
		05	r001i1p01554	mi-bb194	mi-bb195
		06	r001i1p01649	mi-bb197	mi-bb198
		07	r001i1p01843	mi-bb200	mi-bb201
		08	r001i1p01935	mi-bb203	mi-bb204
		09	r001i1p02123	mi-bb210	mi-bb211
		10	r001i1p02242	mi-bb215	mi-bb216
		11	r001i1p02305	mi-bb218	mi-bb219
		12	r001i1p02335	mi-bb221	mi-bb222
		13	r001i1p02491	mi-bb224	mi-bb225
		14			
		15	r001i1p02868	mi-bb206	mi-bb208

addition, the time taken to retrieve all the data from MASS (on the order of months, assuming everything was working) was not feasible. A full data set replicating all the meteorological variables used for the TS1 analysis would be needed to account for future changes, rather than a subset as used for the present-day climatology. Hence, the changes from TS1 to TS2 and TS3 are analysed using changes in mean sea level pressure and weather patterns in Chapter 6.

3.3 Synoptic weather patterns

As introduced in Section 2.4, the mean sea level pressure (MSLP) patterns over Western Europe have been clustered into similar weather patterns by Neal et al. (2016). The clustering algorithm was a simulated annealing variant of k-means clustering, as presented by Philipp et al. (2007). Simulated annealing is a clustering method that approximates the global optimal clusters, rather than converging to local optima as the other methods compared by Philipp et al. (2007) did. They also found that using simulated annealing resulted in more stable and more reproducible clusters than the other methods that they tried.

Figure 3.2 shows the MSLP for each of the 8 patterns (reduced from a larger set of

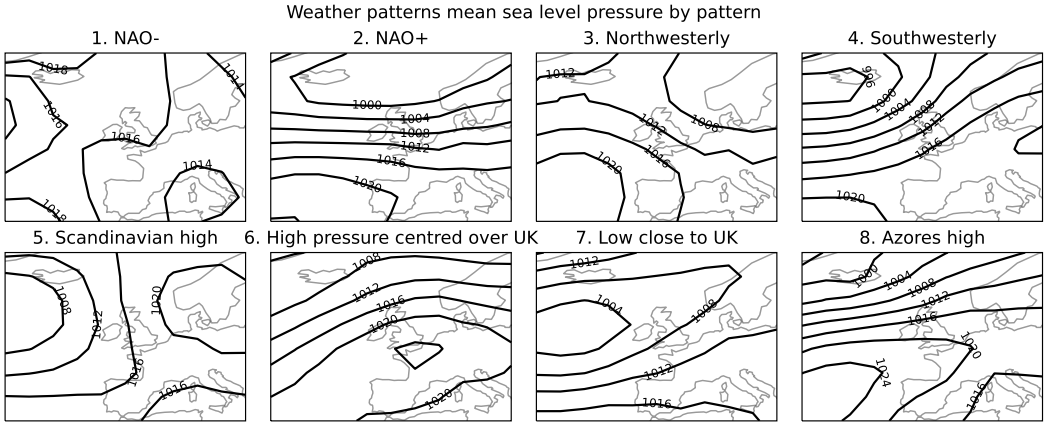


Figure 3.2: Neal et al. (2016) 8 weather patterns for climatology period, 1982–2012 inclusive. MSLP contours at 4 hPa for each pattern, except NAO-, which has contours at 2 hPa intervals.

30: it is these 8 which are used in this thesis, to make visualising the patterns more manageable). Patterns 1 and 2 deal with synoptic pressure regimes analogous to the two phases of the North Atlantic Oscillation (NAO). The negative NAO phase is associated with a flow blocking and trough pattern (Benedict et al., 2004), while a positive NAO is associated with strong westerly flow over Britain and Ireland (Washington and Palmer, 1999). The remaining six patterns all deal with locations of cyclonic and anticyclonic flow relative to northwestern Europe, and the resultant synoptic wind directions over Britain and Ireland. The weather patterns have been produced for the present-day climate data (every day for the period 1850–2020), as well as for each ensemble member in the future UKCP18 data and were available from the CEDA archive for the future climate projections (Met Office Hadley Centre, 2020). The present-day climatology weather patterns data from Wilkinson and Neal (2021) were provided by Jonathan Wilkinson.

In this thesis, the 8 weather patterns are employed to investigate the effects of different synoptic meteorological conditions on lee waves and their characteristics. They are used here to examine broad changes in synoptic meteorology and their effect on lee waves to aid visualisation of the effects on lee wave generation, especially geographically. They are then used to compare how lee waves change in the future climate projections, comparing how lee waves change within the weather patterns, and how the weather within the patterns changes, in the future climate projections.

Chapter 4

Identifying and characterising trapped lee waves using deep learning techniques

4.1 Introduction

This chapter demonstrates the training and application of a U-Net deep learning model to recognise and segment regions of lee wave activity from high-resolution NWP model output. The trained segmentation model is re-trained using synthetic data to diagnose the lee wave characteristics in the NWP model output. An overview of the methodology is given in Section 4.2; the results are presented in Section 4.3; and conclusions are given in Section 4.4.

4.2 Methodology

A U-Net is trained to segment lee waves from data from the Met Office UKV model configuration, specifically the vertical velocities on the 700 hPa pressure surface. Previous work by Sheridan et al. (2017) has shown that the characteristics and impacts of lee waves in UKV is in agreement with observations, as described in Chapter 2. Using NWP model data provides spatially dense and continuous coverage over a long time period, regardless of meteorological or daylight conditions, unlike the use of satellite imagery to observe lee waves from wave clouds. Visible or infrared satellite imagery is an unreliable indicator of lee waves because it requires cloud to

be coincident with lee wave motion, and even then higher level cloud may mask the lee wave cloud patterns. By using vertical velocities from NWP model output, most lee wave cases can be identified without biasing the sample to conditions with suitable cloud cover.

The approach used here to segment lee waves using a neural network relied on having some explicit "truth" data as a target. This is known as supervised learning. In this case, the truth data were a hand-labelled mask of each vertical velocity snapshot containing the location of wave packets which the U-Net tries to predict. While human labellers are good at differentiating between waves and other sources of vertical motion (e.g. convection), precisely identifying where the boundary of the wave packets are is difficult by eye when labelling the data. For example, in Figure 2.2 it is not a trivial task to decide for every pixel, especially those on the edge of a wave region, which pixels contain a wave and which do not. Therefore it is important to assess whether the trained neural network has learned to recognise what lee waves look like rather than learning the precise boundaries of wave regions in the hand-labelled masks.

The Met Office have archived the output of their operational UKV model since 2018, and a subset of this archive of hourly data were used to train and test a model to learn to identify and characterise the patterns of lee waves. Vertical velocity analysis data over Britain and Ireland on the 700 hPa surface from the Met Office UKV output were obtained from 1 January 2018 until 30 June 2022, and data from January, February and July 2021 was labelled by hand. The 700 hPa surface has been used before for lee wave detection from model data over the UK, such as by Vosper et al. (2013). The 700 hPa surface is above the height of the orography in the UK, and so incorrectly interpolated values of vertical velocity within the data due to the chosen pressure surface intersecting with the orography are avoided.

4.2.1 Segmentation model training data

The data were split into sets, one used for training and the other to test the skill of the trained model after training. The training and test vertical velocity slices were labelled to produce binary segmentation masks with 0s for pixels with no waves; and 1s for pixels where there was a wave. The wider forecast area was cropped to

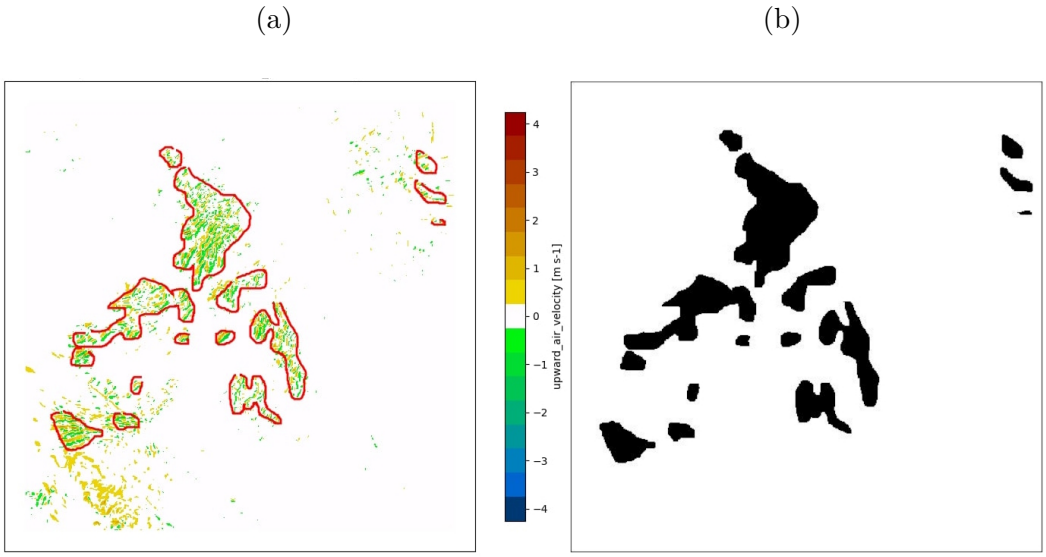


Figure 4.1: Overview of labelling process for an example of test set data. (a) Vertical Velocity data without coastline overlaid, and hand drawn regions of wave activity (in red). (b) Produced segmentation mask. white (0): no waves, black (1): waves.

512 × 512 pixels (1024 × 1024 km) to create square training data, but the cropped area still contained the entirety of Britain and Ireland (shown in Figure 3.1).

512 × 512 pixel binary segmentation masks were created using a custom Jupyter notebook utilising `matplotlib` interactive notebook functionality. A 512×512 pixel array containing zero everywhere was created. Then the human labeller drew around the regions they wished to label as a wave (regions in red in Figure 4.1 [a]). Pixels within each hand-drawn shape (closed by drawing a line from the last point to the first) were then changed to 1 from 0 in the mask array. When done, the binary mask containing 0s for regions with no waves and 1s for regions containing waves was saved (Figure 4.1[b]).

Training data comprised pairs of 335 vertical velocity cross-sections at 700 hPa from 1-18 January 2021, covering different times of day, and corresponding binary segmentation masks. Some examples containing no wave activity were excluded from the training set, so that the number of samples with and without waves was similar, in order not to encourage the model to never predict waves. Despite these measures, in the training set there was still a class imbalance where 10% of pixels were labelled as waves. Two test data sets were created, one from February 2021 and one from July 2021. These test sets contained vertical velocity cross-sections and segmentation masks from 0900 UTC each day within the respective month. The purpose of the July set was to check that the trained model had not learned to

identify waves only within winter months. 16% of pixels were labelled as waves in the February test set, and 3% of pixels were labelled as waves in the July test set. Although there could be some variation between months on the range and scale of waves, this should be unimportant because the model was being trained to recognise the pattern of waves.

4.2.2 Characteristic models training data

For wave characteristic prediction, creating a dataset from model output with pixel-wise true wave characteristics would be exceedingly difficult and perhaps impossible to do correctly. This is often a challenge in supervised learning applications. To work around this, synthetic data were created with explicitly known wavelength, orientation and amplitude, with characteristics selected to mimic the gravity waves seen by eye in the UKV data.

The synthetic data were generated by placing non-intersecting ellipses of differing sizes at random locations within an image of 512×512 pixels, using the code and methodology from Denby (2023). Each ellipse contained a Gaussian wave packet of regular cosine waves with an orientation chosen from a uniform distribution between $0°$ and $180°$ to ensure all orientations were covered equally. Wavelength was chosen at random from a chi-squared distribution with 2 degrees of freedom. Wavelengths up to 80 km were used, so the synthetic data more than spanned typical lee wave wavelengths of 5 to 35 km (American Meteorological Society, 2012), with the chi-squared distribution ensuring that there were more examples of waves in the typical range of wavelengths than there were exceeding 35 km, and that all wavelengths were positive. The amplitude decays to the edge of the wave packets to simulate waves decaying, as seen in the vertical velocity NWP model data. For each example of synthetic data, a number of wave packets with different orientations, amplitudes and wavelengths were produced. For wavelength and orientation prediction, amplitudes were kept constant at $1\,\mathrm{m\,s^{-1}}$ in all wave packets. For the synthetic data created for amplitude prediction, peak synthetic wave amplitude within a wave packet varied in the range 1 to $5\,\mathrm{m\,s^{-1}}$, chosen at random from a uniform distribution, consistent with the range of lee wave characteristics observed in the UKV data.

Pixel-wise random Gaussian noise from the normal distribution with mean 0 and

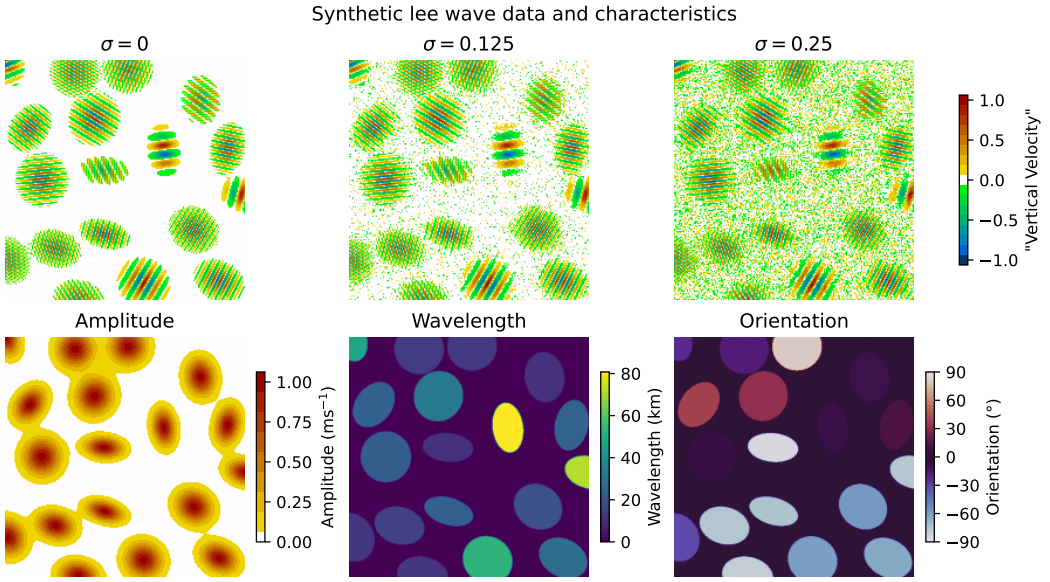


Figure 4.2: Example synthetic normalised vertical velocity input data generated for lee wave characteristic learning, at different standard deviations of normally distributed noise and output wave characteristics that the model was tasked to predict. For wavelength and orientation learning, the amplitude was set to 1. For amplitude learning, this value varied between 1 and 5.

standard deviation σ was also added to the data in order to train models that were more robust to realistic gravity waves embedded alongside other sources of vertical atmospheric motion. σ took values between 0 $\mathrm{m\,s^{-1}}$ and 1 $\mathrm{m\,s^{-1}}$. The noise array was then added to the synthetic data to produce noisy data. Several examples of the noisy data for $\sigma = 0.125\,\mathrm{m\,s^{-1}}$ & $0.25\,\mathrm{m\,s^{-1}}$, and an example without noise ($\sigma = 0\,\mathrm{m\,s^{-1}}$) are shown in Figure 4.2. This ensured a range of noisy data, from $\sigma = 0\,\mathrm{m\,s^{-1}}$ with no noise through to $\sigma = 1\,\mathrm{m\,s^{-1}}$ where the amplitude of the waves is the same as the amplitude of the noise.

It is highly likely that in reality the non-wave sources of vertical motion ("noise") in the UKV data are correlated to the waves. More complicated methods of approximating this distribution in the UKV data may produce better synthetic data, and better models, but for simplicity Gaussian noise was used here.

4.2.3 Network Architecture & Model Training

The basis of the deep learning models used here is the U-Net, a type of neural network commonly used for segmentation problems, which takes 2D data and makes pixelwise predictions. A simplified overview of the U-Net used in this study is shown

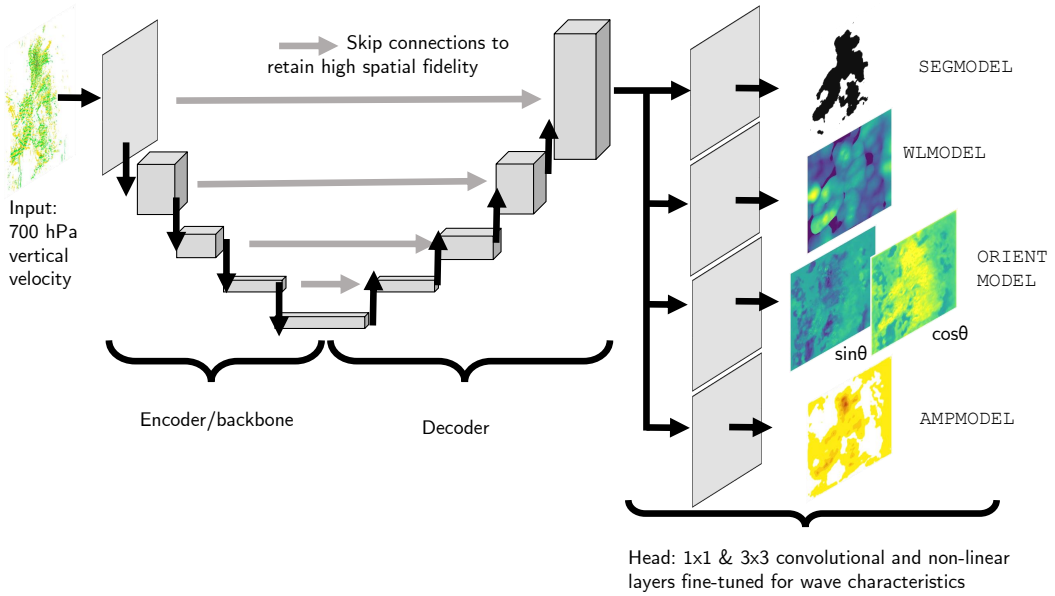


Figure 4.3: Overview of the U-Nets used. The encoder and decoder parts are shown, as well as the head of the model for making final predictions based on the features extracted.

in Figure 4.3. A U-Net consists of two main parts, an encoder and a decoder. The encoder (or backbone) can extract spatially complex patterns by coarse-graining the input data with increasing depth in the network and compositing learned features. The number of channels (the scalar values at a single spatial location) is increased through the encoder to maintain information capacity, to account for the data being spatially coarse-grained in the encoder. The decoder uses the patterns extracted by the encoder and upsamples the data while reducing the number of channels (decreasing the depth), so that a prediction can be made for each pixel. In addition, the upsampled data are combined with data from the encoder at the same level (skip-connections). These skip-connections allow the U-Net to retain high spatial fidelity by combining the up-scaled values in the decoder with more spatially dense values from the encoder. Finally, the head is where the remaining pixel-wise learnt spatial features are further manipulated through pixel-wise transforms to produce predictions per pixel.

The variation of the U-Net used here was the "Dynamic U-Net" implemented in the Python library `fastai` (Howard and Gugger, 2020). This uses a model designed to extract patterns from data (a Resnet34, pretrained on the ImageNet data set: He et al., 2016; Deng et al., 2009) as an encoder, which means that the part of the model dedicated to extracting patterns from data is already trained to do this. `Fastai` is a

wrapper for the Python deep learning library `PyTorch` (Paszke et al., 2019). `Fastai` allows an approach to deep learning that is understandable and easy to access for a user with limited experience with deep learning, yet still produces accurate deep learning models.

4.2.4 Segmentation

The segmentation model (referred to as the `SEGMODEL` from this point) takes the 2D regridded 700 hPa vertical velocity field as an input and outputs a 2D boolean mask of the same shape, containing a prediction of where gravity waves are present. Supervised learning was used to train the `SEGMODEL`, and so labelled data were needed, as described in Section 4.2.1. The labelled training set was divided randomly into a train set (80% of the data) and a validation set (20% of the data), during training. This prevented overtraining by stopping training while the trained model performed similarly on the train and validation set. The test set was only used to assess how well the trained `SEGMODEL` performed at segmenting lee waves on data not used during training.

The training data were augmented using the built in `fastai` augmentation functions including flipping, rotation (up to 360° and a probability of 0.9 of any rotation being applied) and zooming (up to 20× and probability of 0.5 of any zoom being applied) of the data. The vertical velocity data were normalised to have a mean of 0 and standard deviation of 1. Augmentation of the data minimises over-fitting of the model during training (Shorten and Khoshgoftaar, 2019). By augmenting the data, the model was exposed to waves at a range of wavelengths and orientations during training, beyond the original training data. For example, by rotating the lee waves data, the model can learn waves at a variety of angles, not just the waves in the typical southwesterly flow over the UK. Zooming should allow the model to learn to recognise waves of longer wavelengths than those available in the training data. Waves are generated through the same mechanism regardless of the orientation, so rotating the data during training should not affect learning negatively.

Cross Entropy Loss, which prefers models with a high degree of confidence in their predictions, was used as the loss function as it is well suited for boolean prediction problems such as this (Jadon, 2020), rather than other metrics such as the Jaccard

score (which is not differentiable), used later for model evaluation. To prevent overtraining, training continued until the validation loss appeared to be increasing again, using the built-in `fastai` Early Stopping callback, with an epoch window (patience) of 5 epochs. Once the `SEGMODEL` was trained, its performance was assessed using the unseen test data set, consisting of 28 examples of vertical velocity data and hand labelled lee waves from February 2021 and 31 examples from July 2021.

4.2.5 Wave Characteristics

Given that the neural network has learned to recognise lee waves during training, then it should also have learned something about the wave characteristics. By fine-tuning only the layers in the *head* (see Figure 4.3) on synthetic wave data with known characteristics, a model can be produced that predicts a wave characteristic instead of a segmentation. This is an efficient way of training networks to extract multiple characteristics, and if successful, supports the hypothesis that the original model was learning some properties of the waves.

Three copies of the trained `SEGMODEL` were taken: `WLMODEL` to predict wavelength; `ORIENTMODEL` to predict wave orientation; and `AMPMODEL` to predict wave amplitude. In each of the copies, the weights in the encoder and decoder were frozen. Only the weights of the layers in the *head* of the copies (labelled as such in Figure 4.3), consisting of non-linear scaling functions, 1×1 and 3×3 convolutional layers, were trained (fine-tuned) on the synthetic data (see Section 4.2.2) to predict the desired characteristic. In general, these layers in the *head* transform U-Net feature vectors into predictions. By freezing the weights of the spatial feature extracting *encoder* in the model, the contextual information about waves in UKV data that the `SEGMODEL` had learned was retained.

The wavelength network `WLMODEL` was trained to predict the wavelength in kilometres, while the orientation network `ORIENTMODEL` predicted the sine and cosine of a wave's orientation so the orientation could be recovered using the arctangent. By predicting the sine and cosine of the orientation (the direction of wave propagation, perpendicular to the wave fronts) rather than the wave direction in degrees, the discontinuity of angles around 0° and 360° was avoided. Orientation was predicted in the range $(-90°, 90°]$ because waves with orientation 180° apart look identical.

Lee waves are often quasi-stationary, and even for propagating waves the direction and speed cannot be calculated with single time snapshots of data. Since the wave characteristics were not available in the UKV training data, synthetic data were created to re-train these models, as outlined in Section 4.2.2. 550 examples of data were created, with explicitly known orientation and wavelength. 500 examples were used as training data (with a 80/20% train/validation split) and an additional 50 examples used as test data. In a similar vein to training the `SEGMODEL`, the altered models `WLMODEL` and `ORIENTMODEL` were trained using the Early Stopping callback, with an epoch window (patience) of 5. The loss function used was the Mean Squared Error, as these tasks are regression rather than classification.

4.2.6 The S-transform

The neural network predicted wave characteristics (wavelength, orientation, amplitude) were compared against characteristics derived using an established spectral analysis technique: the 1-D Stockwell transform (S-transform). The S-transform provides time-frequency or distance-wavelength localisation of signals present in input data and is thus ideally suited for the measurement of gravity wave packets. The S-transform therefore provides an existing spectral technique to which to compare the predictive skill of the neural networks, but note that the S-transform measurements should not necessarily be taken as "truth". The S-transform is being used here as a means of verifying that the ML models are producing wave characteristics that are realistic and reasonable in line with an existing technique for deriving characteristics of gravity waves.

The 2-D S-transform application developed by Hindley et al. (2016) and Hindley et al. (2019) is used, which provides the dominant local spectral properties (wave amplitude, wavelength, orientation) at every pixel of the input image, but then restricted to those regions recognised as waves by the `SEGMODEL`. This gives undue credit to the S-transform which does not segment waves into wave and non-wave regions as the `SEGMODEL` does. Hindley et al. (2019) describe the below equation to produce the n-dimensional S-transform S for some input n-dimensional data $h(\mathbf{x}) = h(x_1, \ldots, x_n)$, translations $\boldsymbol{\tau} = \tau_1, \ldots, \tau_n$ and spatial frequencies $\mathbf{f} = f_1, \ldots, f_n$.

$$S(\boldsymbol{\tau}, \mathbf{f}) = \int_{-\infty}^{\infty} h(\mathbf{x})\, w(\mathbf{x} - \boldsymbol{\tau}, \mathbf{f})\, e^{-i2\pi \mathbf{f}\cdot\mathbf{x})}\, \mathrm{d}\mathbf{x}$$

In this case, using two-dimensional input (x, y), this becomes

$$S(\tau_x, \tau_y, f_x, f_y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x, y)\, w(x - \tau_x, y - \tau_y, f_x, f_y)\, e^{-i2\pi(xf_x + yf_y)}\, \mathrm{d}x\, \mathrm{d}y$$

w is a function that provides spatial and spectral localisation, typically a Gaussian window (Hindley et al., 2016). One feature of the S-transform application used here is that it can be tuned to provide improved performance for waves present in a given dataset. Specifically, the analysis first computes the Fast (Discrete) Fourier Transform (FFT) of h and selects the N elements (number of frequency voices) with the largest spectral power for further localisation analysis. The larger the number of frequencies N, the higher the fidelity of the analysis but the longer the run length. For images containing simplified large-scale monochromatic waves, only small values of N are required, but for images with numerous small-scale waves with complex structures, higher N values can be useful. Secondly, a scaling parameter c can be used to tune the spectral sensitivity of the S-transform (within w). From a default of $c = 1$, increasing c improves spectral localisation at the expense of spatial localisation, while decreasing c achieves the opposite. For each of these N frequency voices, the S-transform equation above is calculated in Fourier space by multiplying the Fourier-transformed h with a Fourier-domain voice Gaussian for the nth frequency voice. Then the inverse FFT is taken to transform the output back into the original domain (Hindley et al., 2019).

The S-transform was applied to the synthetic and NWP vertical velocity data with three different numbers of frequency voices ($N = 15, 80, 150$) to determine the most appropriate value to capture all relevant waves in the data. A scaling parameter of $c = 0.25$ was used for this initial test, as used in previous studies (Wright et al., 2017; Hindley et al., 2021). Later, the value of c was adjusted to calculate an optimal value to use for further analysis. Using $N = 15$ frequency voices resulted in a cut-off in the output wavelength, with shortest wavelengths of 50 km, while frequency voices of $N = 80$ and above resulted in a cut-off around 10 km. Given that the horizontal

Table 4.1: Comparison of R^2 least squares correlation coefficient (to 3 decimal places) for wavelength, orientation and amplitude derivation using S-transforms with scaling parameter $c = 0.25$, three values for the number of frequency voices N used in the spectral analysis, and machine learned `WLMODEL`, `ORIENTMODEL`, `AMPMODEL` vs known truth on synthetic data without noise.

	Least squares linear regression coefficient R^2 (3 d.p.)			
	Wavelength	**Orientation:** sin	**Orientation:** cos	**Amplitude**
`WLMODEL`	0.997			
`ORIENTMODEL`		0.961	0.979	
`AMPMODEL`				0.997
S-transform ($N = 15$)	0.042	0.198	0.071	0.237
S-transform ($N = 80$)	0.969	0.974	0.941	0.793
S-transform ($N = 150$)	0.968	0.974	0.940	0.799

spacing of the vertical velocity data is 2 km, a cutoff of 10 km is an appropriate limit for the smallest wavelength that can be reliably measured.

To measure the ability of the ML models and S-transform to reconstruct the waves, the least-squares correlation coefficient R^2 with the ML wavelengths on the synthetic data was computed. The R^2 values for the three sets of S-transform characteristics and the machine-learned characteristics were compared against the true values and are given in Table 4.1. From these tests a value of $N = 80$ or 150 frequency voices gave similarly good correlation against the true wavelength and orientation, but 80 frequency voices was computationally cheaper and so is used in the remainder of the chapter.

However, the S-transform assumes a idealised mathematical representation of waves (such as planar, monochromatic waves) and cannot acquire a knowledge of "real world" wave characteristics where the physical scales, orientations and frequencies of waves may vary within one wave cycle, unlike with the machine learning approach. However, it does provide a means of assessing the predictions made by the ML models trained to predict wave characteristics.

4.2.7 Wavelength & Orientation model selection

The synthetic data were much simpler than the original UKV data it was attempting to replicate, as the synthetic data did not contain other sources of vertical velocity, nor superposition of waves. Since the `SEGMODEL` could detect where gravity waves are, it is feasible to suggest that machine learning models may be able to extract the salient features of gravity waves, irrespective of whether there are other physical processes creating vertical velocity variations. Originally, the synthetic data were noiseless, as it was thought that the `SEGMODEL` had retained sufficient learning from its initial training to handle the vertical velocities in the UKV data not associated with lee waves, when predicting characteristics. However, the wavelengths predicted by the model trained on noiseless data produced unrealistically long wavelengths on the UKV data. To address this, seven more wavelength models were trained on noisy data at a range of standard deviations σ in $[0.125, 0.25, 0.375, 0.5, 0.6, 0.8, 1]$. Some examples of the noisy data are shown in Figure 4.2.

The performance of all the ML models and the S-transforms were compared for different noise levels. Figure 4.4 shows the performance of the eight ML wavelength models, and three S-transforms with 80 frequency voices and $c = 0.25$, $c = 1$ and $c = 4$. The best performing machine learning model in each case was the model that was trained on the data most similar in noise level to the corresponding test set at lower values of σ. At values of $\sigma > 0.5$, the picture is less clear, with RMSEs higher than for values of $\sigma \leq 0.5$. Models trained with noise performed better on data without noise than the model trained with no noise performed on data with noise, suggesting that adding noise during training allows the model to better generalise to different levels of noise. Over a range of noise amplitudes ($\sigma = 0.125 - 0.5$), the models are fairly robust at accurately predicting the wavelength from data with different noise levels within this range, as shown by the plots in Figure 4.4. The model trained on noise $\sigma = 0.125$ has $R^2 > 0.8$ for all apart from the noisiest synthetic data. This suggests that the `WLMODEL` trained on no noise was overfitted to the training data (expecting a specific level of noise) and rapidly decreases in skill compared to models trained with some noise. The addition of noise to the training data seems to have mitigated this overfitting. While the Gaussian noise in the synthetic data is relatively simplistic compared to the correlated non-wave sources of vertical velocity in the UKV data, the trained models were all exposed to

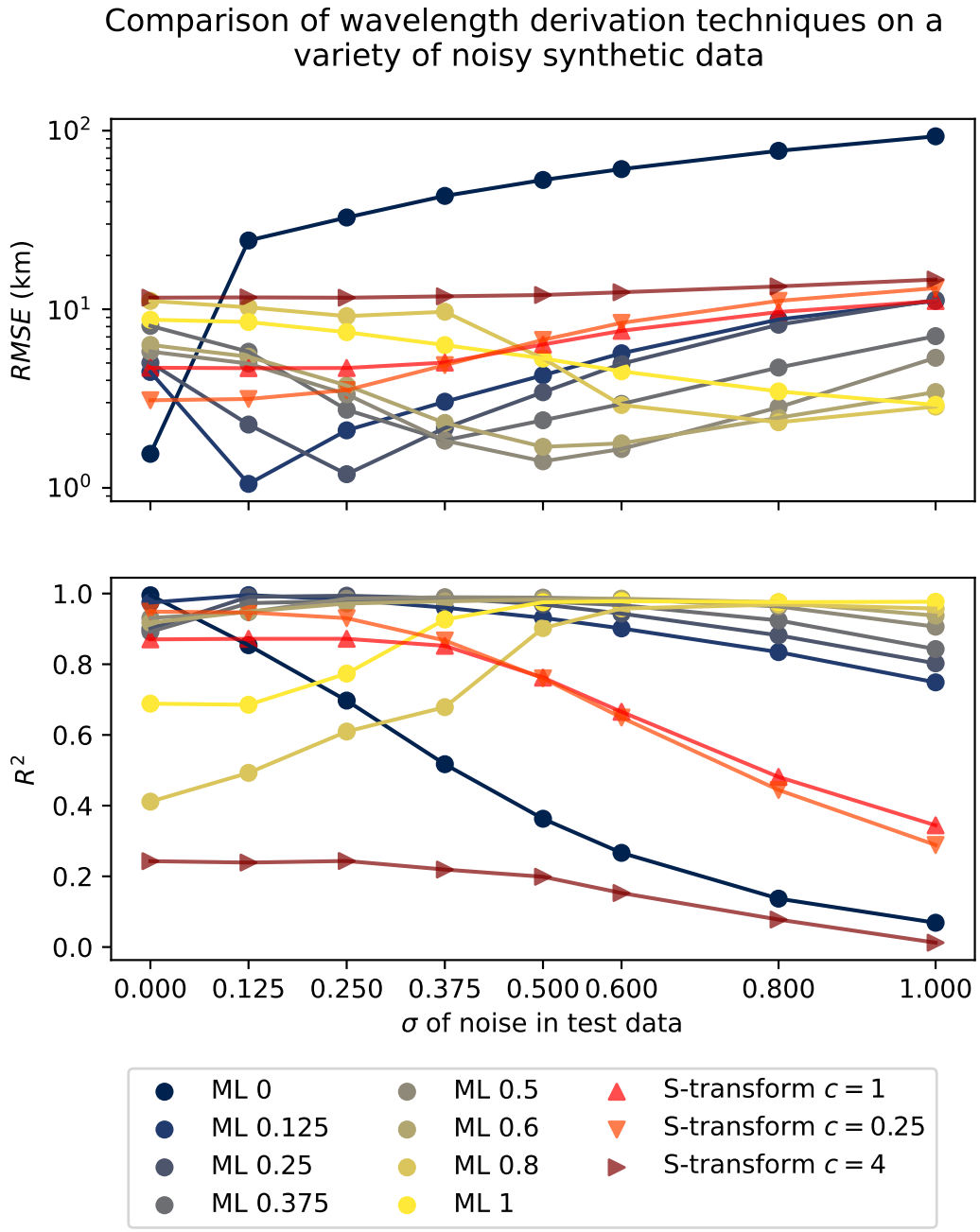


Figure 4.4: Comparison of the root mean squared error (RMSE) (smaller values better) and least squares regression coefficient R^2 (larger values better) performance of wavelength derivation techniques compared with the truth (three S-transformss with 80 frequency voices and different scaling parameters c; and eight ML models) on synthetic data with different levels of noise. Then all are applied to synthetic data with those levels of noise. The number after "ML" indicates the σ of the noise the model was trained on.

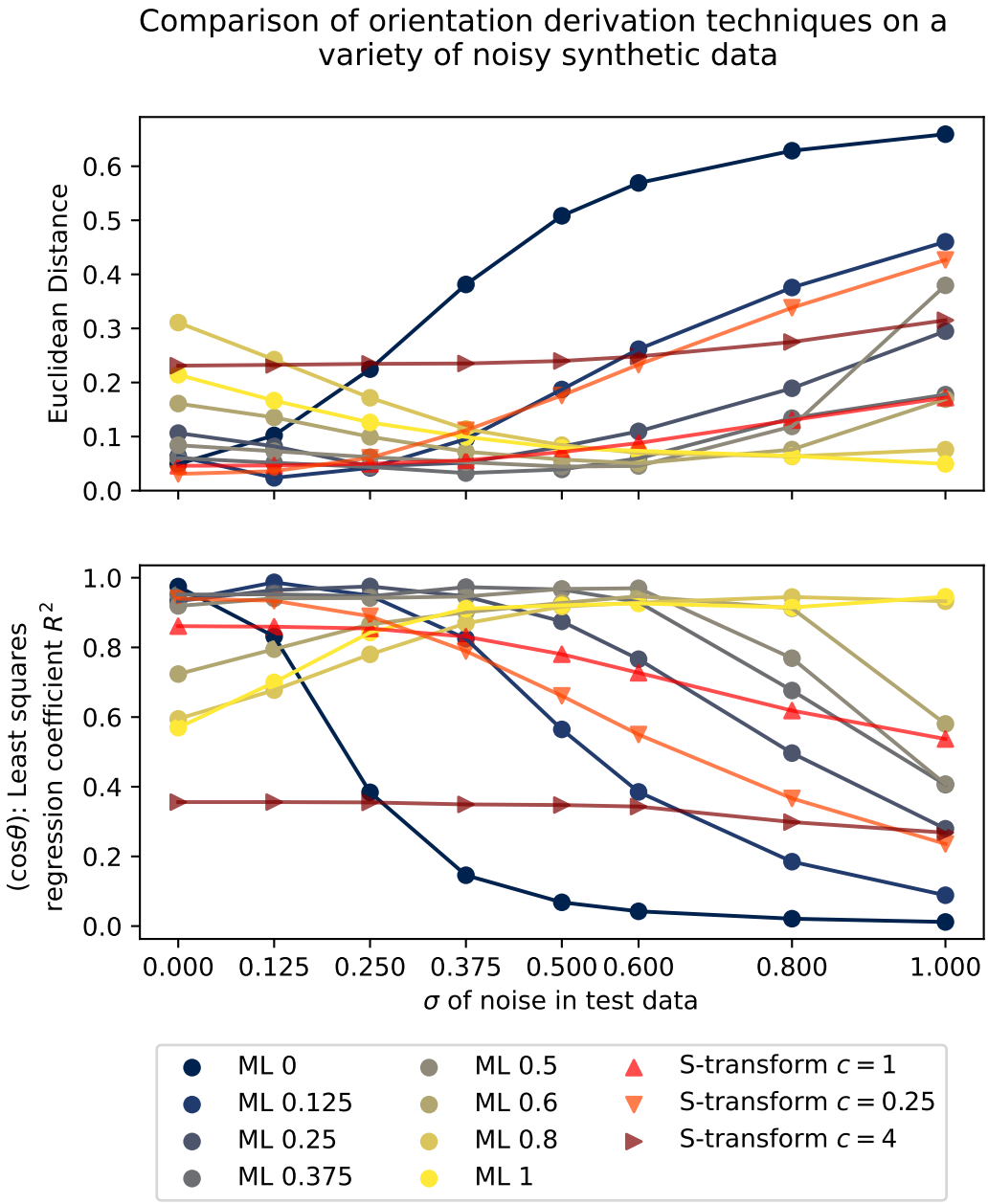


Figure 4.5: Comparison of the Euclidean distance (smaller values better) and least squares regression coefficient R^2 (larger values better) performance of orientation derivation techniques compared with the truth (both ML and S-transforms as in Figure 4.4).

the non-wave sources of vertical velocity during the training of the `SEGMODEL`.

At $\sigma < 0.5$, the S-transform wavelengths showed good correlation with the true wavelengths, with $R^2 > 0.8$. However, the S-transform results consistently had a worse least squares correlation coefficient than the best performing machine learning model, which is likely due to the limited frequency voices (and thus orientations) of using the discrete Fourier transform to calculate the S-transform (Hindley et al., 2019), which results in a slightly less "exact fit" for the input waves.

Figure 4.5 shows a comparison of the orientation derivation techniques on the synthetic data. Due to the circular nature of the data, the two methods used in Figure 4.5 to analyse the accuracy of the predictions are the euclidean distance and the least squares regression coefficient R^2 for the cosine of the orientation. The euclidean distance metric E between the angles θ and ϕ is defined as

$$E(\theta, \phi) = \sqrt{(\cos\theta - \cos\phi)^2 + (\sin\theta - \sin\phi)^2}$$

For example, a pair of angles 90° out of phase with each other would have $E = \sqrt{2}$, and a pair of angles 30° apart would have $E = 2 - \sqrt{3} \approx 0.27$. The R^2 plotted is of the cosine of the angles to avoid discontinuity around $-90°$ and 90°.

There is a smaller difference between the ML and S-transform techniques in Figure 4.5 than in Figure 4.4, with the S-transform $c = 1$ performing better than some of the ML models in the synthetic test data. The skill of an ML model for wave orientation decreases more rapidly when applied to data with other noise levels compared to the wavelength models. That said, the orientation models trained on any noise are more robust to other noise levels than the orientation model trained on noiseless data. The model trained on noisy data with $\sigma = 0.25$ has a least squares coefficient > 0.8 for data up until $\sigma = 0.6$, so is robust to a range of noise amplitudes.

From these tests, the `WLMODEL` trained on data with noise $\sigma = 0.125$ was used to predict UKV wavelengths, and the `ORIENTMODEL` trained on data with noise $\sigma = 0.25$ was used to predict UKV orientation. To select which c to use in the S-transform, local variations in wavelength with $c = 0.25$ were found to be too large while choosing $c = 1$ resulted in a spatially smoother wavelength field. Setting $c = 4$ resulted in oversmoothing that produced inaccurate wavelength estimates. Hence, $c = 1$ was

used for future comparisons in synthetic and model data.

4.2.8 Amplitude

A neural network, `AMPMODEL`, (fine-tuned in the same way as described above for wavelength and orientation) was used to extract the amplitude of the waves from the UKV data. The neural network was trained on synthetic data but with wave packets of variable amplitudes in between 1 and $5\,\mathrm{m\,s^{-1}}$. The synthetic data had a small amount of noise ($\sigma = 0.0625$) added, which resulted in a smoother amplitude model prediction over the UK, compared to a model trained on synthetic data without additional noise. The amplitude of observed waves and the synthetic data decays towards the edge of each wave packet, so the model was trained to predict this smooth envelope. On the synthetic test data, the trained amplitude model scored an R^2 of 0.999.

4.3 Results

This section presents the results of the segmentation model `SEGMODEL` against the hand labelled truth, and the results of the wave characteristics models (`WLMODEL`, `ORIENTMODEL` and `AMPMODEL`), with the wavelength, orientation and amplitude output compared against those from the S-transform. The machine learning models ran significantly faster than the S-transform on the UKV test data. For example, it took an hour and ten minutes for a standard laptop CPU to produce the S-transformed data for the 28 examples in the February UKV test set, while it took the same laptop 5.5 minutes to produce the wave mask, wavelength, orientation and amplitude for the same set, a speed up of 12.7×. Table 4.2 shows the mean and standard deviation of the time taken for the two methods to produce wave characteristics for each example of the February 2021 test data.

4.3.1 Lee wave segmentation

Figure 4.6 shows two examples of lee wave segmentation on vertical velocity data: one example of test data from February 2021 (a) and another from July 2021 (b).

Table 4.2: Comparison between the U-Net and S-transform, and the time taken for the two methods to produce wave characteristics, for each example in the UKV February 2021 test set.

Method	Mean time taken	Standard Deviation
U-Net	12.1 s	2.61 s
S-transform	150.9 s	12.3 s

These results show that the model has learned typical patterns of gravity waves during training. The `SEGMODEL` is skillful at recognising wave-like patterns, and is capable of ignoring non-wave sources of vertical velocity. This is evidenced in Figure 4.6 (a) which shows an occasion where there are large regions of lee waves in the data, which the `SEGMODEL` has recognised as waves. The area to the north of Ireland where the vertical velocity patterns look very different is likely to be convection, and not wave activity, as there is precipitation in the model associated with it. The `SEGMODEL` correctly did not classify these regions as lee waves. Figure 4.6 (b) shows an example where waves are apparent over Ireland, with smaller regions with lee wave like features elsewhere.

Two test sets were used to analyse the performance of the trained `SEGMODEL`. One was from February 2021, and one was from July 2021. This was to check that the model was able to recognise waves from throughout the year. There are typically fewer lee waves in summer months so the results from the two months are presented separately. As a reference, the output from a baseline "model" that always returned no waves everywhere (the `ZEROS` model) are presented alongside the results from `SEGMODEL`.

Four metrics of model performance on the test sets are summarised in Table 4.3, the pixel accuracy, Jaccard Score, Precision score and Recall score. The pixel accuracy is the percentage of pixels that were correctly identified by the model, compared to the hand labelled truth.

The Jaccard Score (or intersection over union) is given by

$$\text{Jaccard Score} = \left(\frac{|P_0 \cap T_0|}{|P_0 \cup T_0|} + \frac{|P_1 \cap T_1|}{|P_1 \cup T_1|} \right) \div 2$$

where P_i is the model's prediction and T_i is the hand-labelled truth for the ith class

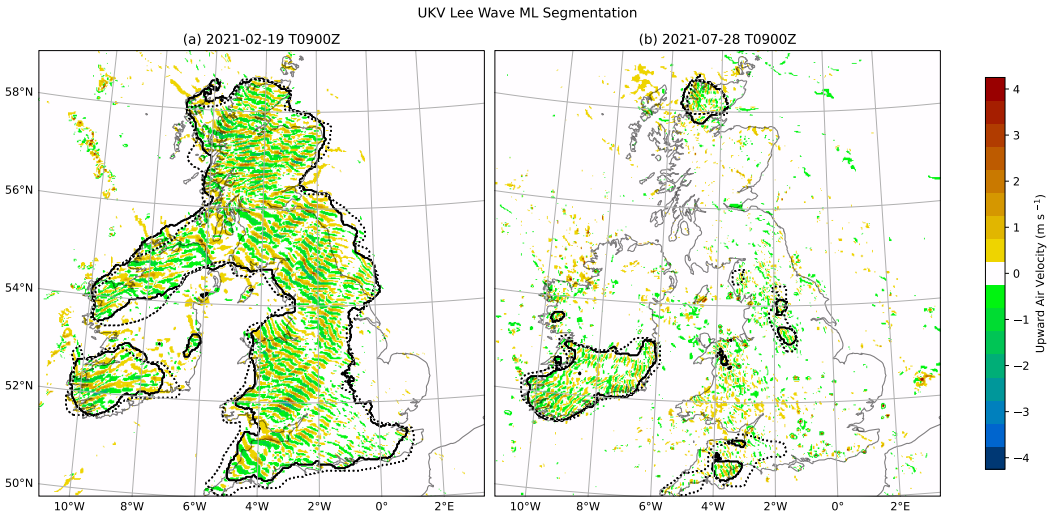


Figure 4.6: Two examples of the SEGMODEL predicted lee wave segmentation over the UK. Vertical velocities at 700 hPa are shown in colour filled contours, with wave regions predicted by the model shown by the black line contour. The dotted line contour shows the hand labelled waves. (a) is an example from the test set in February 2021. Against the "truth" data, the segmentation in (a) achieved a pixel accuracy of 94% and a Jaccard score of 0.87. (b) is an example of data from July 2021 containing waves segmented by the model with a pixel accuracy of 97% and a Jaccard score of 0.78.

of pixel ($i = 0$: no wave; $i = 1$: wave). The Jaccard Score is computed for each class by the area of overlap divided by the area of the union, of the model's prediction and the truth. Then the mean of these is taken to find the Jaccard score for the example of test data. The score shows how similar the prediction is to the hand labels and therefore how good the model is. While it is feasible to have used the Jaccard score as a loss function, the model was trained using cross entropy due to its good performance as a loss function in segmentation tasks (Jadon, 2020).

The Precision score is the number of true positives (correctly identified waves) divided by the number of true positives plus the number of false positives, and the recall score is the number of true positives divided by the number of true positives plus the number of false negatives (Pedregosa et al., 2012).

Table 4.3 shows that on the February 2021 test set, the trained SEGMODEL performed at 95% pixel accuracy. In this test set, only 16.9% of pixels were labelled as waves, even though this was from a winter period with higher wave activity. The ZEROS model "performs" well in the pixel accuracy metric in the February test set, reflecting the small amount of wave activity compared to the background. The SEGMODEL has a Jaccard score of 0.78 compared to ZEROS with a Jaccard score of 0.42 in this set. The

Table 4.3: Performance on the February 2021 and July 2021 test sets, by the SEGMODEL U-Net and ZEROS (a model that never predicts waves).

		Pixel Accuracy	**Jaccard Score**	**Precision**	**Recall**
February 2021 Test Set	SEGMODEL	95.0%	0.78	0.87	0.70
	ZEROS	84.4%	0.42	0.0	0.0
July 2021 Test Set	SEGMODEL	98.0%	0.66	0.48	0.42
	ZEROS	97%	0.49	0.0	0.0

Table 4.4: Confusion matrices for the SEGMODEL predictions for the February and July test sets. The matrices are normalised so the values sum to 1.

		February test set		July test set	
		Prediction		Prediction	
		No wave	Wave	No wave	Wave
Truth	No wave	0.70	0.14	0.97	0.004
	Wave	0.05	0.11	0.01	0.01

SEGMODEL is demonstrating skill by detecting waves in plausible locations compared to never predicting waves. Gravity waves have by their very nature a decaying amplitude envelope which makes defining a hard edge to a gravity wave envelope a poorly defined problem, which in turn means that exactly achieving a Jaccard score of 1 would be difficult. The precision and recall scores indicate a reasonable ratio between the number of correctly identified pixels and those incorrectly identified by the SEGMODEL. The ZEROS model scores 0 for both of these metrics since it did not correctly identify any waves.

On the July 2021 test set, Table 4.3 shows that waves occurred very infrequently from the pixel accuracy of the ZEROS model. However, the difference in Jaccard score shows that the SEGMODEL outperforms the ZEROS model by localising waves when they occurred. This infrequency of wave occurrence is reflected in the precision and recall scores. In addition, Table 4.4 shows the confusion matrices for the two test sets, again demonstrating the high occurrence of non-wave conditions over Britain and Ireland, particularly in July.

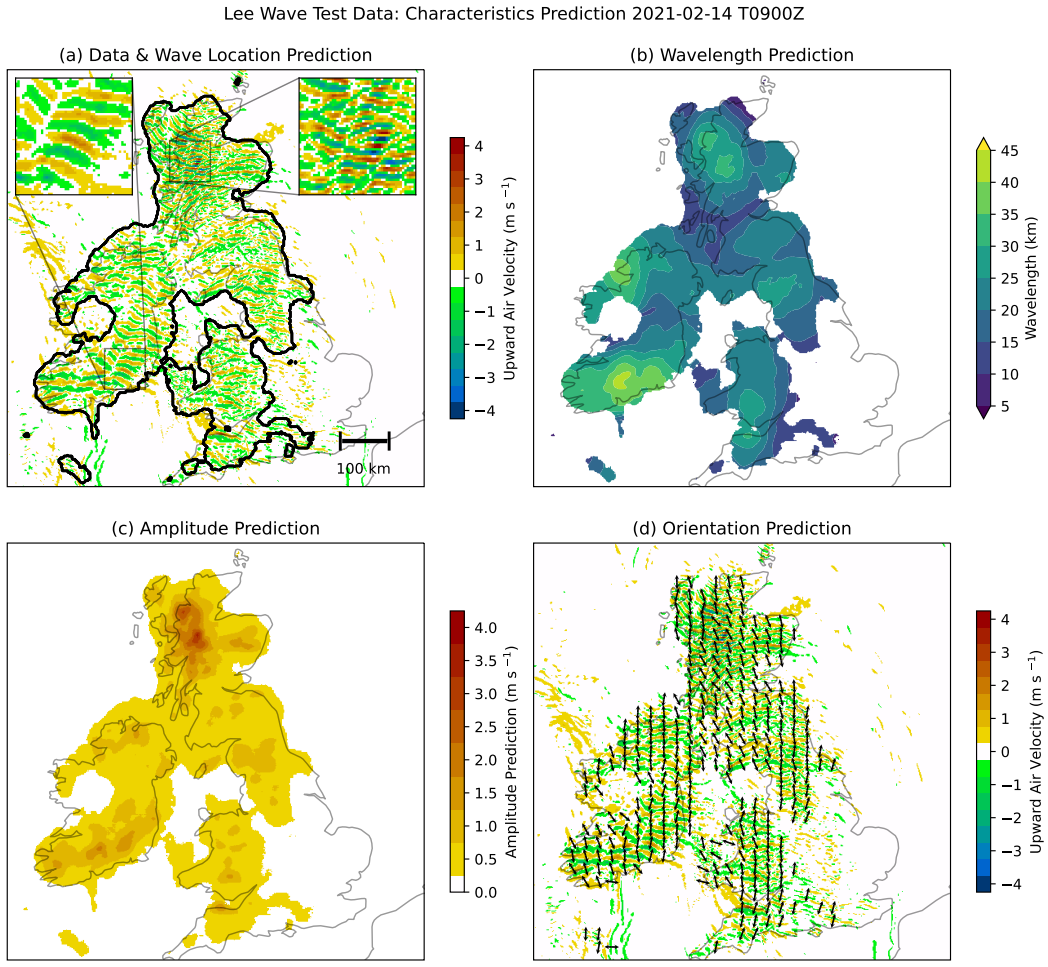


Figure 4.7: Example of lee wave characteristic prediction with: (a) vertical velocity and predicted lee wave regions (black line contour); (b) predicted wavelength for lee wave regions (model trained on data with noise $\sigma = 0.125$); (c) Predicted wave amplitude for lee wave regions; and (d) orientation of lee waves for lee wave regions (perpendicular to wave fronts), (model trained on data with noise $\sigma = 0.25$). The two inset regions in (a) demonstrate the difference in wavelength, amplitude and orientation in Scotland and Ireland for this particular case.

4.3.2 Lee wave characteristics

Having demonstrated the accuracy of the lee waves prediction network, the segmentation U-Net model is next utilised to infer lee wave characteristics: wavelength, orientation and amplitude. These predicted characteristics were then restricted to regions containing waves using the predicted wave masks from earlier.

As discussed in Section 4.2.5, transfer learning was used to fine-tune the final layers of the `SEGMODEL` network to learn wavelengths (`WLMODEL`), orientation (`ORIENTMODEL`) and amplitude (`AMPMODEL`) of the waves. Figure 4.7 shows one such example of the predicted characteristics, from an example of unseen test data.

The predicted characteristics were compared against a spectral technique, the S-transform, in order to have a method to compare characteristic predictions against. The following subsections will deal with each characteristic in turn.

4.3.2.1 Wavelength

The ML model and S-transform approaches were contrasted both on synthetic wavelength data samples (where the true wavelength value is known) and UKV simulation output (where the true value is not known). If the ML approach works well on the synthetic data compared with an S-transform, then this gives confidence that the ML derived wavelengths from UKV data are reasonable. While the S-transform derived wavelengths cannot necessarily be regarded as "truth", they can be used to ensure that the ML model is consistent with the S-transform and produces physically realistic wavelengths.

Figure 4.8 (a) and (b) show a 2D histogram for the synthetic test data set, comparing the ML model predictions and S-transform ($N = 80$ frequency voices and $c = 1$) derived wavelengths against the true wavelengths. The ML wavelengths compare well against the synthetic wavelengths, with $R^2 = 0.996$. There is a high density of points along the $y = x$ line in Figure 4.8 (a). The S-transform derived wavelengths compare with the truth favourably, though less so than the ML wavelengths, which is reflected in the lower R^2 value of 0.889, and the slightly larger scatter of points about the $y = x$ line in Figure 4.8 (b). The S-transform derived wavelengths are too small at wavelengths greater than 80 km, which is not seen in the ML derived

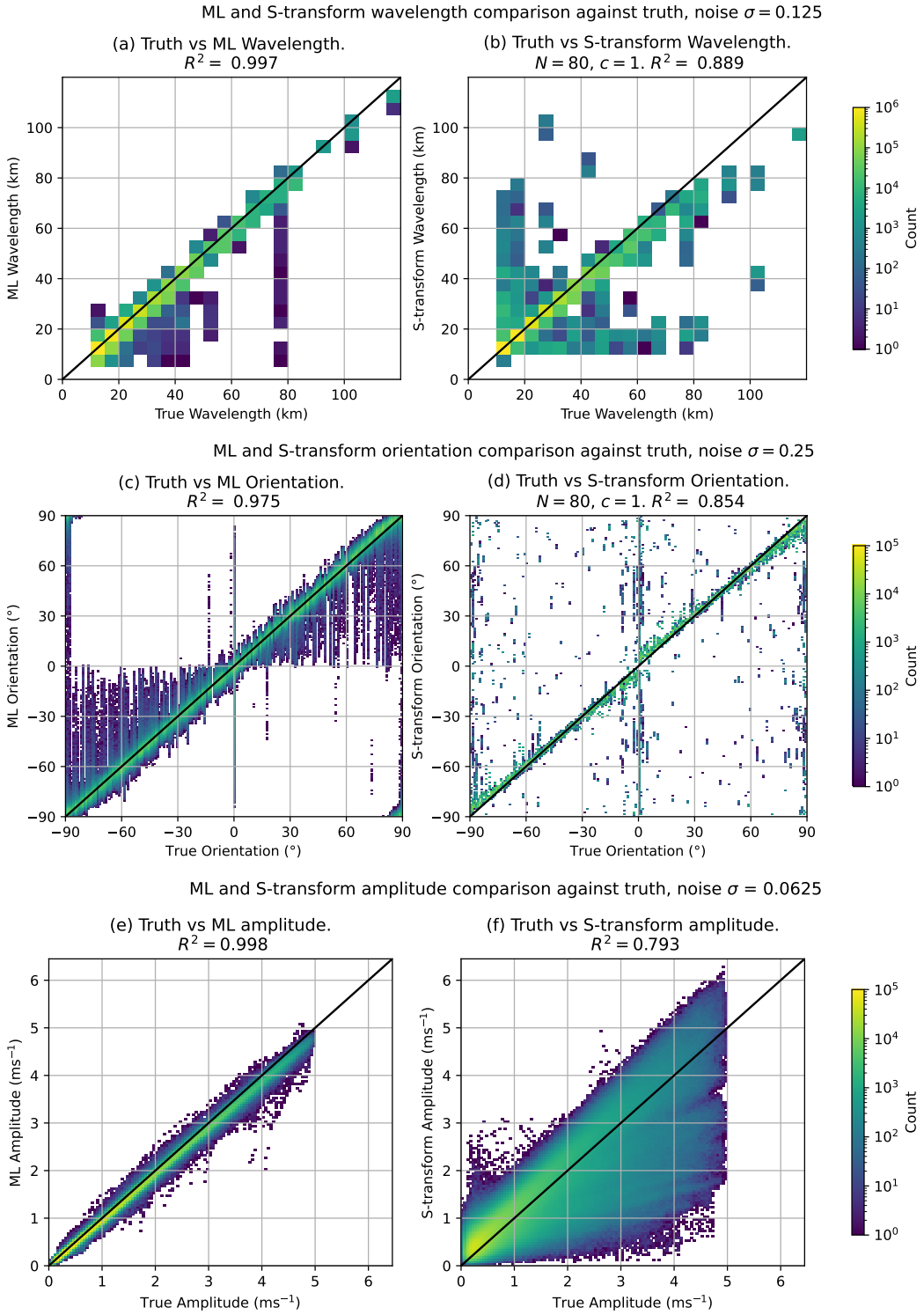


Figure 4.8: Synthetic data: comparison between true characteristics, ML model prediction, and S-transform characteristics for the synthetic test data set (with 80 frequency voices and a scaling parameter c of 1). (a, c, e): Histogram of truth vs ML derived characteristic from the test data set; (b, d, f) Histogram of true characteristic vs S-transform from the test data set. The black line in (a) - (f) is the line $y = x$.

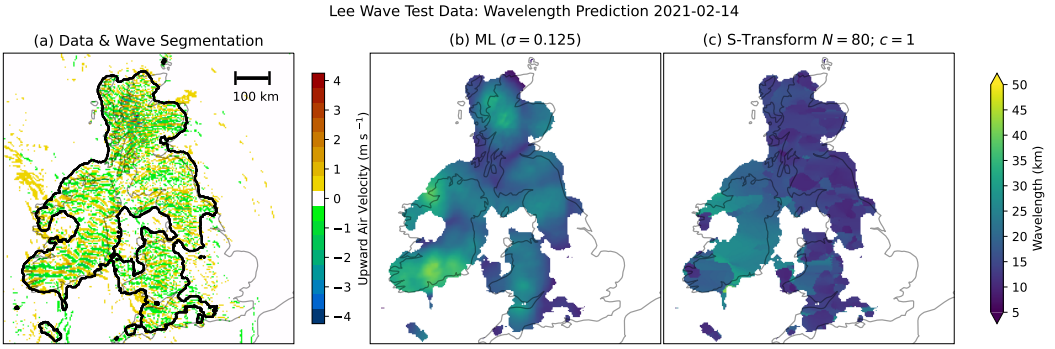


Figure 4.9: One selected example of machine learned wavelengths against S-transform wavelength from UKV data, from 2021-02-14 0900 UTC. (a) 700 hPa UKV vertical velocities and recognised lee wave regions; (b) `WLMODEL` (normally distributed noise, standard deviation $\sigma = 0.125$ in training data) derived wavelengths. (c) S-transform wavelength ($N = 80$ frequency voices, scaling parameter $c = 1$).

wavelengths.

Figure 4.9 shows the S-transform and the ML derived wavelengths for one example of UKV test data. The wavelengths predicted by the `WLMODEL` are reasonable, and relative wavelengths observed by eye correspond appropriately in both the ML prediction and the S-transform. For example in Figure 4.9 there is a region of longer wavelengths over the south of Ireland compared to shorter wavelengths over Scotland, which is predicted as such by the `WLMODEL`. The `WLMODEL` prediction shows a smoother field with greater variation in wavelength over the UK, but longer wavelengths than those produced by the S-transform. The S-transform wavelengths are, by comparison, more uniform compared to the ones from the `WLMODEL`. The sharp boundaries between regions in S-transform derived wavelengths are a product of the S-transform reproducing a clean wave field. However, this results in unrealistic discontinuities in the S-transform wavelengths. For example in Figure 4.9 (c), there is a discontinuity in wavelength over the south of Ireland according to the S-transform, which is not seen by eye in Figure 4.9 (a).

Figure 4.10 shows a histogram for the UKV February 2021 test set, comparing the `WLMODEL` predictions against the S-transform. The histogram only shows locations where the `SEGMODEL` predicts wave activity in the first place. Overall, the `WLMODEL` produces physically reasonable wavelengths (typical lee wave wavelengths are in the range 5 km to 35 km American Meteorological Society, 2012).

The `WLMODEL` wavelengths are slightly longer than the S-transform wavelengths. The

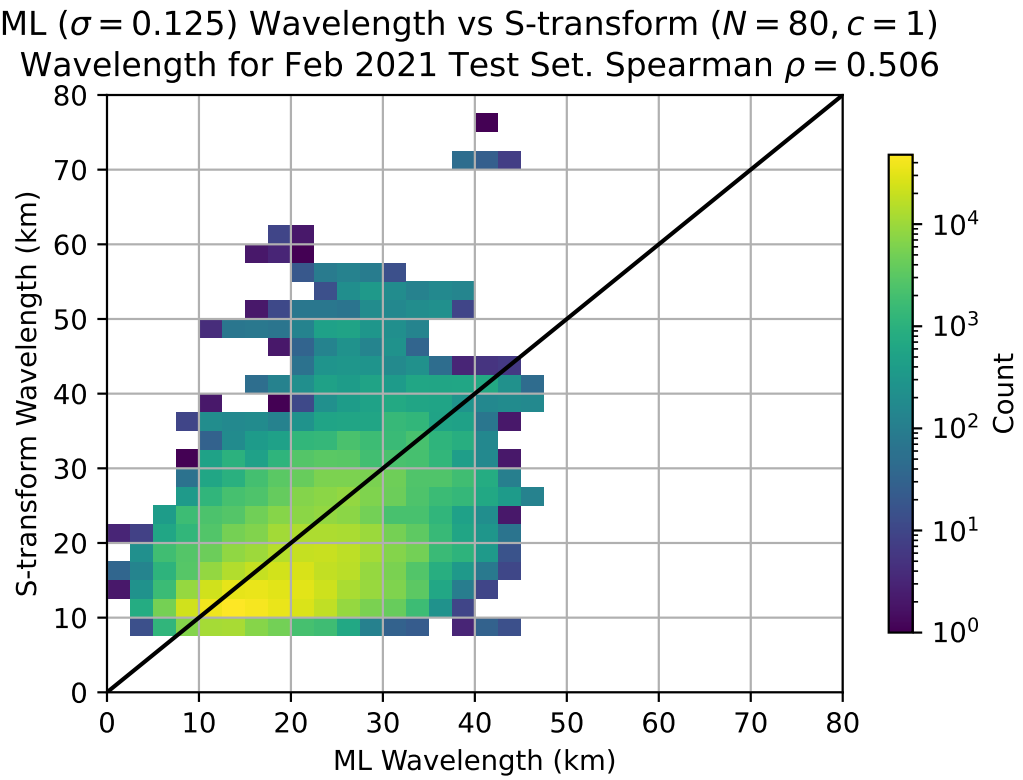


Figure 4.10: A histogram of the `WLMODEL` wavelengths against the S-transform wavelengths for the UKV test set.

range of wavelengths is smaller in the `WLMODEL` output than the S-transform. Overall, the `WLMODEL` gives reasonable indications of wavelength on the UKV data, compared to the S-transform. On the UKV data, while the `WLMODEL` wavelengths are typically longer than the wavelengths derived using the S-transform technique, the longest ML derived wavelengths are less than 50 km compared to the longest S-transform wavelengths being over 70 km. These longer S-transform wavelengths can occur in unrealistic locations (for example one case occurs in a region less than 50 km in diameter). This is despite good correlation in the synthetic test data.

4.3.2.2 Orientation

The angle predictions from the `ORIENTMODEL` and S-tranform were combined with the output from the original segmentation model, so that orientation predictions were only produced for regions containing waves. Figure 4.8 (c) and (d) compare the performance of `ORIENTMODEL` ($\sigma = 0.25$) and S-transform (80 frequency voices and $c = 1$) derived orientations on the synthetic test set. Figure 4.8 (c) shows that the S-transform derived orientations compare well with the true orientation,

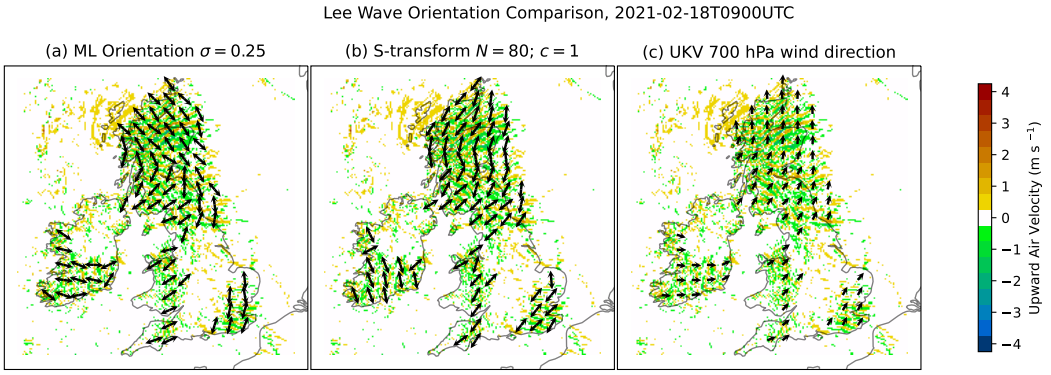


Figure 4.11: Comparison of the `ORIENTMODEL` predicted wave orientation, S-transform orientation and UKV 700 hPa wind direction. The black arrows show the direction of wave propagation/wind direction and are therefore perpendicular to the wave fronts. Wind direction is the closest variable in UKV to compare predicted wave orientation with.

with some seemingly random scatter across the axes. Figure 4.8 (d) shows that the ML derived orientations also compare well with the truth, but with greater spread towards 0°. Both methods of deriving orientation line up well along the $y = x$ line in each subplot.

In simple flows the wave crests would be expected to be perpendicular to the wind direction, and so wind direction can be used as a proxy for wave orientation. Due to the 3-D nature of the orography and the fact that the waves are not monochromatic, this assumption is not perfectly true, however the UKV wind direction at 700 hPa is still useful as an independent sanity check on the derived wave orientations.

Figure 4.11 shows an example of wave orientation (the `ORIENTMODEL` and S-transform) alongside UKV wind direction. `ORIENTMODEL` has done a good job of predicting the angle of the waves by eye. Neither tell the full picture, as the wind direction is not necessarily the same as the orientation of a wave. For example, regions with regular wave-like structures, such as those over Scotland, Wales, South East England and South West England have plausible predicted orientations. However, waves with less structure, such as those over Ireland, have predicted orientations which are less convincing by eye. The S-transform in this case has not captured fully the change in orientation over Scotland, and has the orientation more northerly than the `ORIENTMODEL`, which by eye seems to have captured the northwest/southeast orientation better.

Figure 4.12 (a) shows the `ORIENTMODEL` orientation against the UKV wind direction

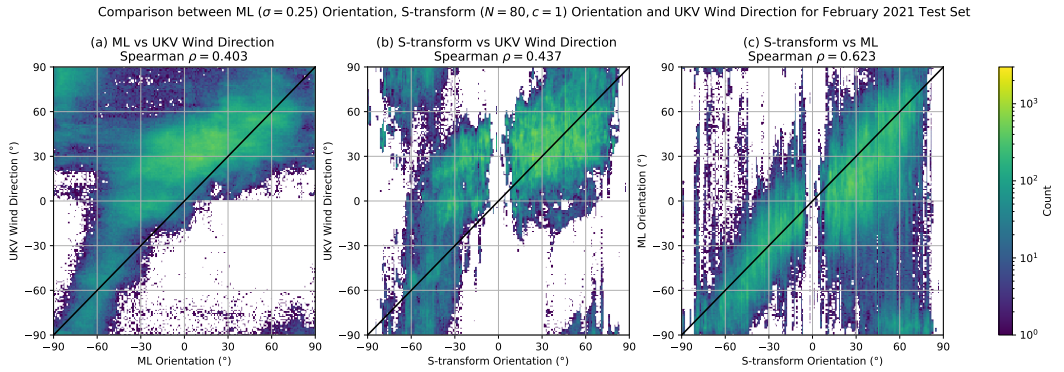


Figure 4.12: Histograms comparing the `ORIENTMODEL` orientation, S-transform orientation ($N = 80$, $c = 1$) and UKV wind direction at 700 hPa for the February 2021 test set in degrees from north (0°). In each case the line $y = x$ is plotted in black.

for the test set. The data contain a high degree of scatter, though there is a relationship by eye between the ML orientation and wind direction. The least squares correlation coefficient $R^2 = 0.116$ is low, however. Figure 4.12 (a) does show that, in general, as the `ORIENTMODEL` orientations veer, the wind direction veers as well, though not quite along the $y = x$ line. Figure 4.12 (b) compares the S-transform orientation against the wind direction. This plot is also noisy, though by eye shows correlation against the wind direction. As stated above, the wind direction is not necessarily a good predictor for the wave orientation, and the fact that the data in Figure 4.12 does not follow a 1:1 line may actually be for a good physical reason, for example because there is a preferential orientation of many of the mountain ranges over the UK. Finally Figure 4.12 (c) compares the `ORIENTMODEL` predictions against the S-transform. A positive trend is shown by eye, approximately along the $y = x$ line. The discontinuity in the S-transform measured orientations around 0° and 90° is due to its formulation using the discrete FFT. When orientated in exactly the x or y directions, wavenumbers in the orthogonal direction are equal to 0, corresponding to an infinite wavelength, which is just the signal mean in that direction and is therefore not able to be localised. This is not a limitation for the ML approach. The data contain a high amount of scatter, but do show some relationship between the ML derived orientation and S-transform, suggesting that the output from the `ORIENTMODEL` is reasonable. However it also suggests that deriving wave orientation is hard for both traditional spectral methods and ML methods.

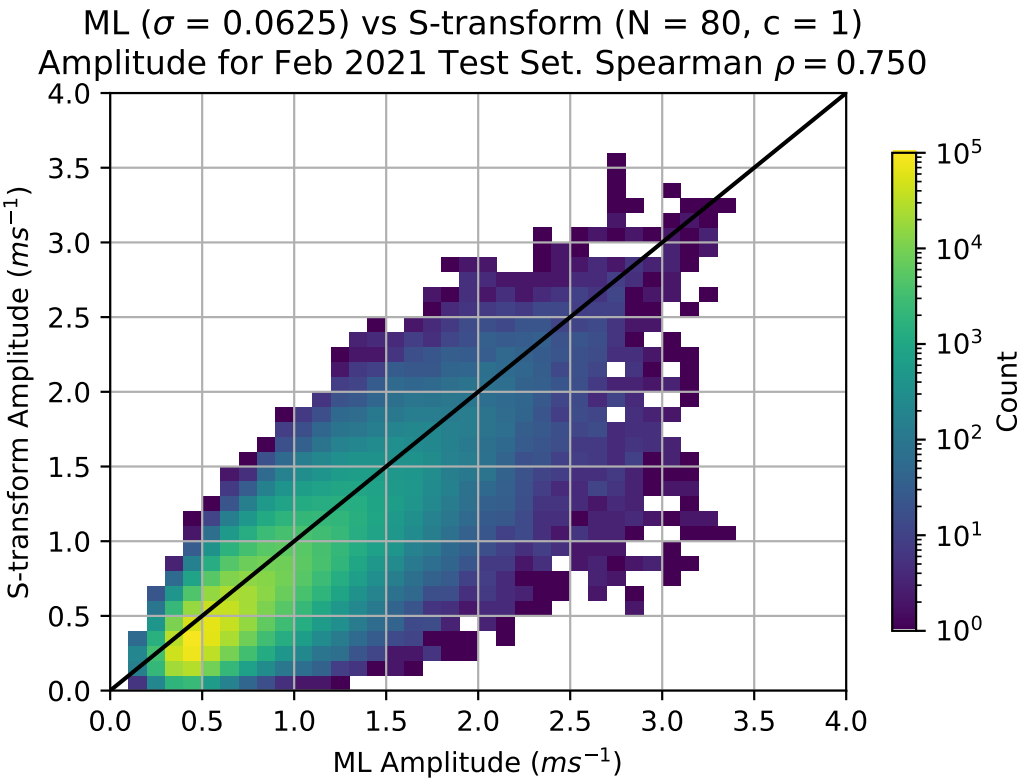


Figure 4.13: Histogram comparing the ML derived amplitudes against those from the S-transform for the February 2021 UKV test set. As with similar plots here, the line $y = x$ is plotted in black.

4.3.2.3 Amplitude

Figure 4.7 (c) shows the wave amplitude predictions by the ML model for 0900 UTC on 14 February 2021. The largest amplitudes tend to be over hilly areas, such as in Scotland. The amplitude predictions produced are reasonable when compared by eye to the vertical velocities in Figure 4.7 (a). An alternative approach using pixel-wise wavelengths to retrieve amplitudes by selecting the maximum vertical velocity within a region the size of the wavelength for each pixel resulted in unrealistic large local variations in amplitude. This meant that there were large regions containing unreasonably large amplitudes. The neural network approach as described here produces amplitudes that are more consistent and smoothly varying over the length-scale of the gravity wave envelopes seen visually in the vertical velocity data.

Figure 4.8 (e) and (f) compare the ML and S-transform amplitudes to the true amplitudes for the synthetic data. The ML model performs well, with an $R^2 = 0.997$, while the S-transform amplitudes had $R^2 = 0.729$ compared to the truth. Figure 4.13 compares the `AMPMODEL` derived amplitudes against the amplitudes from the S-transform. The amplitudes are well correlated, with a Spearman $\rho = 0.750$, and mainly focused around the $y = x$ line. At smaller amplitudes ($\approx 0.5\,\mathrm{m\,s^{-1}}$), the ML model overestimates amplitude slightly compared to the S-transform, while at higher ML amplitudes ($> 2\,\mathrm{m\,s^{-1}}$), there is a higher spread of S-transform predicted amplitudes.

4.3.3 Dimensionality reduction

The learned latent features of the trained `SEGMODEL` were probed using a dimensionality reduction technique, PCA, described in more detail in Section 2.2.1. This was in order to understand what the `SEGMODEL` had learned about lee waves during training. Some simple cosine wave data were generated at different wavelengths (w0-w5; w2 and w3 have identical wavelengths but are rotated by 90°) and orientations (o0-o6; o0 and o6 are identical). These are shown in Figure 4.14.

The intermediate outputs at the end of the encoder (shown in Figure 4.3) were used for this. At this point in the U-Net, the input data has had its features extracted and been coarsened in the x and y direction: each now had the dimensions

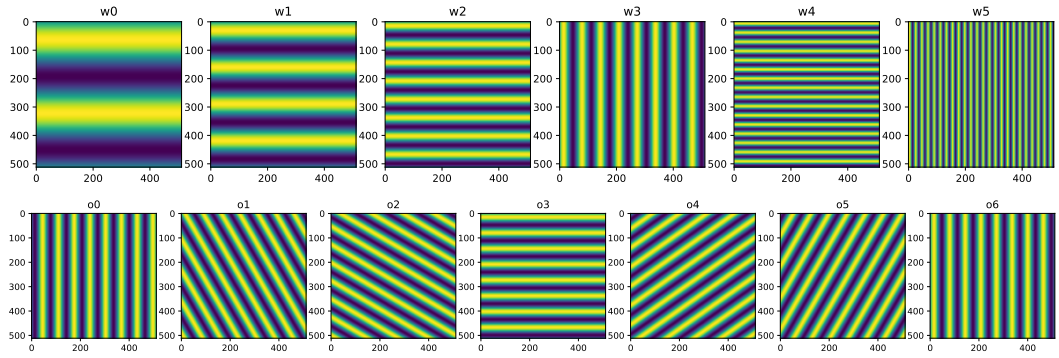


Figure 4.14: Different data for probing the learned latent features of the trained neural network.

$512 \times 16 \times 16$, (512 features and 16 pixels in both the x and y directions). In each case for wavelength and orientation, the first two principal components of the reduced space were taken and plotted as boxplots for each sample, shown in Figure 4.15. The first two principal components of the reduced data with different wavelengths (wo-w5), shown in Figure 4.15 (a) and (b) explain more of the variance than the orientation data (o0-o6), in Figure 4.15 (c) and (d). These were 44% and 15% compared to 23% and 19%. The first principal component of the wavelength data (a), shows the data with longer wavelengths occupying larger spaces. w2 and w3 had identical wavelengths but rotated by 90°, and the first principal component in (a) shows the distributions for w2 and w3 as being similar. Similarly for the orientation, o0 and o6 have similar distributions of principal components in (c) and (d), with the intervening data examples spreading out (o1 and o5, for example) and showing different distributions of principal components.

The change in wavelengths between w0 and w5 are shown clearly by eye in Figure 4.15(a) and the cycle of orientations from o0 to o6 are visible in the first two principal components in Figure 4.15 (c) and (d). These demonstrate that the trained U-Net has been able to learn the difference between different wavelengths and orientations within some wave-like patterns, even within the principal components of the three-dimensional output of an intermediate layer within the U-Net. This shows that during training, the U-Net has learned to distinguish between wave patterns of different wavelengths and orientations, which is then utilised in the fine-tuning of the `SEGMODEL` to produce the characteristics models discussed in Section 4.3.2.

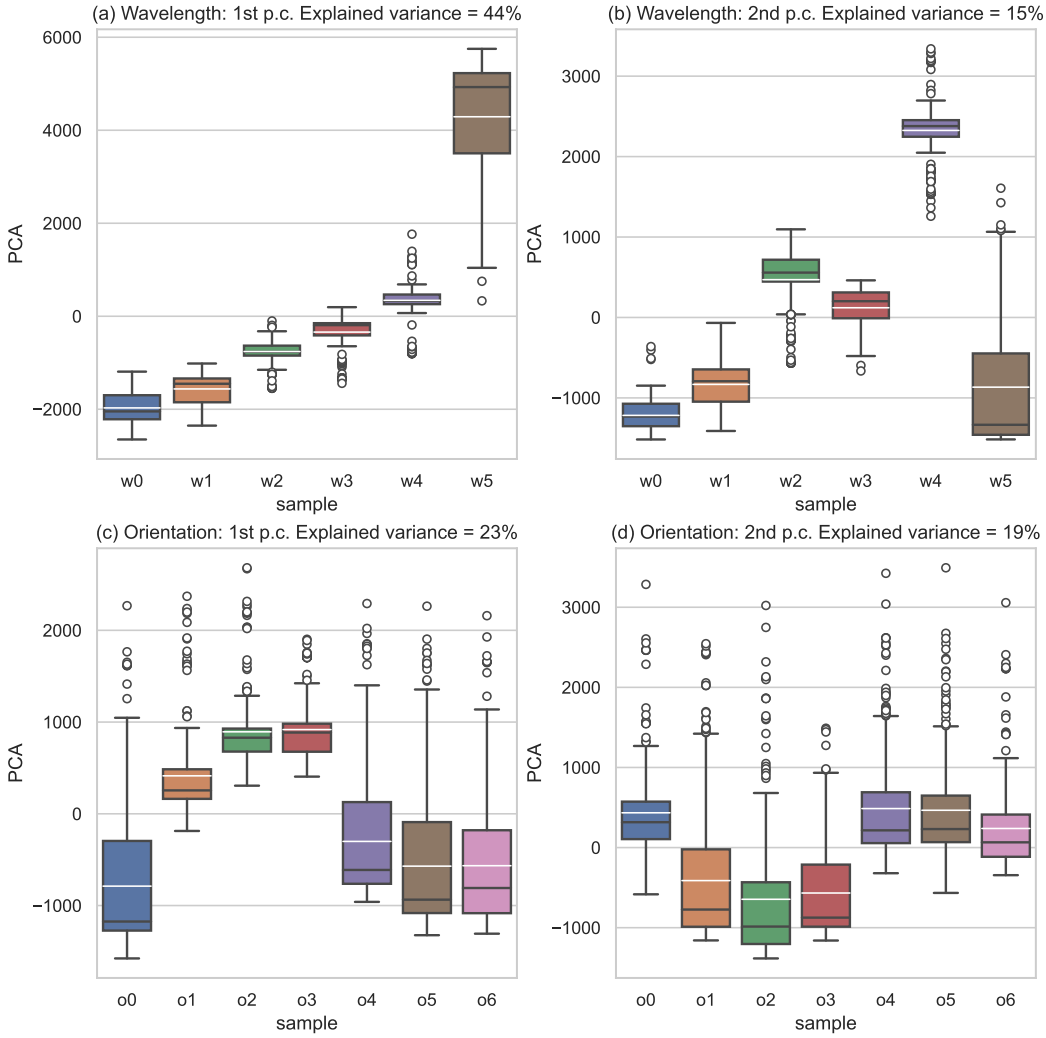


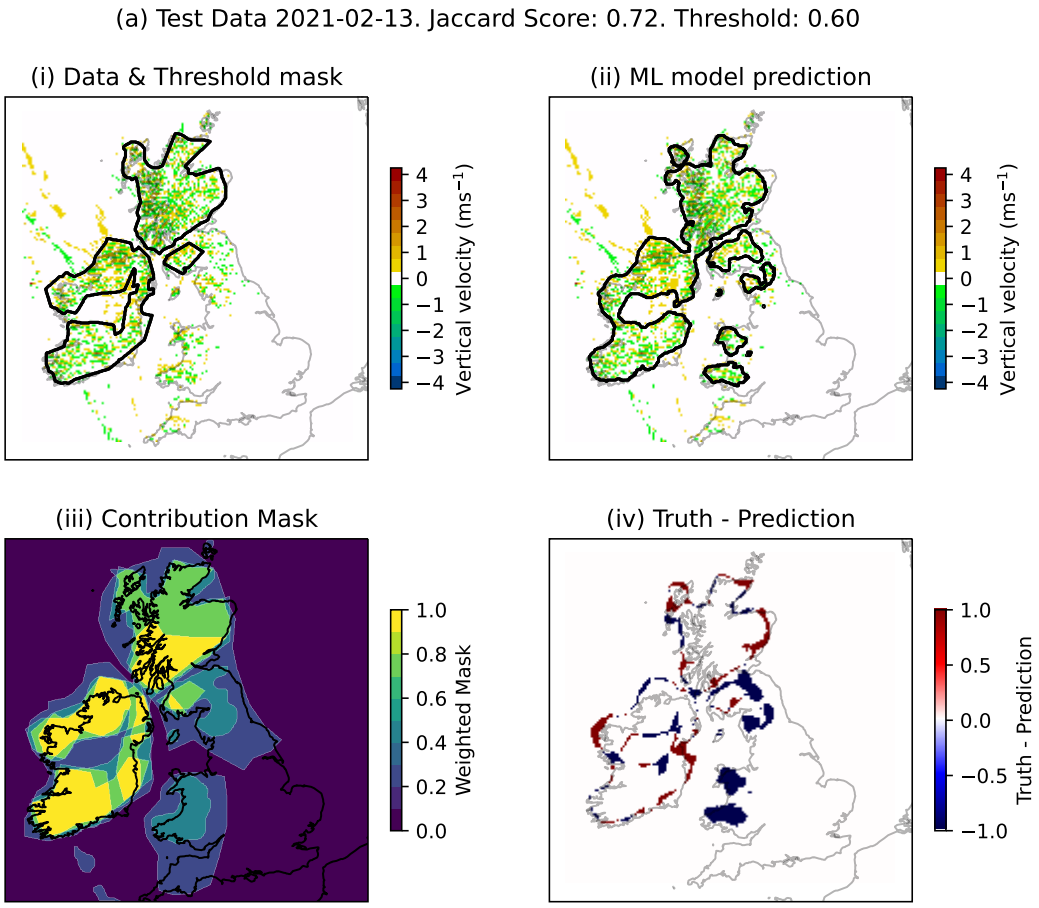
Figure 4.15: Box plots for the first two principal components for synthetic wavelength and orientation data. The median value is shown in black and the mean in white.

4.3.4 Crowdsourced labelling

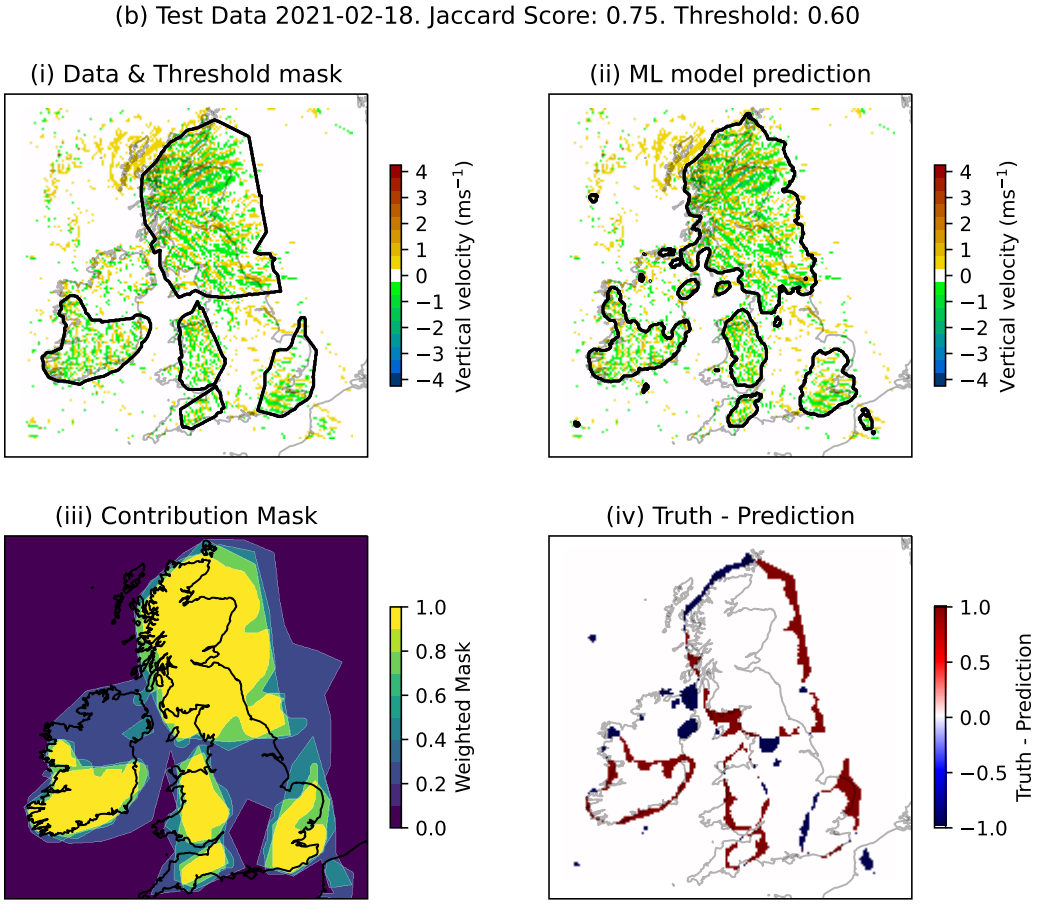
Additional crowd-sourced labels for the test set were generated using the online platform Zooniverse (Simpson et al., 2014), to investigate the subjective nature of the hand-labelling of the segmentation masks used to train the `SEGMODEL`. Six logged in users contributed labels, in addition to some users who were not logged in. The Zooniverse tool works slightly differently to the hand-labelling tool produced for the original mask making (discussed in Section 4.2.1), so there are slight differences in the methodology to produce a mask. The Zooniverse tool tends to produce masks with very straight edges.

The vertical velocity data presented to the labellers were presented upside down, to prevent the labellers using their geographical knowledge about likely wave locations over Britain and Ireland and therefore (consciously or unconsciously) be biased in their identification of waves. A revised test data set was created, where a pixel was only predicted to be a wave if at least 60% of labellers labelled it so, to ensure multiple labellers had to agree, while also not restricting the output to being the intersection of all labellers. Figure 4.16 shows two examples of crowdsourced lee wave labels on the data. Figure 4.16 also shows that there was a deviation between labellers as to what they perceived to be a lee wave, although there was more consensus for regions where waves were very obvious to the eye. When the boundaries between wave regions and the background are not clear cut, this adds ambiguity to the model output since there is some subjectivity between human labellers on what constitutes a wave.

There was some uptake among volunteers to label data, but labelling a large data set in this way would require a large amount of interest in the project in order to have multiple labels per vertical velocity slice. In addition, labellers would need to be aware of lee waves and how to recognise them from vertical velocity data. Crowdsourcing data has proved useful for meteorological data, for example the Weather Rescue project (Hawkins et al., 2019), which recruited members of the public to transcribe historic observations from the Ben Nevis and Fort William weather stations - albeit involving the clear-cut transcription of hand-written observations, rather than the subjective decisions about lee wave boundary labelling as was the case here.



(a) Example of crowdsourced labelling for 2021-02-13.



(b) Example of crowdsourced labelling for 2021-02-18.

Figure 4.16: Two examples of Zooniverse data at a 60% threshold. In each group of four: (i) shows the combined "threshold" hand-labelled mask (where 60% of labellers need to agree for a wave to be labelled as so); (ii) shows the predicted wave mask by the `SEGMODEL`; (iii) shows each hand-labelled mask for that data, where 0 indicates that nobody thought that a wave was in that location, and (1) everybody thought a lee wave existed in that location. (iv) The crowdsourced "truth" minus the `SEGMODEL` prediction, so blue indicates the model predicted a wave when the labellers did not, and red indicates a Zooniverse wave not recognised as such by the `SEGMODEL`.

4.4 Conclusions

This chapter shows the development of a set of deep learning models capable of identifying regions of lee waves over Britain and Ireland from vertical velocity model data. The final layers of the network have subsequently been fine-tuned separately to predict wavelength, orientation and amplitude of the waves. Spectral techniques (for example the S-transform) do not permit the creation of a pixel-wise wave mask as has been done here.

The trained segmentation model has a pixel accuracy of $> 95\%$ when compared against hand labelled truth. Despite being trained with a relatively modest training set of 335 scenes of vertical velocity data covering the whole of Britain and Ireland at 2 km resolution, the `SEGMODEL` U-Net is skilful. The segmentation produced by the `SEGMODEL` is realistic, and in coherent regions, for example those shown in Figure 4.6. The produced segmentation mask is on the same resolution as the NWP data, allowing precise localisation of wave forecasts. An accuracy closer to 100% would be difficult to obtain given the subjective nature of labelling waves near the edges of wave packets.

By using transfer learning, the copies of the trained `SEGMODEL` were fine-tuned on synthetic gravity wave data to estimate wave characteristics of gravity waves in UKV model output. The characteristics models still retain learned weights from being trained on the NWP data, with the synthetic data being used to extract characteristics from the model instead of a segmentation mask. Originally, these characteristics models were trained on data without noise, which, for the wavelength model, resulted in too long wavelengths being predicted on the NWP data. This has been rectified by training the characteristics models on noisy data.

On the UKV data, the wavelength model trained with noise tended to predict shorter wavelengths than with no noise, but longer than wavelengths derived using an S-transform. While the noise used within the synthetic training data are unlikely to be the same as the background vertical velocities in the UKV data, these results do suggest that using noisy synthetic data might help make the ML models more robust to noise in real world applications. How much noise to include in the training data, or how it should be distributed remains an avenue to be explored, though using a small amount of noise such as $\sigma = 0.125$ or 0.25 seems reasonable

given models trained on this magnitude of noise perform well across the different levels of noise in the synthetic data.

The orientation predictions, such as those shown in Figure 4.11 demonstrate that the `ORIENTMODEL` performs well at predicting the orientation of waves. Figure 4.11 shows that the wind direction at 700 hPa is not sufficient on its own to show the wave direction so this method of deriving the wave orientation from the vertical velocities could add value to existing forecasts of waves and rotor activity. Figure 4.12 shows that the wind direction is less correlated with the ML ($\rho = 0.403$) and S-transform ($\rho = 0.437$) predictions compared with the S-transform and ML predictions ($\rho = 0.623$). On the UKV data, the models trained on noisy data still retain a smooth field, but with wavelengths and orientation closer to those from the S-transform. Several wavelength and orientation models could be run on the data to obtain a measure of the uncertainty between the ML model derived characteristics.

The `AMPMODEL` produces a smoothly varying prediction of the wave amplitude, where individual peaks in vertical velocity are smoothed out. The test set, despite being small, contains cases of large amplitude waves (velocities in excess of $3\,\mathrm{m\,s^{-1}}$, as used in Vosper et al., 2013) as shown in Figure 4.7 (c). Large amplitude waves have potential impacts downstream for the formation of rotors and so successfully identifying these is important for forecasters.

While this work has used U-Nets to create an ML model capable of identifying and characterising lee waves, it also highlights the wider potential of these methods to be used in identifying a wide range of weather features and phenomena in high resolution model data. The study also offers useful examples of leveraging the maximum impact from limited hand labelled data by supplementing with augmentation and carefully constructed synthetic data sets. It is also a valuable example of how with fine-tuning an ML model developed to classify features can be used to identifying underlying physical characteristics of the features. While these ideas could easily be applied to other wave problems in geophysical systems, they could equally be applied to a range of other types of feature in the atmosphere.

The following chapters of this thesis explore the application of these trained models to large climate data sets, so as to develop a climatology of lee waves, and then interrogate the conditions that affect the generation of lee waves.

Chapter 5

A climatology of trapped lee waves over Britain and Ireland obtained through deep learning models applied to high resolution model output

5.1 Introduction

This chapter outlines the development of a climatology of trapped lee waves over Britain and Ireland, created by applying the deep learning models developed in Chapter 4 to 31 years of high resolution regional climate model output using the United Kingdom Climate Projections (UKCP18) Local data. The features in the climatology include the frequency of occurrence, wavelength, amplitude and orientation of trapped lee waves, covering a much longer period of time and using a more sophisticated model than the work of Vosper et al. (2013). Variations in lee wave location, frequency of occurrence and characteristics are investigated as a function of time of day, season and weather pattern. Finally, random forests are used to investigate relationships between lee waves and 40 other model variables in order to evaluate the important physical processes controlling the climatology.

The generation of this climatology aims to help answer questions regarding the generation and characteristics of lee waves over Britain and Ireland. The diurnal cycle

within the inversion in the boundary layer could affect the trapping, amplitude or wavelength of lee waves over Britain and Ireland, and this climatology data set allows this to be tested. In addition, by examining different weather patterns within the climatology, the effects of changing synoptic meteorology on the generation and characteristics of lee waves can be investigated, spatially and also through distributions of characteristics. In addition, the relationships between different meteorological variables and lee wave generation and characteristics can be examined, such as wind speed and wave amplitudes, and the Scorer parameter and lee wave wavelengths.

An overview of the method and the data used is given in Section 5.2, the results are presented and discussed in Section 5.3 and some concluding remarks are given in Section 5.4.

5.2 Methodology

5.2.1 Developing a climatology of lee waves

As in Chapter 4, vertical velocity on the 700 hPa surface was used to identify lee waves from the model output. In addition to the vertical velocity slices used by the deep learning models, other variables from model output were used to understand which conditions are important for the production of lee waves (as discussed in Chapter 3).

The ERA-Interim driven UKCP18 Local data (as introduced in Section 3.2.1) were obtained from MASS, specifically the vertical velocities on the 700 hPa surface, for the three-hourly data from 1 January 1982 03 UTC to 30 December 21 UTC inclusive. The trained deep learning models trained in Chapter 4 were applied to the 31 years worth of UKCP18 Local model data to produce a wave segmentation mask (a pixelwise classification into non-wave and lee wave pixels), and the wavelength, orientation (the direction of wave propagation, perpendicular to the wave-fronts) and amplitude of lee waves over Britain and Ireland during 1982–2012.

These segmentation masks and characteristics were then interrogated to investigate how lee waves differ at different times of the day, seasonally, and between synoptic weather patterns. This was done using the `dask` library in Python to produce stat-

istics for the prevalence and characteristics of lee waves geographically, and overall (using distributions and histograms), for Britain and Ireland from the individual segmentation mask and characteristic files.

5.2.2 Relationship of lee waves with other meteorological variables

In order to examine which meteorological variables were important for the generation of lee waves, several upland regions of Britain and Ireland with a high frequency of occurrence of lee waves (shown in Figure 5.1) were subsetted from the data. These regions were deliberately chosen to be similar to those used by the 3DVOM forecast domains as shown in Sheridan et al. (2017), but with the region over Wales being extended and the inclusion of a new region over the south west of Ireland. These were given rough geographical names related to their location (for example, the "Northern Ireland" box does not necessarily align with the political borders of Northern Ireland, but does contain the orography in that part of Ireland). The lee waves produced by the segmentation ML model were compared with other variables from the UKCP18 Local data within these regions in order to explore the relationship between meteorological conditions and lee wave generation.

To do this, as introduced in Section 3.2, a set of variables likely to be correlated to lee wave activity were initially chosen, on available pressure surfaces, (for example, variables used in the calculation of the Scorer parameter were included): the horizontal wind speed on the 200, 300, 500, 650, 750, 850 and 925 hPa pressure surfaces and at 10 m ; the change in horizontal wind direction between each of the aforementioned pressure surfaces; the virtual potential temperature θ_v on the 200, 300, 500, 650, 700, 750, 800, 850 and 925 hPa surfaces; the Brunt-Väisälä frequency; the Scorer parameter; the weather pattern; the half-sine of the month of the year (precisely $\sin\left(\frac{\pi(m-1)}{12}\right)$, where m is the month of the year: $m = 1$ is January and so on); and the orography and its local standard deviation (to capture the variability of orography), to make 40 inputs in total.

The mid-level Brunt-Väisälä frequency at x hPa was calculated as $N(x) = \sqrt{\frac{g}{\theta_v}\frac{\mathrm{d}\theta_v}{\mathrm{d}z}}$, where $\frac{\mathrm{d}\theta_v}{\mathrm{d}z}$ was calculated using the surfaces above and below x hPa, and θ_v at x hPa by interpolating the data from the surfaces above and below x hPa. This produced mid-level values of N, on the 887, 825, 775, 725 and 675 hPa surfaces. A bulk

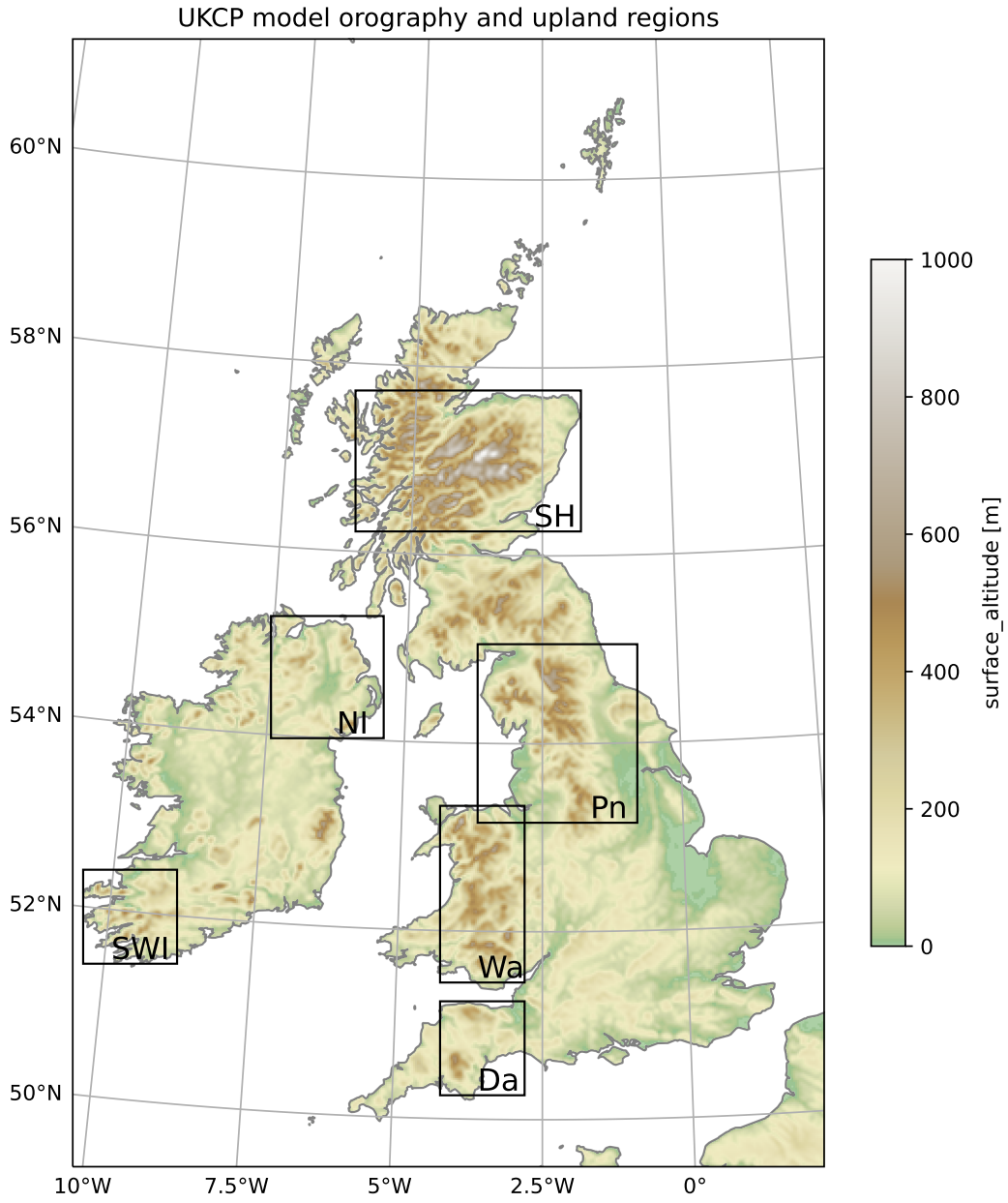


Figure 5.1: UKCP18 Local model orography over Britain and Ireland. Black bounding boxes indicate upland regions used in later analysis and are labelled for identification in the text: SH: Scottish Highlands; NI: Northern Ireland; SWI: South West Ireland; Pn: Pennines; Wa: Wales; Da: Dartmoor.

Brunt-Väisälä frequency was also calculated, by using a mid-profile value of θ_v, at 775 hPa, being representative of most of the vertical profile, and using the highest and lowest values of θ (925 hPa and 650 hPa) to calculate $\frac{\mathrm{d}\theta_v}{\mathrm{d}z}$.

The Scorer parameter was calculated as $l^2 = \frac{N^2}{U^2}$, with the smaller (and harder to calculate accurately) curvature term neglected. However, the influence of changing wind direction was still explored, by including the change in wind direction between each pressure surface as separate inputs. Values of the Scorer parameter were produced on the same pressure surfaces as the Brunt-Väisälä frequency by interpolating U onto the same surfaces as the Brunt-Väisälä frequency (Blockley and Lyons, 1994). The bulk Scorer parameter was calculated using the bulk measure of the Brunt-Väisälä frequency, and using the horizontal wind speed at 775 hPa. This is a mid-profile measure of the horizontal wind speed so most likely to represent most of the profile as well as being above the boundary layer and keeping consistency with the value of θ_v used to calculate the bulk Brunt-Väisälä frequency. The change in the Scorer parameter from 887 hPa to 675 hPa was also calculated. These variables are summarised in Table 5.1.

Random forests (see Section 2.2.2) were trained on these variables (one random forest per region in Figure 5.1) with the aim of predicting whether or not lee waves were present. Random forests were used because of the lack of normalisation or augmentations required on the input data, and because they are quicker to train than neural networks. The nodes in each decision tree are based on the data, for example horizontal wind speed being greater than some threshold. Then, because of the relative ease of probing a random forest's predictions (compared to, for example ANNs), the influence of each of these variables on the random forest's prediction was measured using SHAP values (see Section 2.2.6 and Lundberg and Lee, 2017). Here, the version of SHAP for tree based models (TreeExplainer: Lundberg et al., 2020) was used.

In addition, the correlation between the amplitude of lee waves and the horizontal wind speed aloft was examined. The local maximum lee wave amplitude and 750 hPa wind speeds were grouped for 8×8 pixel blocks (17.6 km $\times$ 17.6 km) for part of the climatology period (1982–87 inclusive). This reduced set was used to reduce the time taken to obtain the additional data, and conserve disk space and processing time, as opposed to using the full data set. Grouping the data into these blocks meant that

Table 5.1: Summary of the inputs for the random forests for meteorological variables on heights and pressure surfaces. In addition to these 35, the weather pattern, half-sine of the month of the year, the difference in Scorer parameter between the 887 hPa and 675 hPa pressure surfaces, the model orography and its local standard deviation were also included.

Height/Pressure surface	Variable				
	$\sqrt{U^2+V^2}$	Wind shear	θ_v	N	$l=\sqrt{\frac{N^2}{U^2}}$
10 m	✓				
925 hPa	✓		✓		
887 hPa		✓		✓	✓
850 hPa	✓		✓		
825 hPa				✓	✓
800 hPa		✓	✓		
775 hPa				✓	✓
750 hPa	✓		✓		
725 hPa				✓	✓
700 hPa		✓	✓		
675 hPa				✓	✓
650 hPa	✓		✓		
575 hPa		✓			
500 hPa	✓		✓		
400 hPa		✓			
300 hPa	✓		✓		
250 hPa		✓			
200 hPa	✓		✓		
Bulk measure				✓	✓

small local variations in wind speed and amplitude were removed, meaning a fairer comparison between broader wind speeds and amplitudes. The crest of a wave and therefore the peak amplitude should occur within these squares in all cases apart from the longest wavelength lee waves, without losing too much fine detail for short wavelength lee waves.

5.3 Results

5.3.1 Frequency of occurrence of lee waves

Figure 5.2 shows the seasonal frequency of occurrence of lee waves over Britain and Ireland. The general location of lee waves changes little between seasons, with lee waves tending to occur more often over hilly areas (Scotland, Northern England, Wales, South West Ireland). However, the winter months contain the highest frequency of occurrence of lee waves and the summer months the least. Over the Highlands of Scotland, lee waves occur 60% to 70% of the time during winter in the UKCP18 Local data, compared with 40% to 50% of the time during summer months. This is consistent with Vosper et al. (2013), where lee waves were forecast 57% of the time over the Grampians in Northern Scotland.

Figure 5.3 shows a histogram relating the UKCP18 Local model orography and the frequency of occurrence of lee waves per pixel (normalised by column). The standard deviation σ of the model orography for the 5×5 pixel box surrounding each grid cell were calculated and plotted against the grid cell's frequency of wave occurrence during the period of interest. There is a correlation (least squares $R^2 = 0.711$ and Spearman rank $\rho = 0.662$) between the local orography variability σ, and the frequency of occurrence of waves. There is spread within the histogram: this is likely due to lee waves propagating well downstream past the orography, and the influence of meteorology which will vary across the country.

5.3.2 Diurnal effects

Figure 5.4 shows the frequency of occurrence of lee waves for each of the hours of the day in the UKCP18 Local data. There is no discernible diurnal variation

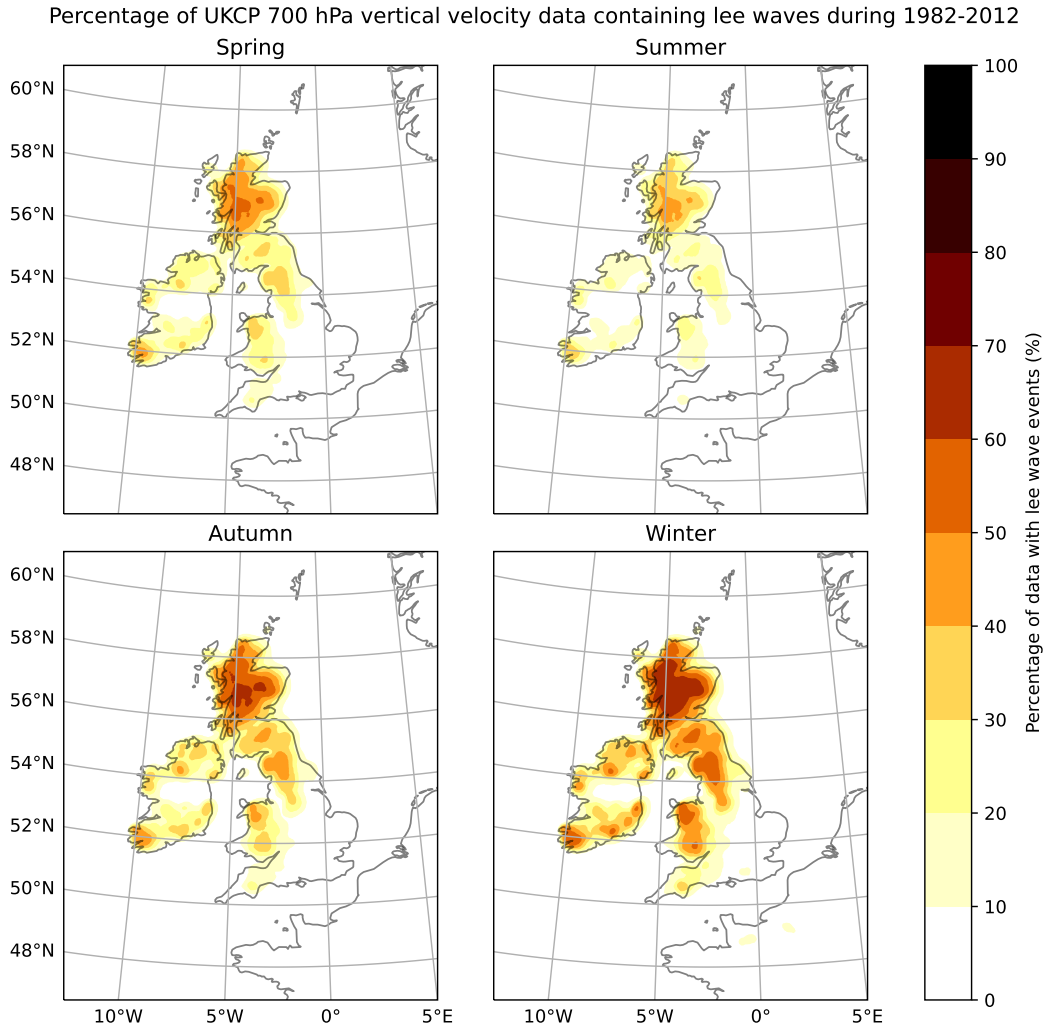


Figure 5.2: Maps showing the frequency of occurrence of lee waves by season.

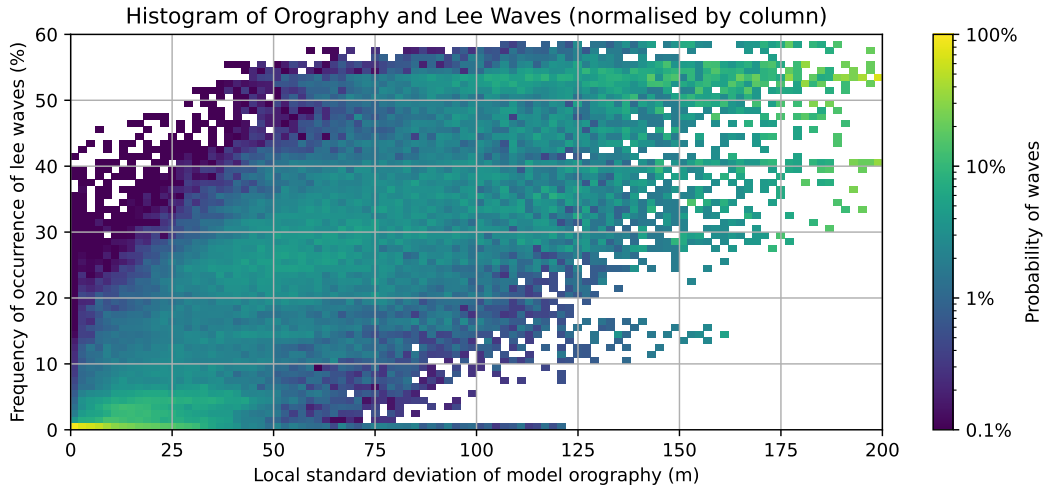


Figure 5.3: Histogram of local standard deviation in the UKCP18 model orography and the frequency of occurrence of lee waves per 5×5 pixel box. This plot is normalised by column (so the sum of each column is 100%).

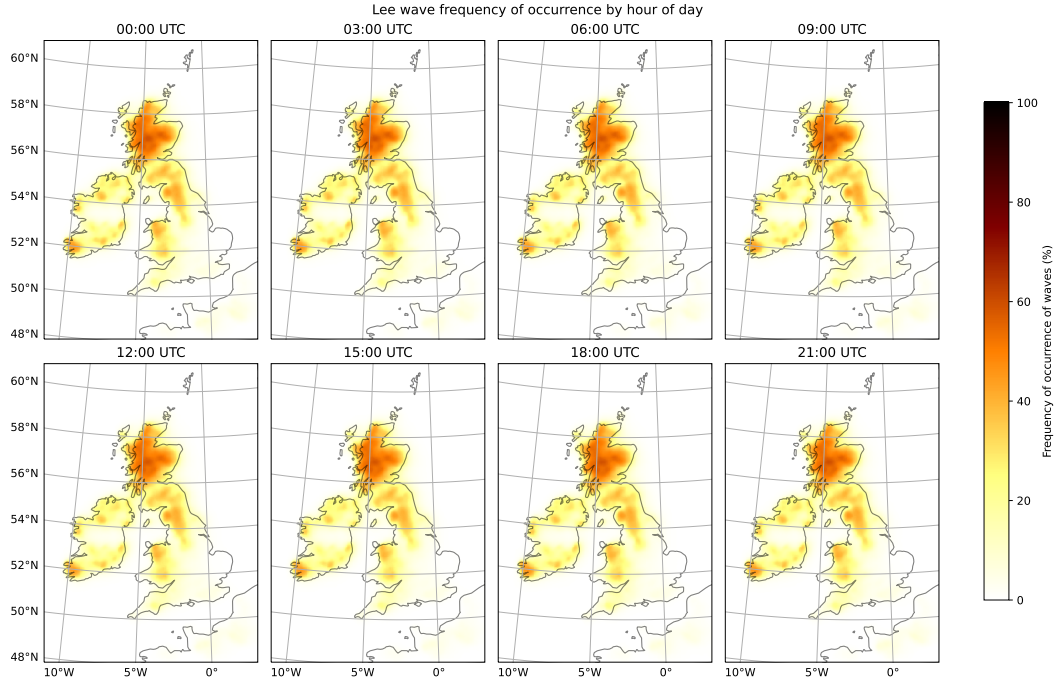


Figure 5.4: Map plots showing the frequency of occurrence of lee waves by hour of the day in the UKCP18 Local data over the climatology period 1982–2012.

geographically in the frequency of occurrence of lee waves over the entirety of Britain and Ireland.

Figure 5.5 shows the probability density functions (PDFs) for the frequency of occurrence and characteristics of lee waves for each time of day, and confirms that there is very little diurnal variation in either the frequency of occurrence or the characteristics of lee waves in the data. The temporal resolution of the data is 3 hourly, so unpicking the potential effects of sunrise and sunset on lee waves is not immediately obvious. However, the data at 0600Z and 1800Z cover either side of sunrise and sunset respectively, depending on the time of year, and there is no variation in the distribution of the data at these times to suggest that sunrise and sunset affect the generation of lee waves materially. There is no change between distributions of wave characteristics. Within the range of amplitudes observed in Figure 5.5 (b), from $< 0.1\,\mathrm{m\,s^{-1}}$ to $> 5\,\mathrm{m\,s^{-1}}$, the lack of a diurnal change except at very low probabilities and high amplitudes suggests there is little diurnal change in lee wave amplitude during the day. Likewise for the wavelength (Figure 5.5[c]) of lee waves, most observed lee waves have a wavelength in the range $10-20\,\mathrm{km}$ but no diurnal change. Most lee waves are either north east/south west; or north west/south east aligned (Figure 5.5[d]), but there is not an observed diurnal cycle

of lee wave orientations.

Hence, there is little evidence to suggest a diurnal cycle in lee wave frequency of occurrence or characteristics over Britain and Ireland, similar to the conclusions of Worthington (2006), who found very little diurnal change in mountain wave amplitude, albeit using a different method (they used VHF radar observations).

5.3.3 Weather patterns

There is a variation in the frequency of occurrence of lee waves according to the synoptic weather pattern (introduced in Sections 2.4 and 3.3, and in more detail in Neal et al., 2016). Figure 5.6 shows the frequency of occurrence of lee waves geographically by weather pattern (0% indicates that lee waves never occur in a given weather pattern, and 100% indicates that lee waves always occur under that pattern), with the insets showing for each season the total number of days on which the respective weather pattern occurred during the climatology period 1982–2012.

The weather patterns most conducive to lee waves being generated are the NAO+ and Southwesterly patterns. The Scottish Highlands receive the most lee waves in all conditions, but the trend is not consistent across regions: the Pennines receive a similar amount of waves in NAO+ conditions, but comparatively less so under the Azores high pattern. While the spread of NAO− weather patterns is fairly constant between seasons, NAO+ conditions, which have a high frequency of occurrence of lee waves, occur more often during the winter than in the summer. This may explain the seasonal changes of lee waves in Figure 5.2: more lee wave conducive weather occurs during the winter than in the summer. Under Scandinavian high conditions, lee waves occur infrequently over most of Britain and Ireland except for in Scotland and a small patch in south west Ireland.

Figure 5.7 shows an overview of the frequency of occurrence and characteristics of lee waves under different synoptic weather patterns. Each subplot shows a PDF of the weather pattern's lee wave frequency of occurrence or characteristic. For the frequency of occurrence Figure 5.7 (a) shows that the most lee waves occur during NAO+ and Southwesterly conditions, and the least during the NAO− and Scandinavian high patterns. The following subsections will consider the characteristics shown in Figure 5.7 (b)–(d) in more detail.

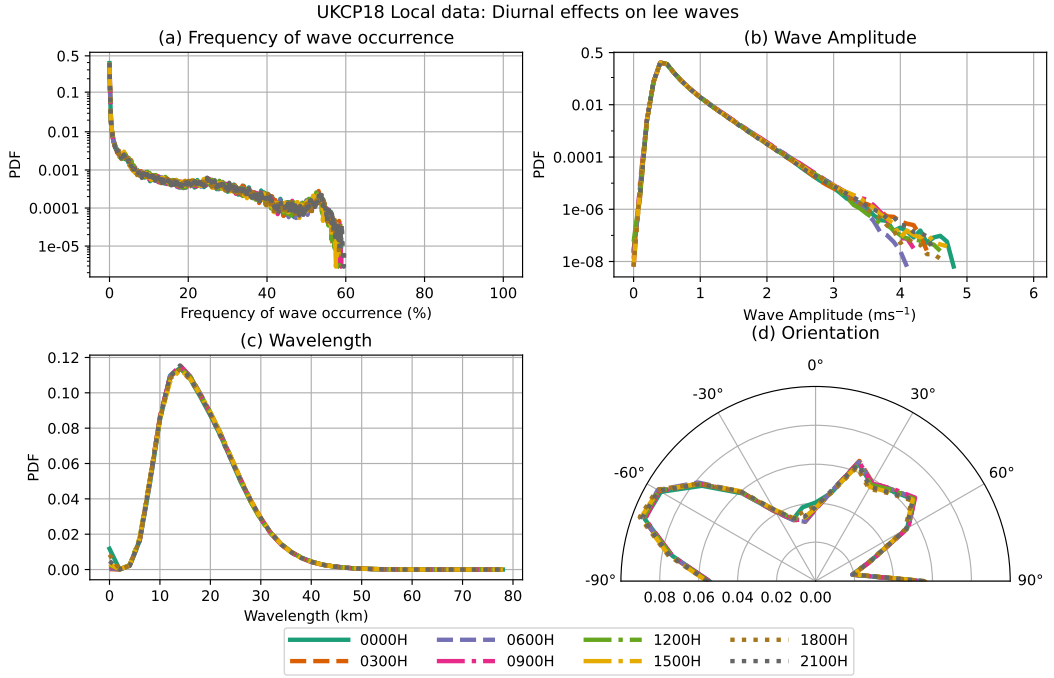


Figure 5.5: PDFs of lee wave frequency and characteristics by hour of the day.

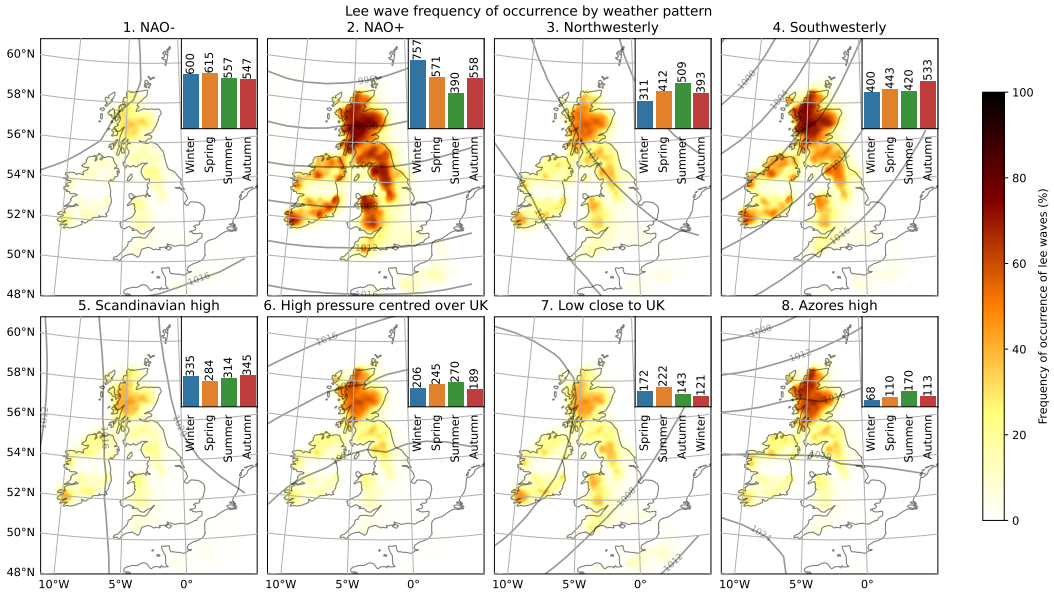


Figure 5.6: Lee wave frequency of occurrence by weather pattern (Neal et al., 2016) from the UKCP18 model data over the climatology period 1982–2012. Mean sea level pressure contours are plotted for each pattern. The inset axes for each subplot show the number of days 1982–2012 where each weather pattern was prevalent.

Table 5.2: Table showing how often lee waves of different maximum amplitudes occur, for each region. This includes times when there is no wave activity in a region, and the number in brackets excludes examples where there is no wave activity in a region.

Region	**Percentage of time with maximum wave amplitude:**			
	$\geq 1\,\mathrm{m\,s^{-1}}$	$\geq 2\,\mathrm{m\,s^{-1}}$	$\geq 2.5\,\mathrm{m\,s^{-1}}$	$\geq 3\,\mathrm{m\,s^{-1}}$
Scottish Highlands	34.7% (49.9%)	5.4% (7.7%)	1.66% (2.38%)	0.383% (0.551%)
Northern Ireland	11.3% (25.5%)	0.4% (0.9%)	0.06% (0.14%)	0.006% (0.012%)
SW Ireland	14.6% (28.3%)	0.8% (1.6%)	0.18% (0.36%)	0.023% (0.045%)
Pennines	20.1% (37.4%)	1.4% (2.6%)	0.28% (0.53%)	0.041% (0.076%)
Dartmoor	5.6% (18.9%)	0.2% (0.8%)	0.04% (0.12%)	0.006% (0.019%)
Wales	18.9% (38.3%)	1.4% (2.9%)	0.23% (0.46%)	0.025% (0.052%)
Overall	51.0% (59.7%)	10.0% (11.7%)	2.89% (3.39%)	0.624% (0.732%)

5.3.3.1 Amplitude

Figure 5.7 (b) shows that the amplitude of lee waves is distributed similarly between weather patterns. The largest amplitude lee waves occur under Northwesterly and Southwesterly flows, however the changes in amplitude between weather patterns are small compared to the actual mean amplitudes. The vast majority of lee waves have amplitude $< 1\,\mathrm{m\,s^{-1}}$, and the largest amplitudes in the data are $\approx 5\,\mathrm{m\,s^{-1}}$.

Figure 5.8 shows the 95th percentile amplitude of lee waves for each of the weather patterns. No weather pattern shows a distribution of 95th percentile amplitudes noticeably stronger than any other. The regions where the stronger amplitudes occur tend to be places with higher orography, such as the Scottish Highlands in each weather pattern, the Pennines and Wales (particularly in the NAO+ pattern).

Table 5.2 shows how often lee waves of different amplitudes occur in the regions outlined in Figure 5.1. While being rare in general, the strongest lee waves occur most frequently in the Scottish Highlands, where high amplitude lee waves (amplitude in excess of $3\,\mathrm{m\,s^{-1}}$) occur 0.383% of the time within the region (0.551% of the time when only counting when lee waves are generated). The Pennines have the next most frequent occurrence of high amplitude lee waves, at 0.041% of the time, almost ten times less likely than in the Scottish Highlands region. High amplitude lee waves are the rarest in Northern Ireland and Dartmoor.

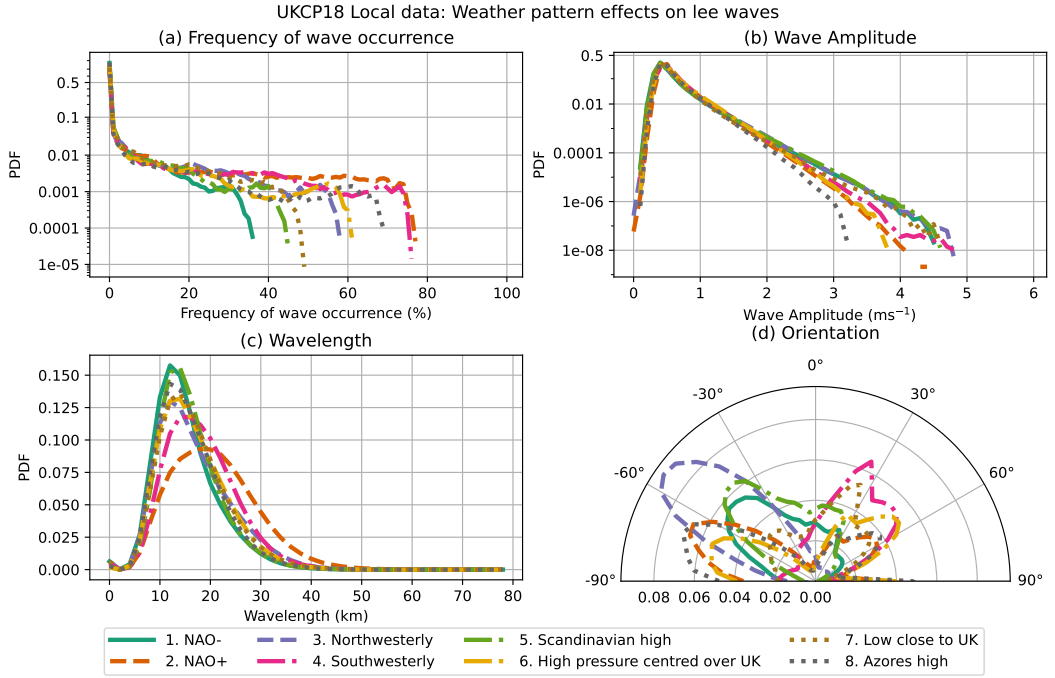


Figure 5.7: PDFs of lee wave frequency of occurrence and characteristics under different synoptic weather patterns.

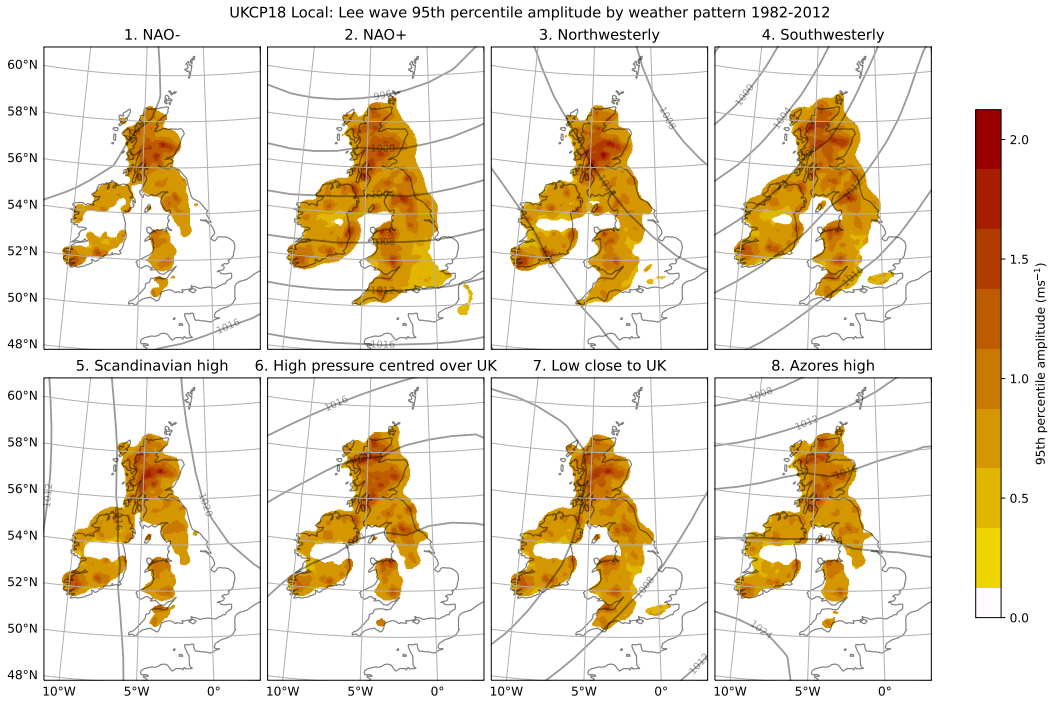


Figure 5.8: Maps showing the 95th percentile lee wave amplitude for each weather pattern.

5.3.3.2 Wavelength

Figure 5.7 (c) shows the PDF of the wavelengths of lee waves under the different weather patterns. Lee waves under most regimes (all except NAO+ and Southwesterly) have a modal wavelength in the range 10 to 15 km. Lee waves produced under Southwesterly flow have a modal wavelength of approximately 15 km, while the longest modal wavelengths of approximately 20 km occur under NAO+ conditions.

Figure 5.9 shows the mean wavelength of lee waves over Britain and Ireland by weather pattern. The longest wavelengths occur in general in NAO+ conditions, to the east of Ireland, and on the lee (east) side of the Pennines. The shortest wavelengths tend to occur under NAO− conditions, though with some longer wavelengths in Scotland. Apart from the NAO+ case (and to a lesser extent, Southwesterly conditions), the mean wavelengths remain generally similar between patterns. The distribution of wavelengths in Figure 5.7 are broadly similar for all the patterns except for NAO+ and Southwesterly. According to theory, the wavelength is related to the Scorer parameter (e.g. World Meteorological Organization, 1993). However, outside the NAO+ and Southwesterly patterns, there is little change in the wavelengths produced overall. This could be because the length scale of the orography influences the wavelength of the lee waves produced because the length of the orography controls whether or not lee waves are produced at their theoretical wavelength, and therefore the wavelengths are less influenced by the synoptic weather than other characteristics.

5.3.3.3 Orientation

Figure 5.7 (d) shows that there is a different orientation pattern depending on the weather pattern. The orientations under some weather patterns have a unimodal distribution, such as the Southwesterly or Northwesterly patterns. Others, such as the NAO+ and Low close to UK patterns, have bimodal orientation distributions. Since the 8 patterns are themselves groupings of 30 more varied ones, there could be more than one pressure regime grouped together, resulting in the bimodal wave orientations under certain patterns.

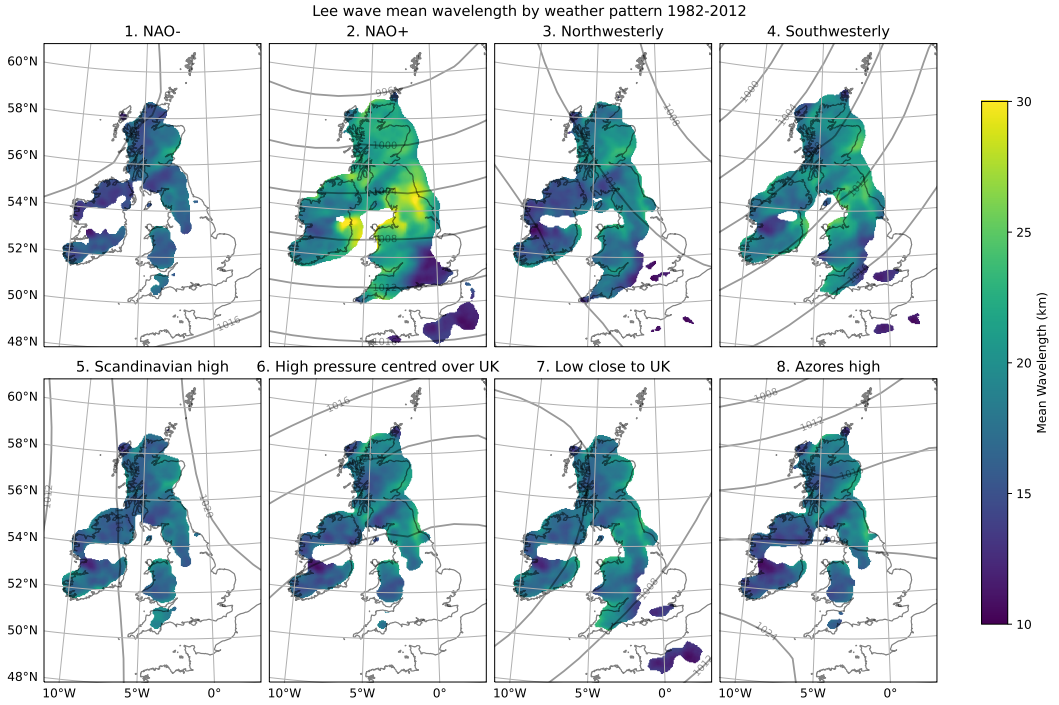


Figure 5.9: Map showing the mean lee wave wavelength by weather pattern.

Figure 5.10 shows the mean orientation of lee waves for each weather pattern over Britain and Ireland, with MSLP contours overlain for the respective pattern. Lee waves are broadly aligned with the MSLP contours, notably so in Northwesterly and Southwesterly conditions. However, there is a large spread in the orientation predictions, shown by the red arrows (indicating $\sigma > 45°$) in Figure 5.10. The least variation in wave direction occurs under the Northwesterly and Southwesterly weather patterns, and the most variation in the Low close to the UK, again perhaps due to the clustering to form the 8 patterns resulting in a large variation in MSLP patterns being clustered together.

5.3.4 Correlation between lee waves and other NWP variables

To investigate any correlation between lee waves and other meteorological variables, random forest models are used to investigate how well lee waves can be predicted from other NWP variables. A random forest is trained on a set of meteorological variables deemed to have influence on the generation of lee waves in each of the regions shown in Figure 5.1 (one random forest per region), as introduced in Section 5.2.2: the horizontal wind speed on different pressure surfaces; the directional wind shear between surfaces; the virtual potential temperature θ_v; the Brunt-Väisälä frequency;

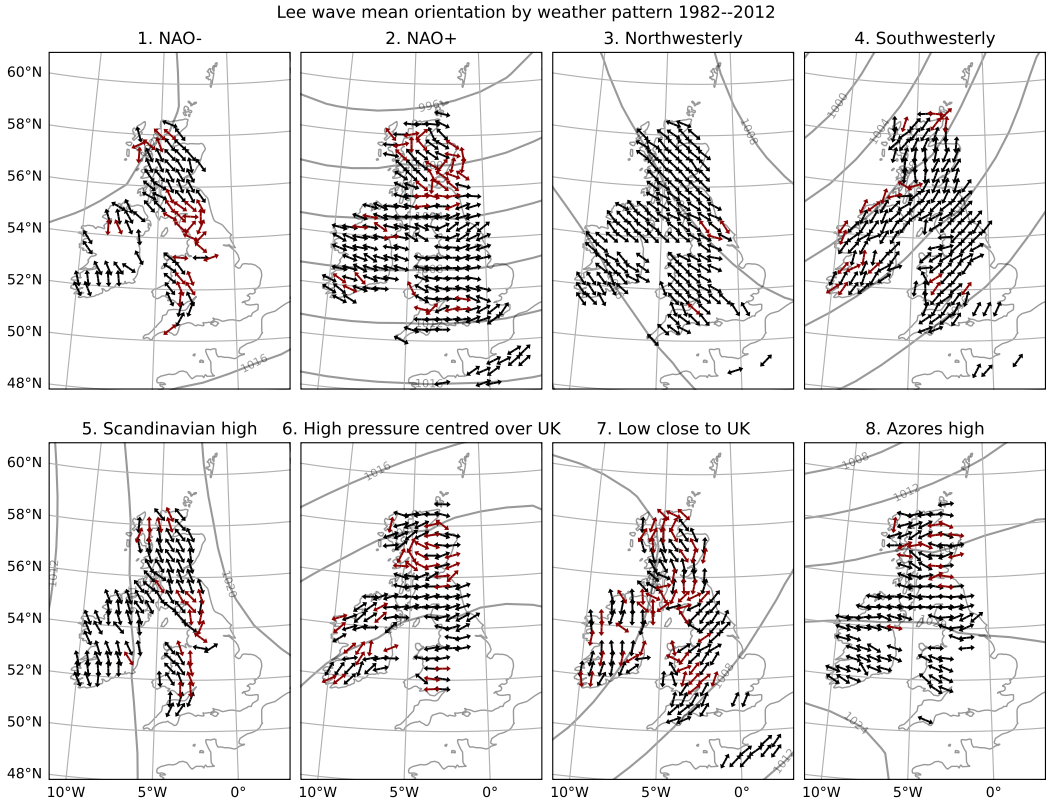


Figure 5.10: The mean orientation of lee waves for each weather pattern. Arrows in maroon indicate that the standard deviation, σ of the orientation at this point is $> 45°$, while black arrows indicate $\sigma \leq 45°$.

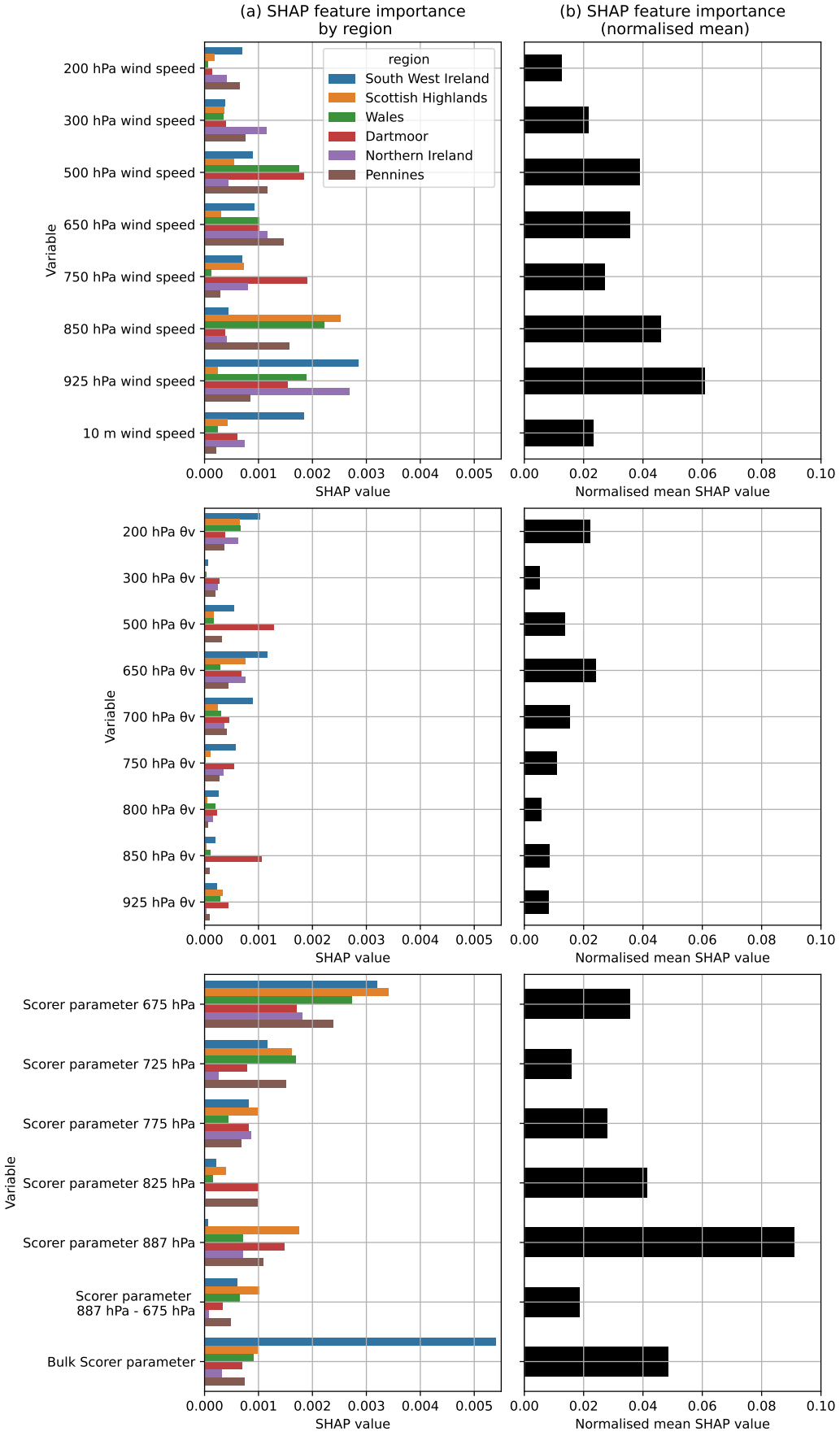


Figure 5.11: (a) SHAP values for each of the inputs into the random forest to predict lee wave generation, by region of Britain and Ireland, and (b), each region's SHAP values are normalised so that the sum of the SHAP values for each region is 1, and then the mean of these are plotted (continued overleaf).

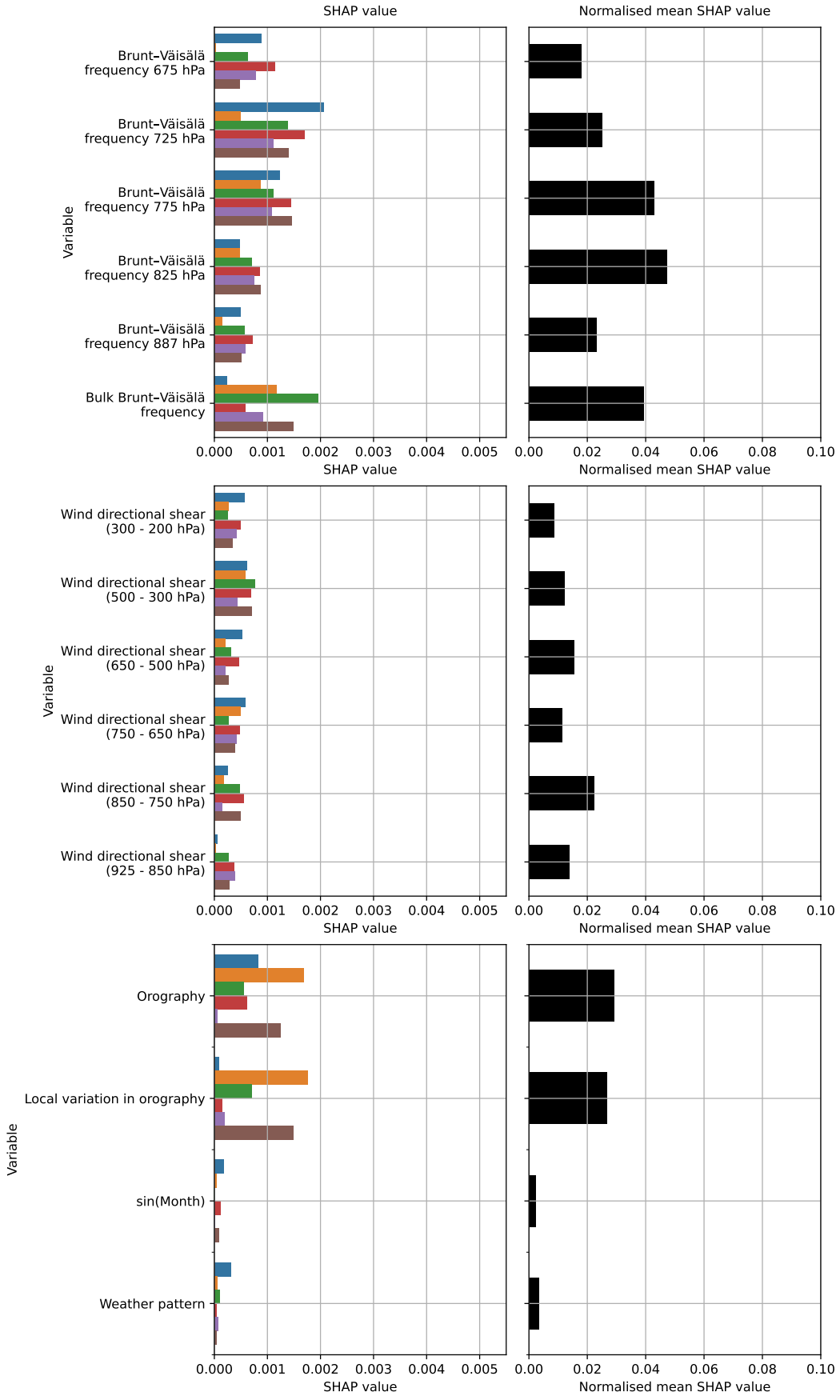


Figure 5.11: (a) SHAP values for each of the inputs into the random forest to predict lee wave generation, by region of Britain and Ireland, and (b), each region's SHAP values are normalised so that the sum of the SHAP values for each region is 1, and then the mean of these are plotted.

Table 5.3: Table showing the variables with the largest SHAP value by region.

Region	Variable with largest SHAP value
South West Ireland	Bulk Scorer parameter
Scottish Highlands	Scorer parameter 887 hPa
Wales	Scorer parameter 887 hPa
Dartmoor	750 hPa wind speed
Northern Ireland	925 hPa wind speed
Pennines	Scorer parameter 887 hPa

a measure of the Scorer parameter (only the first term - see Section 5.2.2); the orography and its local standard deviation; the weather pattern; and the half-sine of the month of the year. These variables were chosen because of their involvement in the Scorer parameter (the stability, horizontal wind speed, turning effects) or the generation of lee waves seen earlier in this section (the month of the year, weather pattern and orography). The relative importance of each of these are calculated using SHAP values (Section 2.2.6), and are shown by region in Figure 5.11 (a), and a normalised average over the regions in Figure 5.11 (b). Table 5.3 shows the variables with the largest SHAP value for each of the 6 regions considered: in each region the variable with the largest SHAP value are all either horizontal wind speeds or a measure of the Scorer parameter.

The Scorer parameter between the 925 hPa and 850 hPa surfaces (i.e. at 887 hPa), in this case taken to be the ratio between the squares of the Brunt-Väisälä frequency and horizontal wind speed, has the largest SHAP importance value in the Scottish Highlands, Wales and the Pennines. A bulk measure of the Scorer parameter has the highest importance value in South West Ireland. In Northern Ireland, the 925 hPa horizontal wind speed is the most important while in Dartmoor, the 750 hPa wind speed is the most important. The Scorer parameter at 887 hPa has at least the third highest SHAP value in every region examined here. The wind speed with the larger SHAP values differs depending on the region: for example regions with generally taller orography, such as the Scottish Highlands and Wales, have higher wind speed SHAP values at 850 hPa, while regions with lower orography such as Northern and South West Ireland have higher SHAP values for wind speeds at 925 hPa and 10 m respectively. The Brunt-Väisälä frequencies at 825 hPa and 775 hPa have higher SHAP values than the other pressure surfaces. The SHAP values of the Brunt-Väisälä frequencies are more consistent (but also lower) between surfaces and

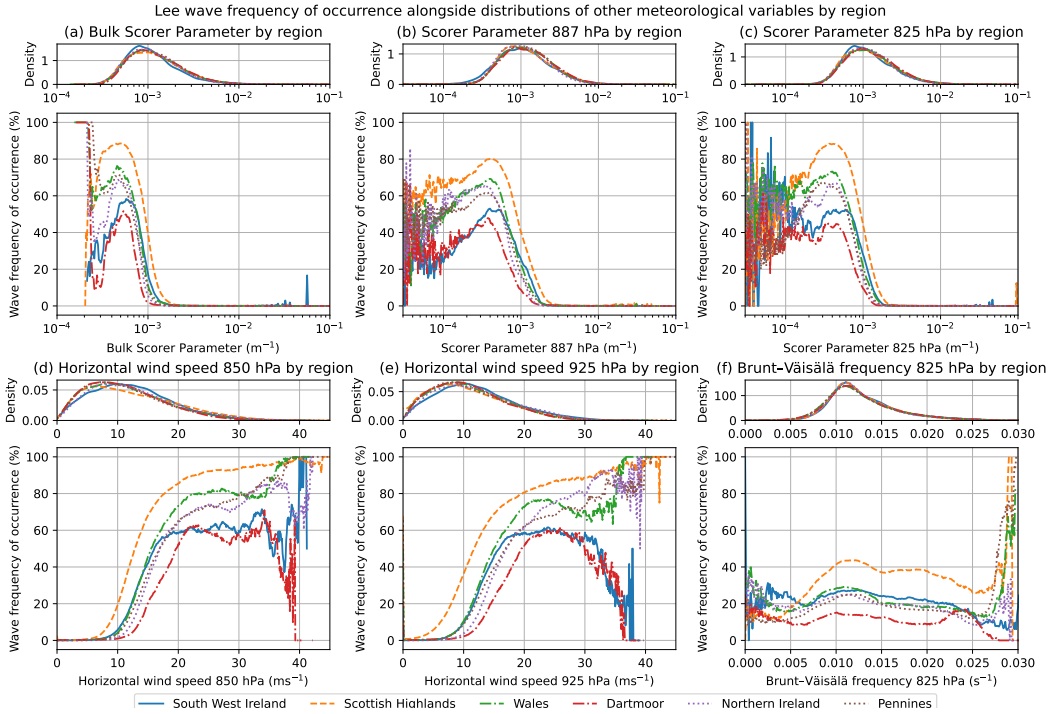


Figure 5.12: Distributions showing the relationship between the variable with the highest SHAP values and lee wave occurrence (bottom of each pair of subplots). The PDF of that variable is also shown by region (top).

regions than the horizontal wind speeds. The virtual potential temperature, and changes in the wind direction have lower SHAP values than the horizontal wind speeds on pressure surfaces below 500 hPa. While SHAP is not a perfect measure of understanding the importance of different features due to it sharing credit between variables, it does attempt to address this more than other methods such as permutation importance. It has also picked out the Scorer parameter as being important for lee waves, which is not unsurprising so the results from SHAP are not unexpected.

To further unpick this, distributions of the variables with the six highest SHAP values were obtained and plotted, along with the lee wave occurrence as a function of these variables, for each region. These are shown in Figure 5.12. In every case, the occurrence of lee waves does not reach 100% in any region outside the very tails of the distributions, and there is a notable difference between regions. This may be due to the regions (shown in Figure 5.1) not being fully covered by topography. However, additional tests (not shown) restricting to pixels where the model orography was > 100 m only partly reduced the differences between regions. These results suggest a combination of conditions are important for lee wave occurrence in different regions, and lee wave occurrence cannot be described by one other variable alone. However,

the differences in lee wave occurrence in different regions does relate to the orography in the regions: there are fewer lee waves in Dartmoor and South West Ireland, which have relatively small hills than the Scottish Highlands where there are much larger hills, and a higher occurrence of lee waves.

Different measures of the Scorer parameter (Figure 5.12 [a]–[c]) show a relatively steep gradient (particularly the bulk measure in [a]) around $10^{-3}\mathrm{m}^{-1}$. The distribution of bulk Scorer parameters, mostly in the range of $10^{-3}\mathrm{m}^{-1}$ to $3 \times 10^{-4}\mathrm{m}^{-1}$ suggest that wavelengths in the range 5 km to 20 km are supported by the model, approximately in line with the range of wavelengths shown in Figure 5.5 (c) and Figure 5.7 (c). Values of the Scorer parameter higher than $10^{-3}\mathrm{m}^{-1}$ would approximately correspond to wavelengths shorter than 5 km, which cannot be resolved in the model anyway (World Meteorological Organization, 1993). The drop off in lee wave occurrence after this point may suggest that any lee waves that are present may not be resolved in the model, however these small-wavelength waves are likely to be of small amplitude. The horizontal wind speed at 850 hPa (Figure 5.12 [d]) and 925 hPa (Figure 5.12 [e]) also shows a cut-off between wave occurrence and increasing wind speed, again with regional differences. There is no such change in lee wave occurrence for the Brunt-Väisälä frequency in Figure 5.12 (f), suggesting that it is changing horizontal wind speed rather than the Brunt-Väisälä frequency that tends to control the Scorer parameter for lee wave occurrence in Britain and Ireland.

5.3.5 Relationship between lee wave amplitude and horizontal wind speed

Figure 5.13 shows a comparison between the amplitude of lee waves and the 750 hPa wind speeds for 1982–1987, for coarse-grained data (8×8 pixel regions as described in Section 5.2.2). This shorter subset of the full climatology data was obtained to save on disk space and time consumed pulling data from MASS. The 750 hPa surface was chosen as it is one of the closest pressure surfaces available to the data used to predict the amplitudes (700 hPa). However, repeating the analysis using the horizontal wind speeds on the 650 hPa surface and the 10 m wind speeds produced similar histograms to those in Figure 5.13.

The raw data shown in Figure 5.13 (a) is dominated by the high number of cases of small amplitude lee waves ($< 1\,\mathrm{m\,s}^{-1}$), with a large spread of surface wind speeds

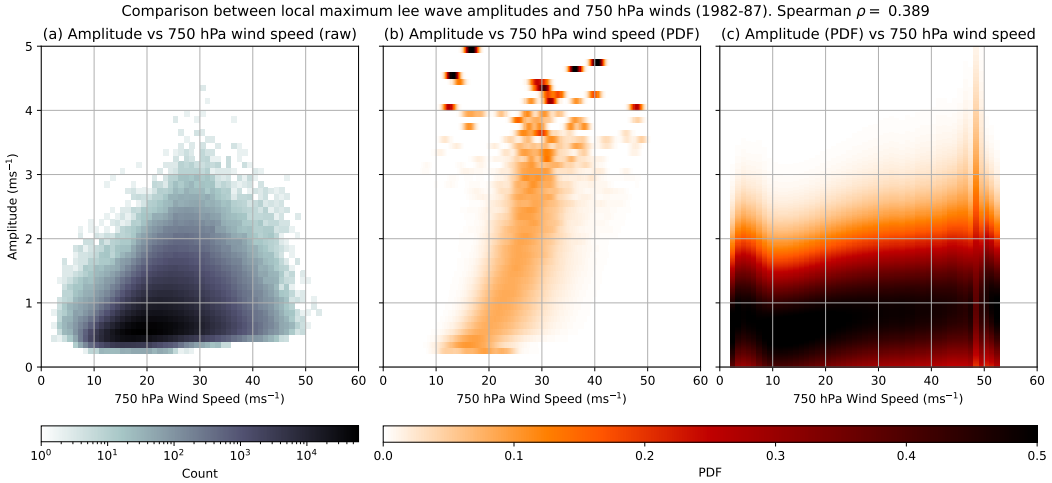


Figure 5.13: Histograms to compare between the local (17.6 km × 17.6 km, or 8 × 8 grid-points) maximal amplitude and 750 hPa wind speed for 1982–1987 over Britain and Ireland. (a) shows the raw data, while (b) and (c) are normalised: by amplitude (b: each row shows a probability density function (PDF) of 750 hPa wind speeds for each amplitude), and by wind speed (c: each column shows a PDF of amplitudes for each 750 hPa wind speed) .

at lower amplitudes. The Spearman rank correlation coefficient is $\rho = 0.389$, which does not show a strong correlation between the amplitude of lee waves and the 750 hPa wind speed. When normalising by amplitude in (b), there is a positive trend where stronger lee waves have stronger 750 hPa wind speeds associated with them. For example, the modal 750 hPa wind speed for lee waves with an amplitude of $1\,\mathrm{m\,s^{-1}}$ is $\approx 22\,\mathrm{m\,s^{-1}}$, and the modal 750 hPa wind speed for lee waves with an amplitude of $3\,\mathrm{m\,s^{-1}}$ is $\approx 30\,\mathrm{m\,s^{-1}}$. However, when normalising by the surface wind speed in (c), there are a similar distribution of lee wave amplitudes for each wind speed. This shows that high wind speeds are necessary for strong amplitude lee waves, but certainly not sufficient, with the majority of events of high horizontal wind speeds on the 750 hPa still corresponding to low amplitude lee waves.

5.4 Conclusions

This chapter presents a new climatology of lee waves, produced using machine learning models trained to detect and characterise lee waves from NWP model output. This study builds on previous work to investigate lee waves over Britain and Ireland (such as Vosper et al., 2013; Worthington, 2006) in several ways. This study creates a climatology over a longer period (1982–2012) than Vosper et al. (2013)

and Worthington (2006), with full coverage over the entirety of Britain and Ireland. This study uses data from a more sophisticated NWP model than the dry, linear 3DVOM model used in Vosper et al. (2013). The machine learning techniques used to develop the climatology of lee waves detect and characterise lee waves resolved by the NWP model, rather than using peak wave amplitudes. While the findings in this chapter support the conclusions of previous studies, this work goes further, for example investigating the relationship between lee waves, their characteristics, and weather patterns as well as exploring which underlying meteorological variables are driving the variability in lee waves.

This chapter verifies some of the findings from Vosper et al. (2013): for example both studies found that lee waves are relatively common in the Scottish Highlands, the north of England/Pennines and north Wales; that lee waves occur more often in the winter months than the summer (Figure 5.2); amplitudes stronger than $3\,\mathrm{m\,s^{-1}}$ occur rarely (Figure 5.7 [b]); and that the strongest amplitude lee waves occur during westerly flow: which corresponds to the strongest amplitude lee waves produced in NAO+ and Southwesterly regimes (Figures 5.7 [b]). In addition, both pieces of work show that the amplitude of lee waves increased with increasing horizontal wind speed (Figure 5.13).

The work in this chapter did not find any evidence within the UKCP18 Local data to suggest a diurnal cycle in trapped lee waves or their characteristics, consistent with the findings of Worthington (2006). The analysis presented here expands on that from Worthington (2006) to cover a much longer time period (1982–2012) than they used, over a larger region (the entirety of Britain and Ireland compared to a region centred around central Wales and one VHF radar). This lack of a diurnal cycle of lee waves is not necessarily replicated elsewhere in the world: Ruff and Ólafsson (2019) attribute observed diurnal changes in downslope wind storms to gravity waves. It may also be that any diurnal cycle may be too small to be captured by the NWP model, although the lee waves simulated in the output of the MetUM used here are generally in good agreement with observations, as shown by Sheridan et al. (2017). Another possibility is that the impacts of the diurnal cycle are most significant in a shallow layer near the surface and so they only impact on the shortest wavelengths of gravity waves which are not resolved in the model data used.

Lee waves occur more often in autumn and winter months than in the spring and

summer months: this is likely due to the prevalence of weather patterns more conducive to strong winds in the winter than the summer. The NAO+ and Southwesterly weather patterns, which have some of the highest prevalence of lee waves across the country (Figure 5.6), occur more often in the autumn and winter (40% of days) than the spring and summer (33% of days) during the climatology period 1982–2012. The weather patterns have more influence on lee waves than any diurnal effects: with relatively few lee waves being produced during NAO− and Scandinavian high conditions, compared to NAO+ or Southwesterly conditions. In general, the frequency of occurrence of lee waves is determined by the orography: there is correlation (Figure 5.3) between the local orography and whether lee waves occur. However, the geographic pattern of lee wave occurrence changes depending on the weather pattern. The spatial difference in lee waves under weather patterns is likely due to the relative strong winds experienced in different locations; for example the greater occurrence of lee waves over Scotland compared to (for example) Wales in Figure 5.6 under Azores high conditions and Southwesterly or NAO+ conditions are likely due to Scotland experiencing higher wind speeds under Azores high conditions than Wales, and more equal wind speeds between the two under Southwesterly or NAO+ regimes. Hence weather patterns with stronger wind speeds experience more conditions conducive to trapped lee waves.

There is some variation in lee wave amplitude depending on the weather pattern: while the amplitude of most lee waves is less than $1\,\mathrm{m\,s^{-1}}$ (Figure 5.7 [b]), the mean amplitudes in NAO+ conditions are stronger than the mean of the complete data set over most of Britain and Ireland. Wavelengths of lee waves are the longest under NAO+ conditions with Southwesterly conditions exhibiting a similar distribution of wavelengths, with a larger spread than the other patterns. Wavelengths under all the other weather conditions present a smaller distribution of wavelengths (as in Figure 5.7 [c]), but with some differences as to where lee waves are located (Figure 5.9). For example, under NAO− conditions, lee waves tend to be confined to Scotland, portions of Ireland and the west coast of England and Wales. In contrast, lee waves in other patterns (such as Northwesterly, Low close to the UK, and Azores high) are more likely to propagate further, such as over the Pennines and towards the eastern coast of Britain. The orientation of lee waves is broadly in line with the synoptic wind direction (Figure 5.10), though with some large spread within the data depending on the regime, but this is likely due to the multimodal distribution of

the orientation under some weather patterns. The Northwesterly and Southwesterly patterns have lee waves oriented north-west/south-east and south-west/north-east respectively. In Figure 5.7 (d), the orientation of lee waves under these two patterns have a unimodal distribution, while the lee wave orientations in the rest of patterns tend to exhibit more bimodal behaviour (such as that seen for lee waves in NAO+ conditions).

When predicting the occurrence of lee waves based on other meteorological variables, the Scorer parameter, horizontal wind speed and Brunt-Väisälä frequency came out as being more important (having larger SHAP values) than the directional wind shear, virtual potential temperature or the month of the year (despite there being a seasonal cycle of lee waves). There is a height dependence, with surfaces around and below where the lee waves were originally predicted (700 hPa) being more important than those well above (compare the SHAP values for horizontal wind speeds at 850 hPa and 925 hPa with 200 hPa and 300 hPa in Figure 5.11). The smaller SHAP values for the month of the year and weather patterns are likely due to the smaller variations within these values. They do affect the generation of lee waves, as seen in Figures 5.2 and 5.6, but the variation of other meteorological conditions is more important. For example, the weather patterns drive the synoptic conditions which result in more conducive conditions for lee waves.

The relationship between lee wave occurrence is shown for the Scorer parameter and the horizontal wind speed, but not the Brunt-Väisälä frequency in Figure 5.12, suggesting that the horizontal wind speed tends to control the Scorer parameter rather than the stability. However, there is a variation in occurrence of lee waves between regions for these variables, and the cut-offs in the distributions of the Scorer parameter correspond to reasonable wavelength lee waves.

The relationship between amplitude and horizontal wind speeds is explored in Figure 5.13. However, despite a dominance of smaller amplitudes in the data, there is correlation between the 750 hPa wind speeds and the amplitude of the lee waves at 700 hPa, when the data is normalised by lee wave amplitude, but no such similar correlation when normalising by the 750 hPa wind speed in Figure 5.13. This shows that strong wind speeds are a necessary but not sufficient indicator of high amplitude lee waves.

The next chapter will explore the application of these trained machine learning models to the future time slices of the UKCP18 Local data to examine the effects, if any, on lee waves over Britain and Ireland under climate change.

Chapter 6

How do lee waves change under a future climate scenario?

6.1 Introduction

This chapter investigates how the frequency of occurrence and characteristics of lee waves over Britain and Ireland may change in the future, by using the high resolution regional climate simulations from UKCP18 under the RCP 8.5 scenario and expanding on the work of Chapters 4 and 5. Under climate change, the number of years with severe storms affecting Britain and Ireland is projected (in the UKCP18 projections) to increase (Bloomfield et al., 2024). In addition, Manning et al. (2023) found the severity of windstorms were projected to increase in a future climate (for example, a current one in ten year event is projected to be a one in five or six year event under RCP 8.5). These projected changes in occurrence of high wind speed events could mean more occurrences of high amplitude lee wave events and therefore more rotor events in a future climate, due to the changing distribution of strong wind speeds (as discussed in Section 5.3.5).

The machine learning models from Chapter 4 are used to produce data from UKCP18 on the frequency of occurrence and characteristics of lee waves under RCP 8.5. Chapter 5 used data from the first time slice of the UKCP18 Local data, TS1 (1982–2012), where ERA-Interim re-analysis was used to drive the CPM to produce the data for the present-day climate. This chapter uses the output from the UKCP18 Local simulations for the future climate time slices, TS2 (2021–2041) & TS3 (2060–

2080).

This chapter is not intended to be a repeat of the methods in Chapter 5 but for future climate data, but instead to highlight how lee waves may change from the present climate to a future, high emissions scenario, climate. The physics is not changing: the formation of lee waves still depends on the Scorer parameter decreasing with height. However, changes in the synoptic meteorology in a future climate scenario may have an impact on how often lee waves are generated, or the characteristics of lee waves in this future scenario. For example, a projected increase in the occurrence of high amplitude ($> 3\,\mathrm{m\,s^{-1}}$) lee waves may strongly impact aviation activity through the increased risk of rotors and turbulence under RCP 8.5. In addition, the future climate data is an ensemble, so the uncertainty within the predictions can also be assessed.

An overview of the methodology is given in Section 6.2, the results are discussed in Section 6.3 and some concluding remarks are given in Section 6.4.

6.2 Methodology

In the same vein as in Chapter 5, lee waves over Britain and Ireland were identified from UKCP18 Local data, specifically the vertical velocities on the 700 hPa surface. The vertical velocity data, from each of the 12 ensemble members for TS2 and TS3, were acquired from MASS. The machine learning models described in Chapter 4 were applied to the vertical velocity data to produce masks of lee wave occurrence, amplitude, wavelength and orientation for each of the ensemble members. From this, a climatology was built up of lee waves within each ensemble member for TS2 and TS3.

This large data set was analysed in several ways: by comparing the headline change in lee wave occurrence from TS1 to TS2 (2021–2041) and TS3 (2060–2080) using probability density functions (PDFs); changes within the diurnal cycle of lee wave occurrence and characteristics using PDFs and maps; and finally using the weather patterns (described in Section 2.4 and Section 3.3) from Neal et al. (2016) to examine changes to lee wave occurrence and characteristics by weather pattern, again using PDFs and maps. The weather patterns proved useful for examining the ef-

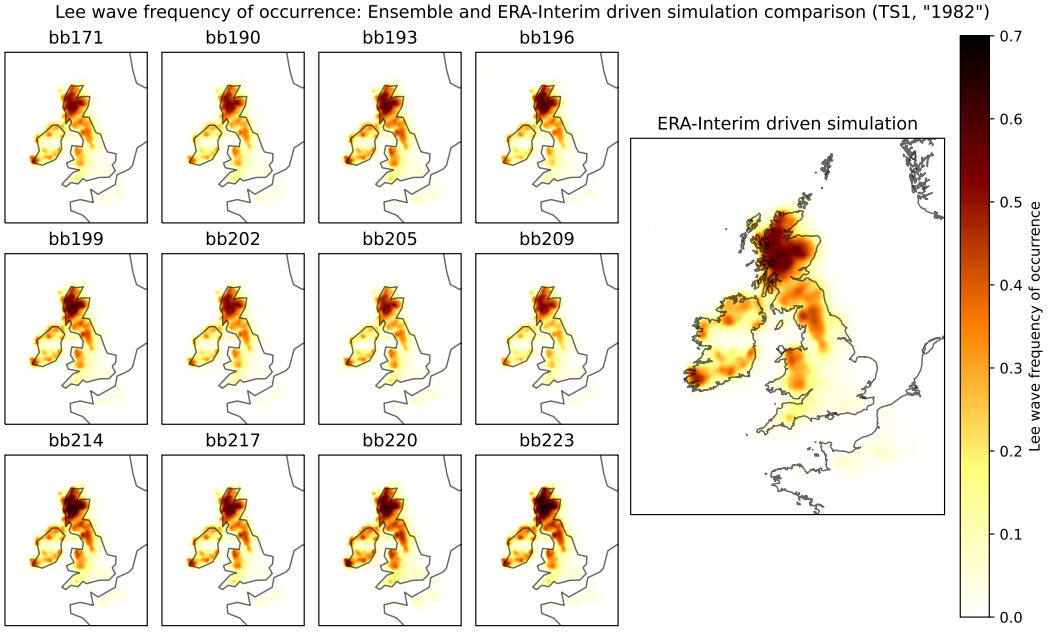


Figure 6.1: Comparison of lee wave frequency of occurrence between ERA-Interim driven simulation and the PPE for "1982".

fects of different synoptic weather conditions on the generation and characteristics of lee waves over Britain and Ireland for TS1 in Chapter 5. In this chapter, they are also employed to investigate how the synoptic weather changed within these patterns from TS1 to the future time slices, and whether modelled changes in lee wave characteristics are attributable to changes in the occurrence of particular weather patterns. The weather pattern data were available every day, for each ensemble member in the future climate simulations.

These distributions and map-style plots were, as in Chapter 5, produced by using the Python library `dask` to process the large number of masks (segmentation and characteristics) outputted by the machine learning models from Chapter 4. Together, these data will give an overview of how lee wave activity is projected to change over Britain and Ireland from the current climate to a future climate scenario under RCP 8.5.

6.2.1 Examining the differences between the ERA-Interim driven simulation and the PPE

Despite still using the same model setup, there is a difference between the present-day (ERA-Interim driven) and future (PPE) climate simulations because they are

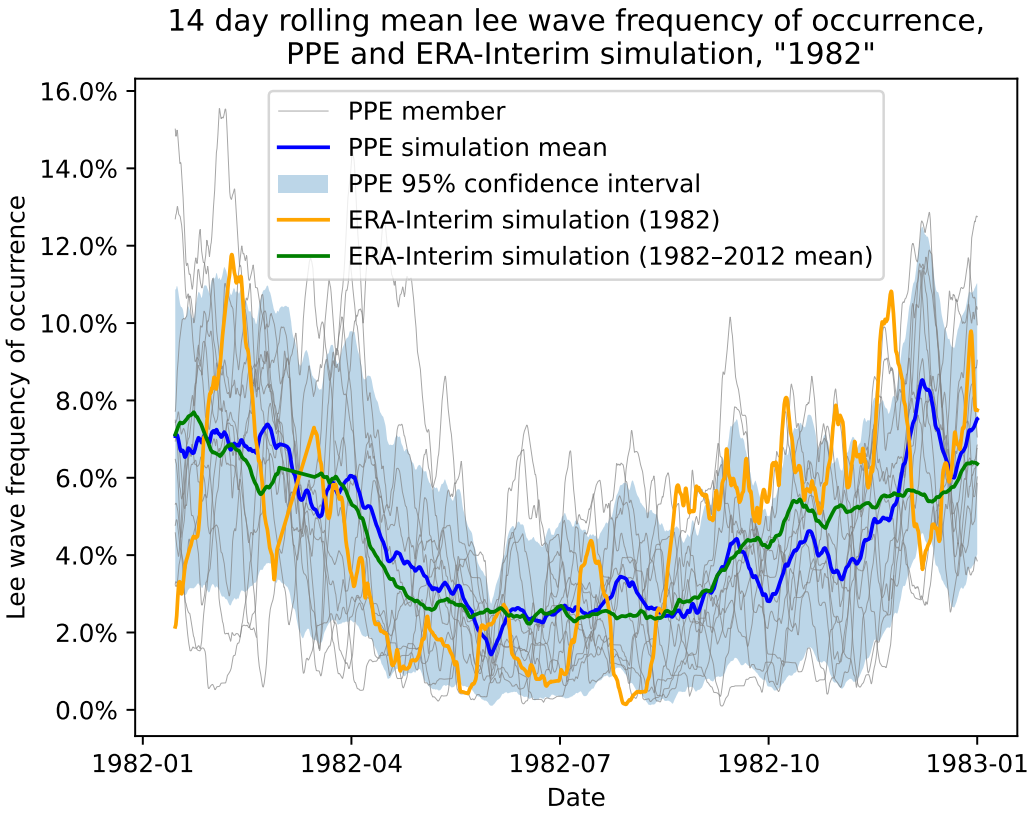


Figure 6.2: 14 day rolling mean lee wave frequency of occurrence for the PPE and the ERA-Interim driven simulation, for "1982".

driven by different boundary conditions, because there will be biases in the climate model in the PPE compared to the reanalysis-driven model used for TS1 in Chapter 5. To analyse this difference, a subset of the PPE data for TS1 was compared with the ERA-Interim driven ensemble. Figure 6.1 shows the lee wave frequency of occurrence between each of the 12 ensemble members in the PPE, for "1982" and the ERA-Interim driven simulation for 1982. The ERA-Interim driven simulation compares well against the other ensemble members, with a similar frequency of occurrence of lee waves occurring in the same (mountainous) areas in both the ERA-Interim driven simulation, and within the PPE. Figure 6.2 shows the lee wave frequency of occurrence for the PPE and the ERA-Interim driven simulation (for 1982, and the mean occurrence for the whole simulation). The seasonal cycle in lee wave frequency of occurrence is shown in both the ERA-Interim driven simulation, and the PPE. The ERA-Interim simulation in 1982 is within the 95% confidence interval of the ensemble mean 79.0% of the time and the ERA-Interim driven simulation mean (for the full TS1 period) is within the 95% confidence interval 100% of the time, suggesting a similar climatology in both the ERA-Interim driven simulation and the PPE for TS1. Hence, while care needs to be taken in comparing the ERA-Interim driven simulation with the PPE for the future climate simulations, both the ERA-Interim driven simulation and the PPE show similar occurrences in lee waves over Britain and Ireland throughout TS1, meaning that this comparison is appropriate.

6.3 Results

The overall projected changes in lee wave occurrence and characteristics from TS1 to TS2 and TS3 are small. Figure 6.3 shows PDFs for the lee wave frequency of occurrence and characteristics in the three time slices considered in this study: TS1 of the present-day climate, and TS2 and TS3 of the future climate under RCP 8.5. There is little change within the probability distributions from TS1 to the future climate scenarios for frequency of occurrence (a) and wavelength (c). For the amplitude (b), while the bulk of the distributions look very similar, at the stronger amplitudes within the tail of the distributions, there are cases of amplitudes stronger than $5\,\mathrm{m\,s^{-1}}$ in TS2 and TS3, which do not occur in the data for TS1. 7 of the 12 ensemble members in TS2 had PDFs containing amplitudes stronger than $5\,\mathrm{m\,s^{-1}}$,

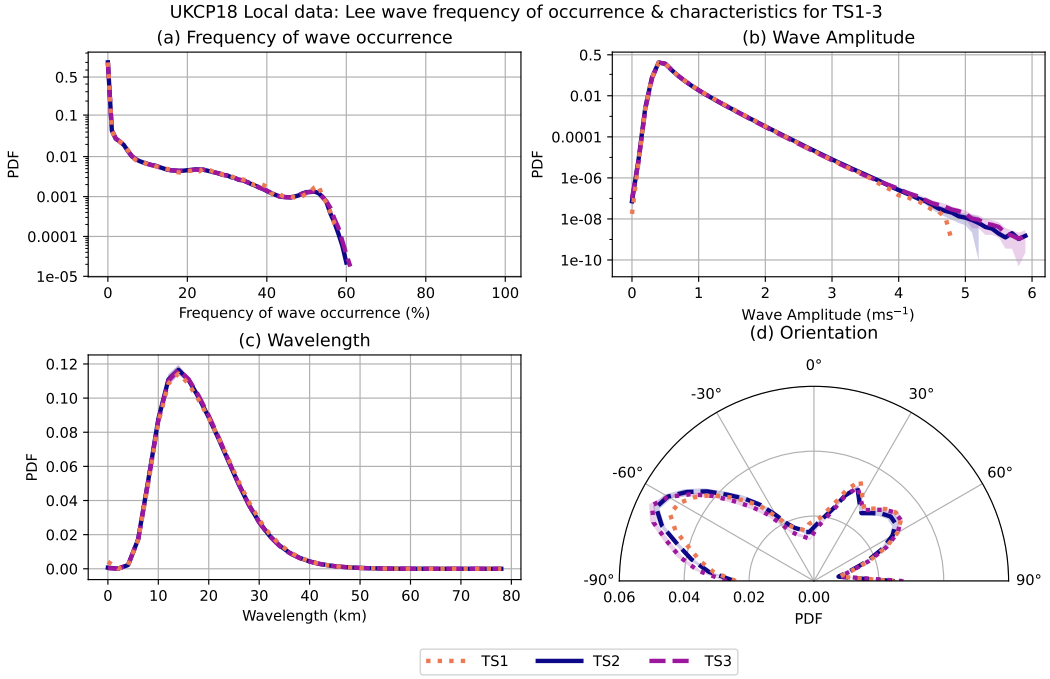


Figure 6.3: Headline PDFs of lee wave frequency of occurrence and characteristics for TS1 (1982–2012), TS2 (2021–2040) and TS3 (2060–2080). Shading shows the 95% confidence intervals for the ensemble mean for TS2 and TS3.

Table 6.1: Table comparing the percentage of the time (to 3 significant figures) where lee waves of amplitude exceeding $1\,\mathrm{m\,s^{-1}}$, $3\,\mathrm{m\,s^{-1}}$ and $5\,\mathrm{m\,s^{-1}}$ occurred somewhere in the domain over Britain and Ireland, for TS1, TS2 and TS3. For TS2 and TS3, the ensemble means are given as well as a 95% confidence interval from the ensemble.

Time Slice	**% of the time where lee waves of amplitude occurred (3 s.f.)**		
	$\geq 1\,\mathrm{m\,s^{-1}}$	$\geq 3\,\mathrm{m\,s^{-1}}$	$\geq 5\,\mathrm{m\,s^{-1}}$
TS1	51.0%	0.624%	0%
TS2	51.3% ± 1.49%	0.711% ± 0.110%	0.00389% ± 0.00201%
TS3	51.2% ± 1.65%	0.691% ± 0.0995%	0.00524% ± 0.00242%

while in TS3 the PDFs of 9 members contained amplitudes stronger than $5\,\mathrm{m\,s^{-1}}$. It should be emphasised that these cases of stronger amplitudes are very rare within the data, and while the majority (in both TS2 and TS3) of ensemble members contain these stronger amplitudes, the rest of the distribution changes little. There is a significant shift in orientation (d) from TS1 to TS2 and TS3, where there are more lee waves with a west-north-west/east-south-east orientation and fewer north/south oriented waves in the future scenario.

Table 6.1 shows how often lee waves of amplitude greater than $1\,\mathrm{m\,s^{-1}}$ and $3\,\mathrm{m\,s^{-1}}$ were produced somewhere in the UKCP18 Local domain (shown in Figure 3.1) for TS1, TS2 and TS3. For TS2 and TS3, the ensemble mean occurrences as well

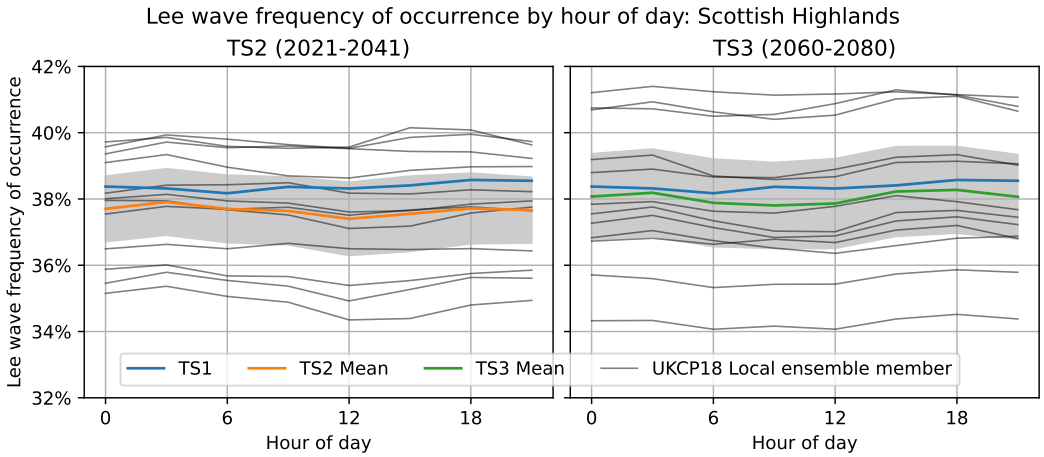


Figure 6.4: Mean diurnal cycle for each ensemble member for the two future time slices of lee wave frequency of occurrence, averaged over the Scottish Highlands, with the ensemble mean, and the present-day time slice overlain for comparison. The 95% confidence interval for the ensemble mean is shaded in grey.

as a 95% confidence interval are shown. While the ensemble mean occurrences of amplitudes stronger than $1\,\mathrm{m\,s^{-1}}$ and $3\,\mathrm{m\,s^{-1}}$ in TS2 and TS3 represent a slight increase from TS1, the 95% confidence interval of the ensemble in both TS2 and TS3 includes the values from TS1.

The following sections investigate any diurnal effects and effects due to weather patterns (Neal et al., 2016) on lee waves in the future climate scenario.

6.3.1 Diurnal effects

Figure 6.4 shows the mean diurnal cycle of lee wave frequency of occurrence for each ensemble member for TS2 and TS3, and the data from the present-day (TS1 [1982–2012]) climatology, in each case averaged across the Scottish Highlands region. There is little change between ensemble members at each time of day, and the current climatology (TS1) is well within the 95% confidence interval (grey shading) of the ensemble mean of the future climatology for both time slices. While the spread between ensemble members increases in TS3 from TS2, the ensemble mean frequency of occurrence of lee waves over the Scottish Highlands changes little from TS1 to TS2 and TS3, remaining within 1 percentage point of 38%.

Figure 6.5 shows PDFs for lee wave frequency of occurrence and characteristics by hour of the day, for TS3 as a proportion of the distributions in TS1 (i.e. as a

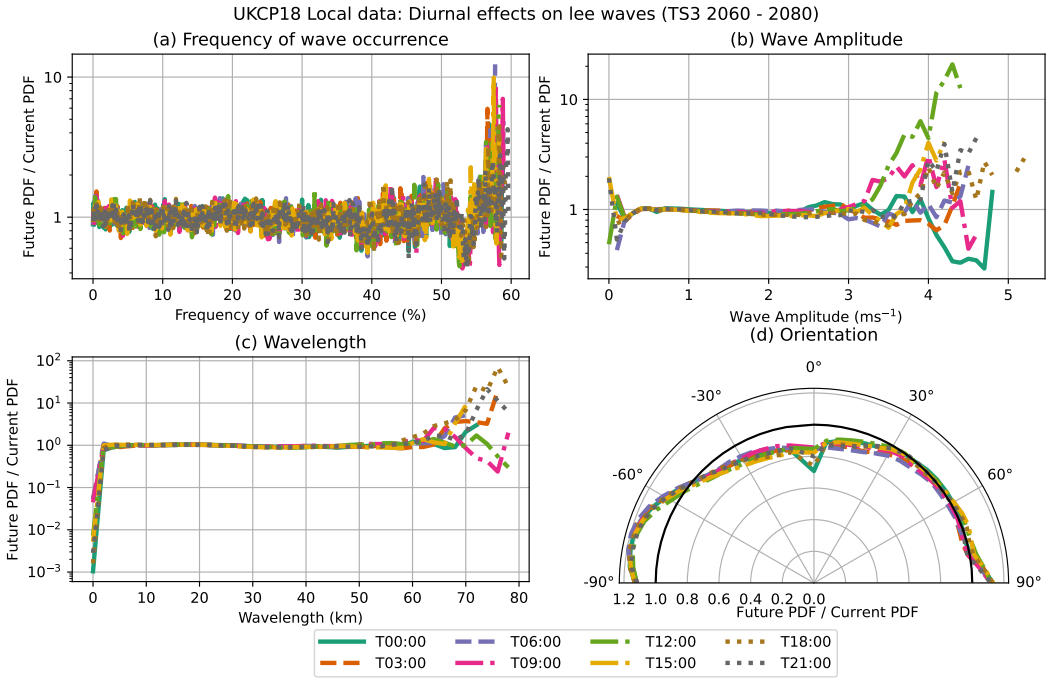


Figure 6.5: PDFs of lee wave frequency of occurrence (a) and characteristics (b)–(d) for TS3, by hour of the day as a proportion of the PDFs from the TS1 climatology.

proportion of the distributions plotted in Figure 5.5: the raw distributions for TS3 can be seen in Appendix A). In Figure 6.5, 1 would indicate no change from TS1 to TS3.

There is no systematic change in diurnal cycle of lee wave occurrence or characteristics from TS1 to TS3 (the plots for TS2 look very similar to this, showing a similar amount of change from TS1). The very small changes in amplitude (b) are at very small probabilities: lee waves of amplitude stronger than $3\,\mathrm{m\,s^{-1}}$ occur somewhere within the domain less than 1% of the time according to Table 6.1. The distributions of lee wave wavelength (c) and orientation (d) remain similar at every time during the day in the data, but the data overall shows a shift towards more east-west oriented lee waves in the future data from more north-westerly/south-easterly seen in TS1.

6.3.2 Weather patterns

Since the occurrence of the weather patterns changes from TS1 to TS2 and TS3, it was examined whether changes in lee wave frequency of occurrence were attributable to the changes in the weather patterns. Here we focus on TS3 since we expect the

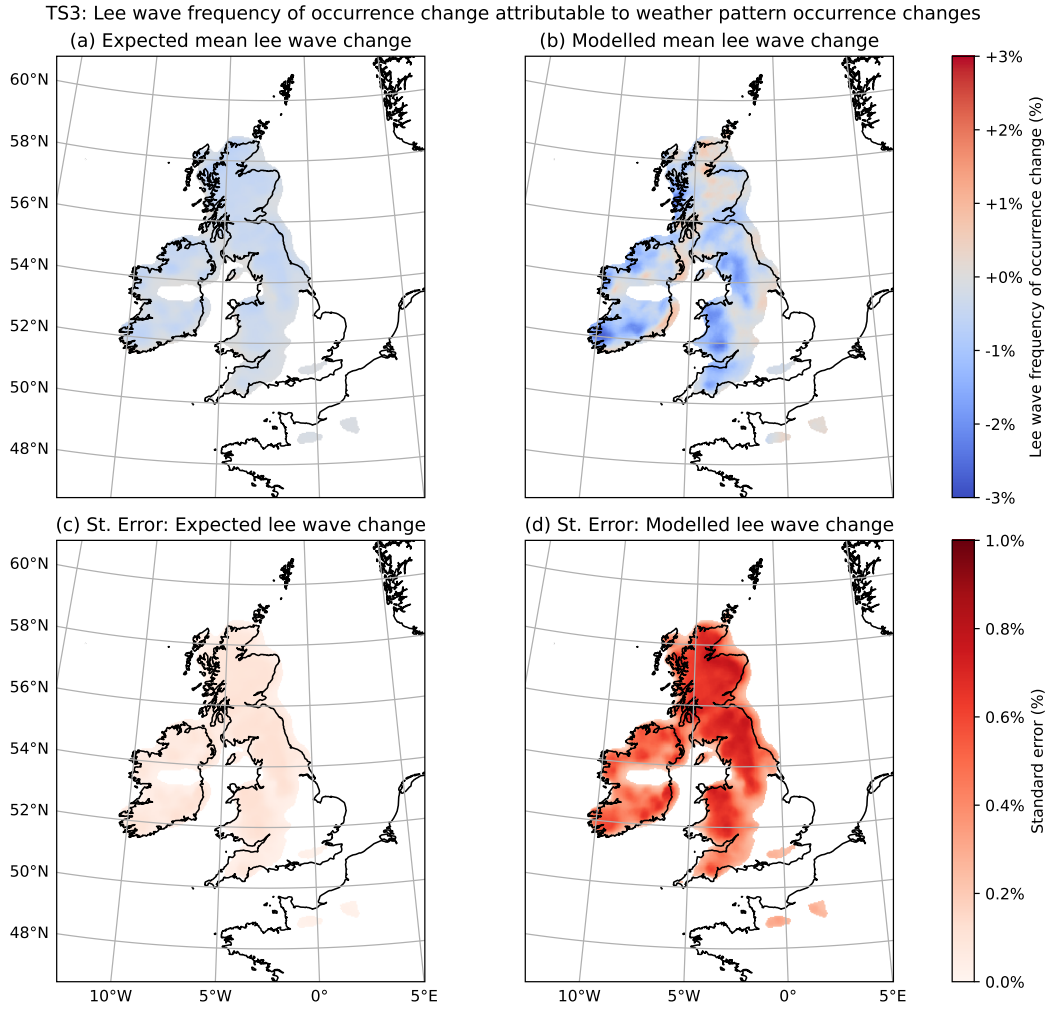


Figure 6.6: Maps showing the change in percentage points due to the changing prevalence of weather patterns in TS3, and the modelled change in lee wave prevalence in TS3. In (a), the expected change in lee wave occurrence if only the change in weather pattern occurrence, (b) the modelled change in lee wave occurrence from the UKCP18 Local data. In (c), the standard error on the ensemble mean is shown for just the weather pattern occurrence and (d) within the modelled data.

most significant differences to occur at the later time slice. The spatial distribution of lee waves under each weather pattern for TS1 were obtained (i.e. as shown in Figure 5.6). Then, by using the frequency of occurrence of each weather pattern in TS3, an expected overall map of lee wave occurrence was produced assuming the changes in lee wave occurrence depended only on the changing occurrence of the weather patterns from TS1 to TS3. For example, NAO+ conditions occurred 19.3% of the time in TS3, down from 20.1% in TS1. To produce this "expected" change map, the spatial pattern of NAO+ lee waves from TS1 were expected to occur 19.3% of the time in TS3.

Figure 6.6(a) shows what the change would be having adjusted the TS1 spatial

distribution of lee waves under the weather patterns to TS3 and (b) shows the ensemble mean lee wave frequency of occurrence change for TS3. The change in (a) is small, at most a decrease in lee wave occurrence of 1 percentage point. This contrasts with the modelled change in lee wave occurrence in (b), from the UKCP18 data. While the signs mostly agree (decrease in waves), the magnitude of the decrease is 3 percentage points in the hilliest regions. (c) and (d) show the standard error between the ensemble members. For (c), this only depends on the occurrence of the different weather patterns between ensemble members, and in (d), the variation in the physics of the different ensemble members, as well as the weather pattern occurrence.

Hence, the modelled change in lee wave occurrence in the UKCP18 Local data in TS3 is significantly more than that which can be explained by only a change in distribution of the weather patterns in the future climate data, and there must be changes in lee wave occurrence within the weather patterns too.

Figure 6.7 shows the PDFs of lee wave occurrence and characteristics by weather pattern, as a proportion of the TS1 climatology for TS2 and TS3. There is a much greater change within the weather patterns for occurrence and characteristics than by the time of day (Figure 6.5). The raw distributions (equivalents of Figure 5.7 for TS2 and TS3) can be seen in Appendix A. Figure 6.8 shows the change in occurrence and characteristics from TS2 to TS3, removing the effect (if any) from the differently driven models (ERA-Interim for TS1 and the PPE for TS2 and TS3).

The following subsections take each of these characteristics in turn.

6.3.2.1 Lee wave frequency of occurrence

Figure 6.9 shows how the lee wave frequency of occurrence changes from TS1 to TS2 and TS3 within the different weather patterns, compared to TS1. The most lee waves occur in the NAO+, Southwesterly, and the High pressure centred over UK patterns, and the least in NAO− and Scandinavian high, similar to in TS1. The biggest change in lee wave frequency of occurrence happens in the Azores high pattern, which only accounts for 4.3% (4.1%) of days in TS2 (TS3), the smallest share of the weather patterns. Under Azores high conditions, the frequency of occurrence of lee waves decreases by more than 10 percentage points in the Scottish

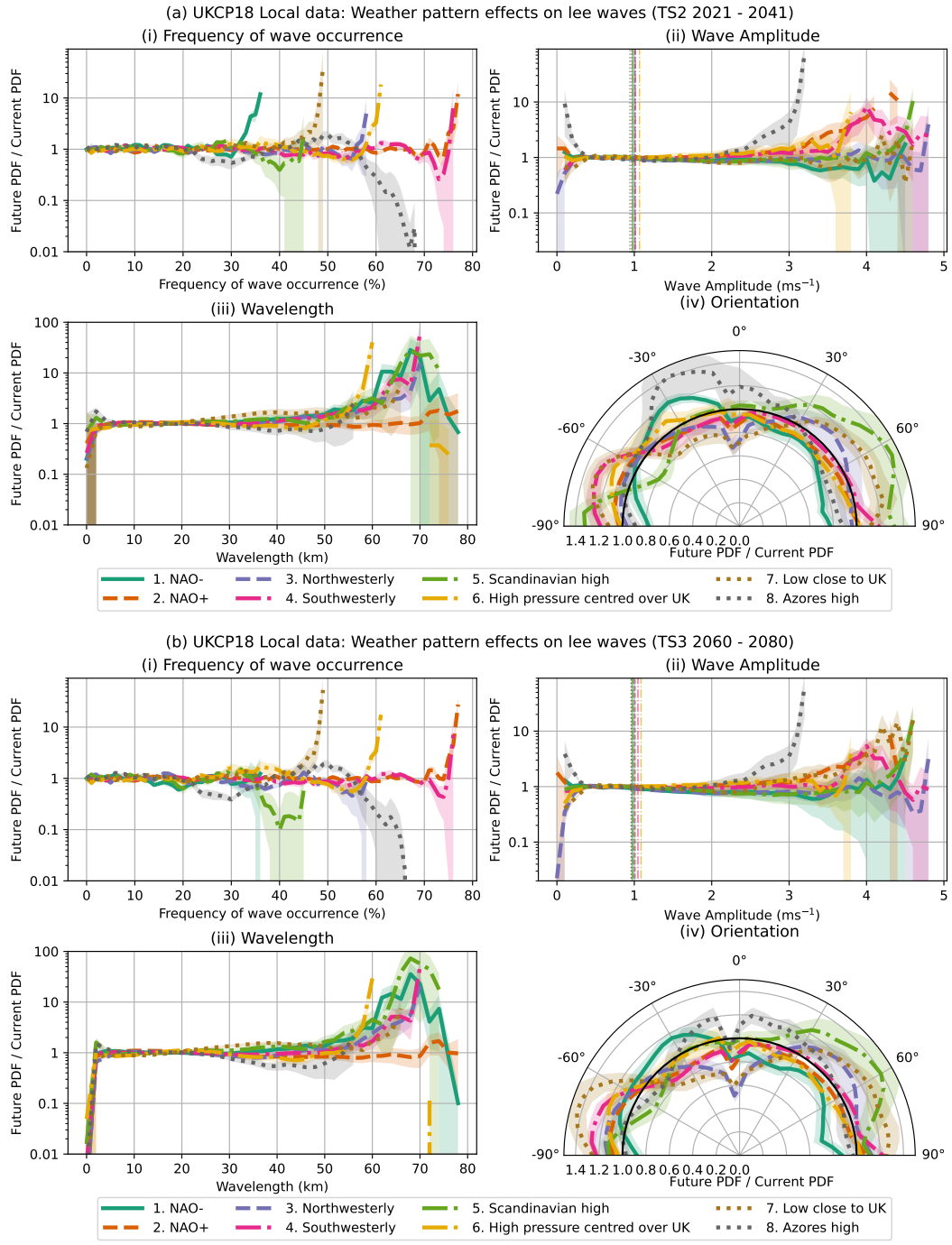


Figure 6.7: PDFs of lee wave frequency of occurrence (i) and characteristics (ii-iv) for TS2 (a) and TS3 (b), by weather pattern, as a proportion of the PDFs from the TS1 climatology. Values of 1 indicate no change from TS1 to the future. The 95th percentile amplitude of lee waves for each weather pattern in the time slice is marked as a vertical line in (ii).

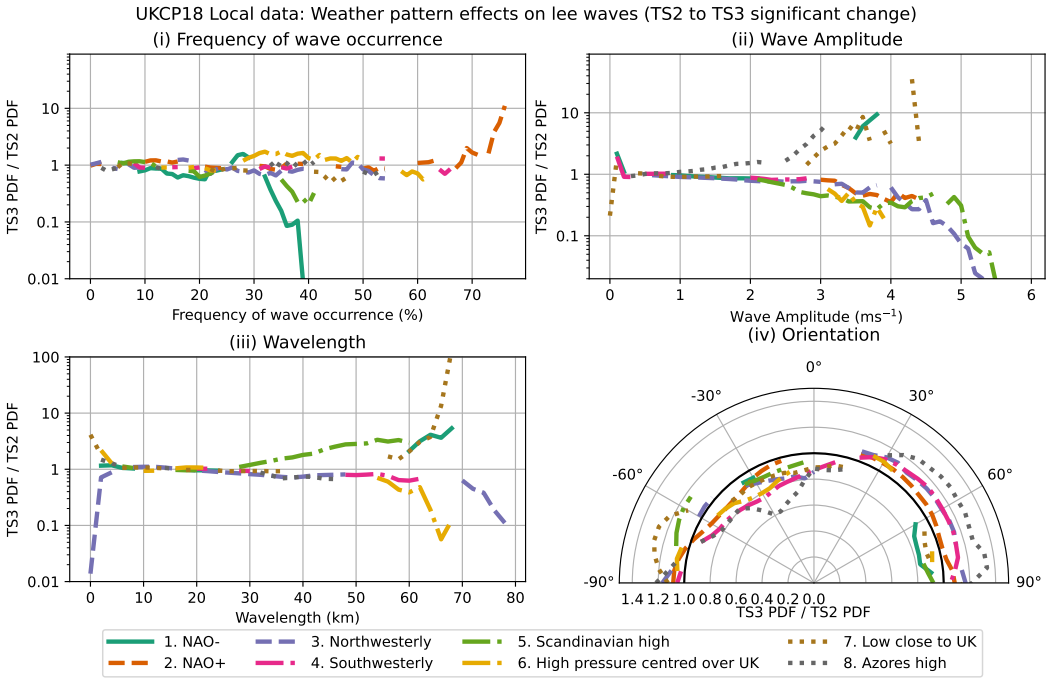


Figure 6.8: PDFs of lee wave frequency of occurrence (i) and characteristics (ii-iv) for TS3, by weather pattern, as a proportion of the PDFs from the TS2 climatology. Only sections where the change from TS2 to TS3 was significant (the entirety of the 95% confidence interval on the mean change had the same sign) are shown.

Highlands compared to TS1, in both TS2 and TS3. In TS1, the highest frequency of occurrence of lee waves occurred under the NAO+ pattern, and the frequency of occurrence of lee waves increases in TS3 compared to TS1. There are also more lee waves generated in TS3 during the High pressure centred over UK and Low close to UK weather patterns compared to TS1. From TS2 to TS3 in Figure 6.8(i), while there are significant changes projected in the occurrence of lee waves under different weather patterns, these are relatively small and show no systematic change across weather patterns. There is a projected increase in occurrence of lee waves under the NAO+ pattern from TS2 to TS3, which matches modelled increases in Figure 6.7. The changes from TS2 to TS3 in Figure 6.8 are smaller than those in Figure 6.7 for TS1 to TS2 and TS3.

Figure 6.9 also shows the projected change of MSLP contours within the weather patterns. Solid lines show the ensemble MSLP contours for the weather pattern, and dotted the MSLP contours for TS1. For example, the orientation of the MSLP under the High pressure over UK pattern are projected to change, as is the mean position of the low under the Low close to UK pattern. In Figure 6.7 (a & b [i]), the frequency of occurrence of lee waves in the future climate simulations remains

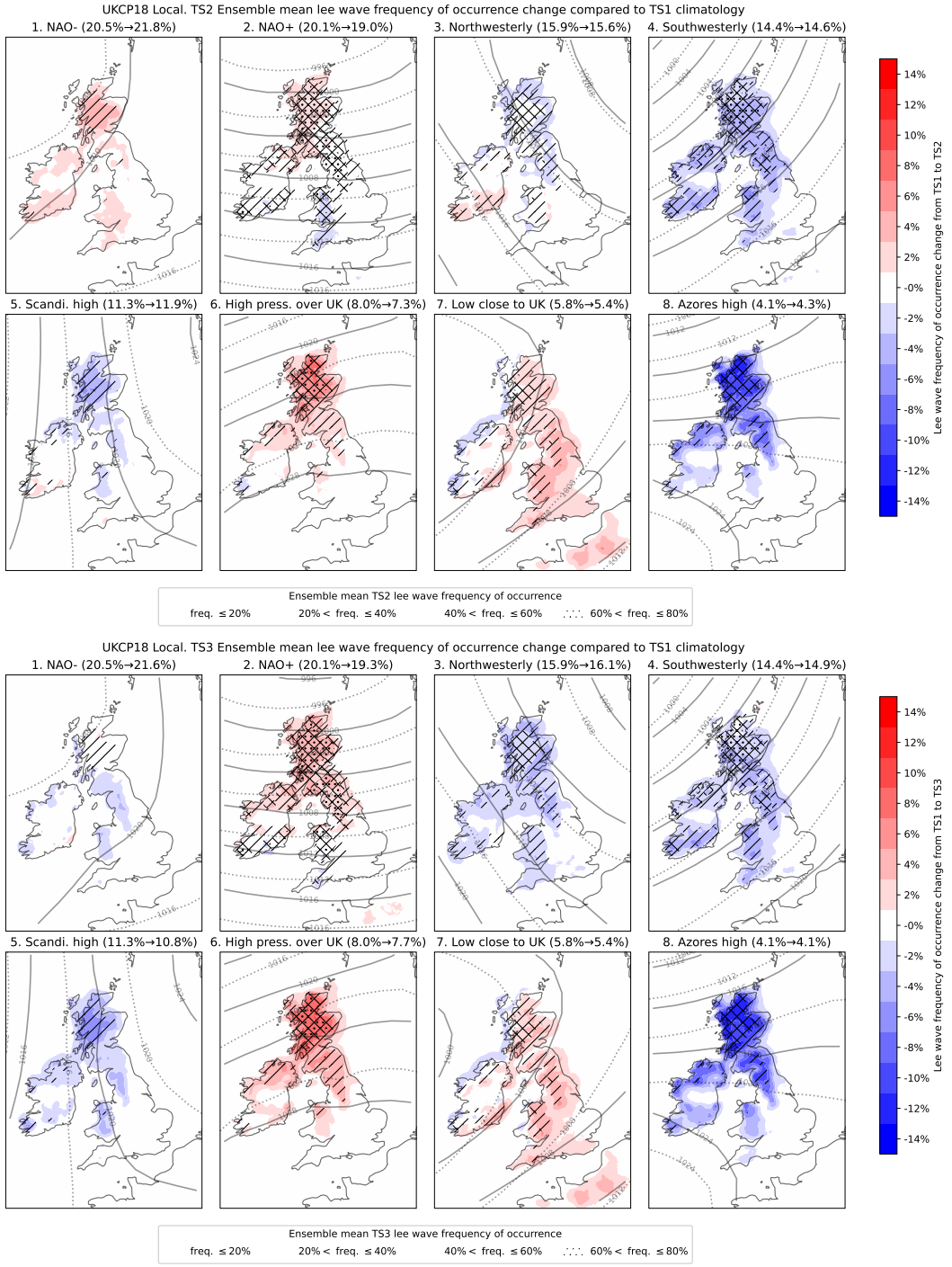


Figure 6.9: TS2 & TS3 ensemble mean change in lee wave frequency of occurrence compared to TS1, by weather pattern. Colour contours show the percentage point change in lee wave frequency of occurrence in TS2 & TS3 compared to TS1 (for example +2% here may represent a change from 20% in TS1 to 22% in TS2). Hatching shows the lee wave frequency of occurrence in TS2 or 3. The percentage in brackets after each pattern's name indicates how often that pattern occurs in that time slice, and how often it occurred in TS1 (TS1→TS2/3). Grey contour lines show the MSLP pattern for that period: dotted for TS1 and solid for TS2/3.

similar to TS1 for pixels with wave activity up to 30% of the time, and general increases in occurrence at the tail of the distribution for all the patterns except Azores high. In the Azores high pattern, there is a strong decrease in occurrence of lee waves from TS1 for both TS2 and TS3. This change could be attributed to the changing patterns of horizontal wind speeds from the present-day to future climate scenarios under these two patterns.

Figure 6.10 shows PDFs of 750 hPa horizontal wind speeds under the High pressure centred over UK (a) and Azores high (b) weather patterns for TS1 and each ensemble member in TS2 and TS3. Under the High pressure centred over UK pattern, the occurrence of wind speeds in the range $5\,\mathrm{m\,s^{-1}}$ to $15\,\mathrm{m\,s^{-1}}$ decreases in TS2 and TS3 from TS1. This decrease in occurrence of horizontal wind speeds in the range $8\,\mathrm{m\,s^{-1}}$ to $12\,\mathrm{m\,s^{-1}}$ is significant at the 95% confidence level from TS1 to TS3, but not from TS1 to TS2 or from TS2 to TS3. There is also higher occurrence of 750 hPa horizontal wind speeds stronger than $15\,\mathrm{m\,s^{-1}}$ in TS2 and particularly TS3, and this change from TS1 to TS3 for wind speeds in the range $25\,\mathrm{m\,s^{-1}}$ to $30\,\mathrm{m\,s^{-1}}$ is significant at the 95% confidence level. Under the Azores high pattern, the occurrence of 750 hPa wind speeds between $10\,\mathrm{m\,s^{-1}}$ and $20\,\mathrm{m\,s^{-1}}$ decreases from TS1 to TS2 and TS3, and the occurrence of 750 hPa wind speeds slower than $10\,\mathrm{m\,s^{-1}}$ increases in TS2 and TS3 from TS1. The decrease in occurrence of horizontal wind speeds in the range $10\,\mathrm{m\,s^{-1}}$ to $14\,\mathrm{m\,s^{-1}}$ from TS1 to TS2 and TS3 is significant at the 95% confidence level.

The modelled changes in lee wave frequencies of occurrence under these two patterns in Figure 6.7 (a & b [i]) and Figure 6.9 suggest that the shift to slower horizontal wind speeds in the Azores high pattern may be linked to the decrease in occurrence of lee waves under Azores high conditions. The pattern is less clear for the High pressure centred over UK pattern: there is still a decrease from TS1 to TS2 and TS3 in occurrence of 750 hPa horizontal wind speeds between 5 and $15\,\mathrm{m\,s^{-1}}$, but an increase in 750 hPa horizontal wind speeds faster than $15\,\mathrm{m\,s^{-1}}$ could be linked to the modelled increase in lee wave occurrence under the High pressure centred over UK pattern. Lee wave generation depends on the Scorer parameter, of which the horizontal wind speed contributes: so there could be changes in stability occurring too.

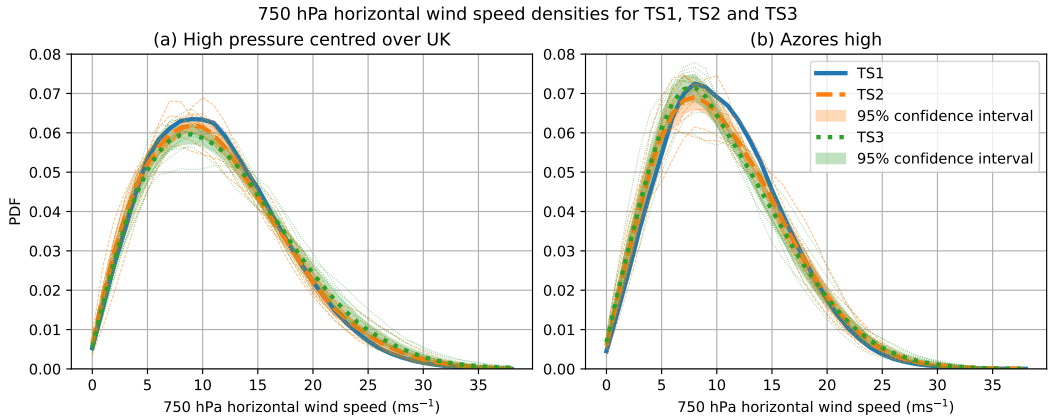


Figure 6.10: PDFs showing the 750 hPa wind speed under TS1, TS2 and TS3 for the High pressure centred over UK (a) and Azores high (b) weather patterns. For TS2 and TS3, each ensemble member is plotted (thin lines), along with the ensemble means (bold lines) and the 95% confidence interval on the mean (shading).

6.3.2.2 Amplitude

Figure 6.7 (a & b [ii]) shows the change in PDFs of the lee wave amplitude for each weather pattern from TS1 to TS2 and TS3, with the 95th percentile lee wave amplitude in TS2 and TS3 for each weather pattern shown. The amount of change from TS1 to TS2 and TS3 is similar, but with large uncertainties at amplitudes stronger than $3\,\mathrm{m\,s^{-1}}$. The cases of lee wave amplitudes stronger than $2\,\mathrm{m\,s^{-1}}$ increase for the NAO+, Southwesterly, High pressure centred over UK and Azores high patterns, and decrease for the NAO− regime. The Low close to the UK pattern shows a decrease in amplitude between 2 and $4\,\mathrm{m\,s^{-1}}$ for TS2 and an increase for TS3 compared to TS1. The pattern is unclear for amplitudes stronger than $3\,\mathrm{m\,s^{-1}}$ in general – but cases of lee wave amplitudes stronger than $3\,\mathrm{m\,s^{-1}}$ occur rarely (Table 6.1) so care must be taken to not read too much into the data. Figure 6.8 (ii) shows the significant (at the 95% confidence level) change in lee wave amplitude by weather pattern from TS2 to TS3. There are more modelled cases of lee wave amplitudes stronger than $2\,\mathrm{m\,s^{-1}}$ in the Azores high pattern and Low close to the UK pattern in TS3 than TS2, with the other patterns showing similar, but significant decreases in amplitudes stronger than $2\,\mathrm{m\,s^{-1}}$ from TS2 to TS3.

Figure 6.11 shows maps displaying the projected change in the ensemble mean 95th percentile amplitudes for the 8 weather patterns for TS2 (a) and TS3 (b), where this change is significant at the 95% confidence level. As Figure 6.7 (a & b [ii]) shows, for each weather pattern, the 95th percentile lee wave amplitude is approximately

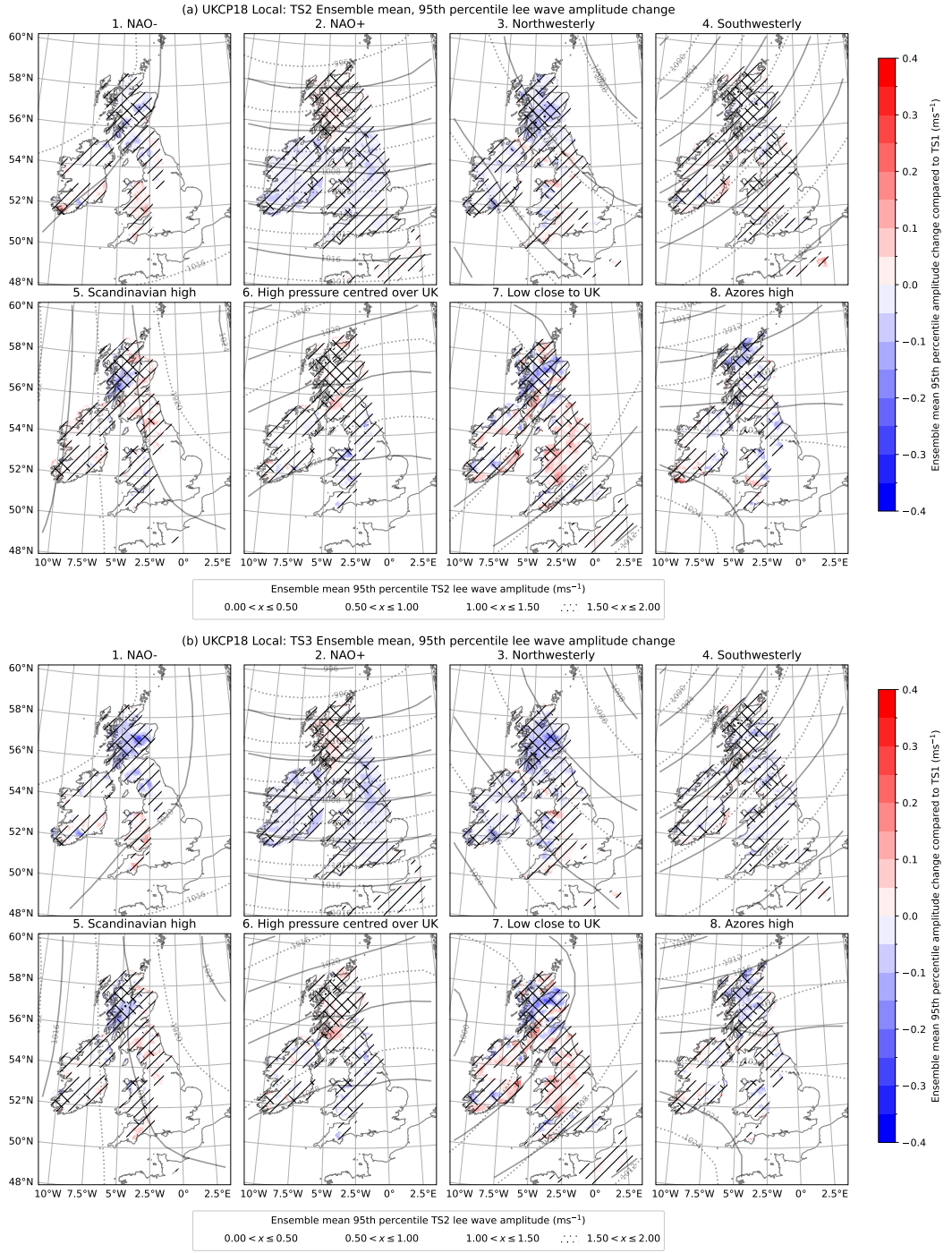


Figure 6.11: The change in the ensemble mean 95th percentile lee wave amplitude by weather pattern for TS2 (a) and TS3 (b). Colour contours show the change (where significant at the 95% confidence level) in the amplitude from the TS1 climatology, and hatching shows the respective amplitude for the future climatology period. As in Figure 6.9, the solid contours show the MSLP pattern for TS2/3, and dotted for TS1.

$1\,\mathrm{m\,s^{-1}}$, so the projected changes to higher amplitude lee waves are at the tail of the amplitude distribution, far beyond the 95th percentile as shown in Figure 6.11. Under each weather pattern, the 95th percentile amplitude decreases over Scotland, except for a small but significant increase in NAO+ conditions in the West Highlands, and the far North Highlands in Low close to UK conditions. The Scottish Highlands are typically where the largest amplitude waves are seen, shown by dots within the hatches in Figure 6.11 (a) and (b). Over the rest of Britain and Ireland, the 95th percentile amplitude within each weather pattern decreases in most cases (all apart from Low close to UK conditions), especially by TS3, but there is no clear pattern to the changes in amplitude in TS2 and TS3. There are changes to the MSLP contours, both in orientation and gradient, which could affect the generation of lee waves at different amplitudes, and as shown in Chapter 5, the wind speed seemed to have more of an effect on amplitudes than the stability.

6.3.2.3 Wavelength

Figure 6.7 (a & b [iii]) shows PDFs of the ensemble mean lee wave wavelength for each weather pattern for TS2 and TS3, as a proportion of the TS1 climatology distribution. The distribution of wavelengths does not change much for wavelengths shorter than 50 km, compared to TS1, which covers most cases of lee waves. For very long wavelength ($> 50\,\mathrm{km}$) lee waves, the change is similar for both TS2 and TS3, with all patterns showing an increase compared to TS1 in the future. The change in wavelength under NAO+ conditions is smaller than the rest of the patterns – but the longest wavelength lee waves in TS1 occurred under NAO+ conditions, as Figure 5.7 (c) shows, albeit there being few cases of wavelengths longer than 50 km overall. In Figure 6.8(iii), there is little significant projected change in wavelengths shorter than 50 km for all patterns except for Scandinavian high conditions which shows a significant increase in wavelengths between 30 km and 60 km. At wavelengths longer than 50 km, although rare, there is a significant projected increase in long wavelength waves under Low close to the UK conditions, and a decrease in long wavelength lee waves under High pressure centred over UK and Northwesterly conditions from TS2 to TS3.

Figure 6.12 shows spatially the change in the ensemble mean lee wave wavelength for each of the weather patterns for TS2 (a) and TS3 (b) from TS1 as colour contours.

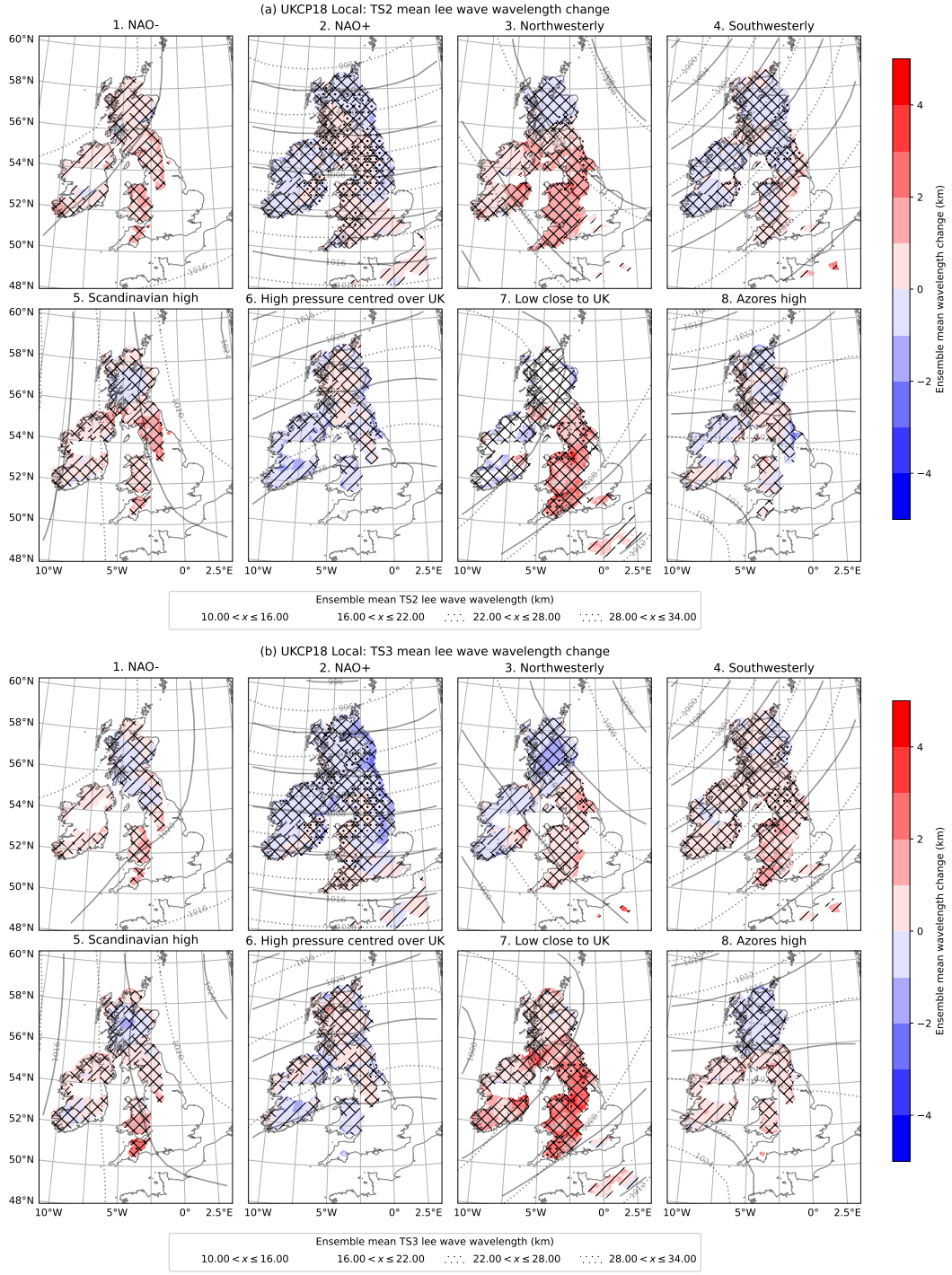


Figure 6.12: The change in mean lee wave wavelength from TS1 to TS2 (a) and TS3 (b) by weather pattern. Hatches show the ensemble mean wavelength for the respective time slice. As in Figure 6.9, the solid contours show the MSLP pattern for TS2/3, and dotted for TS1.

The respective mean lee wave wavelength for each weather pattern for that time slice is also shown as hatches. The mean wavelength changes by at most 5 km, with the largest changes seen in the Low close to the UK weather pattern, where the mean wavelength increases from TS1 to TS3. Despite the decrease in mean wavelength compared to TS1 under NAO+ conditions, the longest mean wavelengths still occur under NAO+ conditions in TS2 and TS3.

The change in wavelength from TS1 to TS2 and TS3 is mostly significant at the 95% confidence level, illustrated by the hatched areas for TS2 and TS3 in Figure 6.13. This shows that the changes in wavelength are consistent between ensemble members at the 95% confidence level, but there is no uniform pattern across Britain and Ireland geographically or between weather patterns. This change in wavelength may be attributable to the changing distribution of the Scorer parameter under different weather patterns (due to the changing geostrophic winds from the present-day to future climate) affecting the wavelengths that lee waves are generated at in the future climate simulations.

6.3.2.4 Orientation

Figure 6.7 (a & b [iv]) shows the PDFs of lee wave orientations under the different weather patterns for TS2 (a) and TS3 (b) compared to TS1. The shift to more east/west orientations in TS2 and TS3 is again evident in most of the weather patterns – particularly the Low close to UK cases. Conversely, the NAO− pattern has a slight decrease in east/west oriented waves in the future simulations compared to TS1. Orientations in the NAO+, High pressure over UK and Northwesterly patterns remain similar to those in TS1. In TS2, the orientations of lee waves in Azores high conditions gets more northwesterly, but this pattern is not continued in TS3. This pattern is also shown in Figure 6.8(iv), where the orientations of lee waves are shown to be significantly more east-west oriented from TS2 to TS3 in NAO+, Southwesterly, Northwesterly and Azores high conditions, but less so under NAO−. The remaining patterns do not show a significant decrease in east-west oriented waves from TS2 to TS3.

Figure 6.14 shows the mean orientation by weather pattern for TS2 (a) and TS3 (b), with the mean orientations from TS1 shown as semi-transparent arrows. Any area

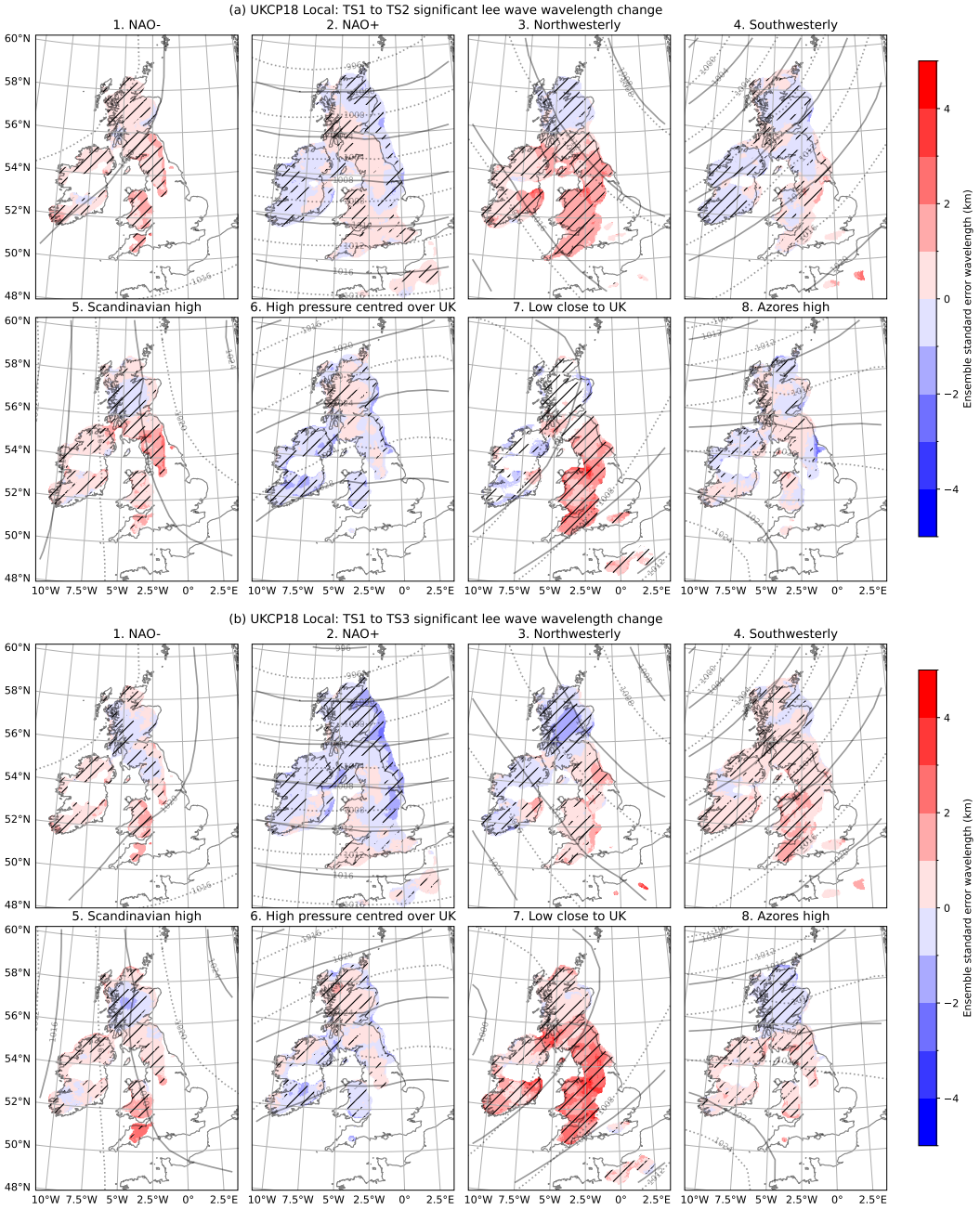


Figure 6.13: Maps showing only the significant wavelength change from TS1 to the future climate (95% confidence level). Colour contours as in Figure 6.12, but hatching showing locations of significant change.

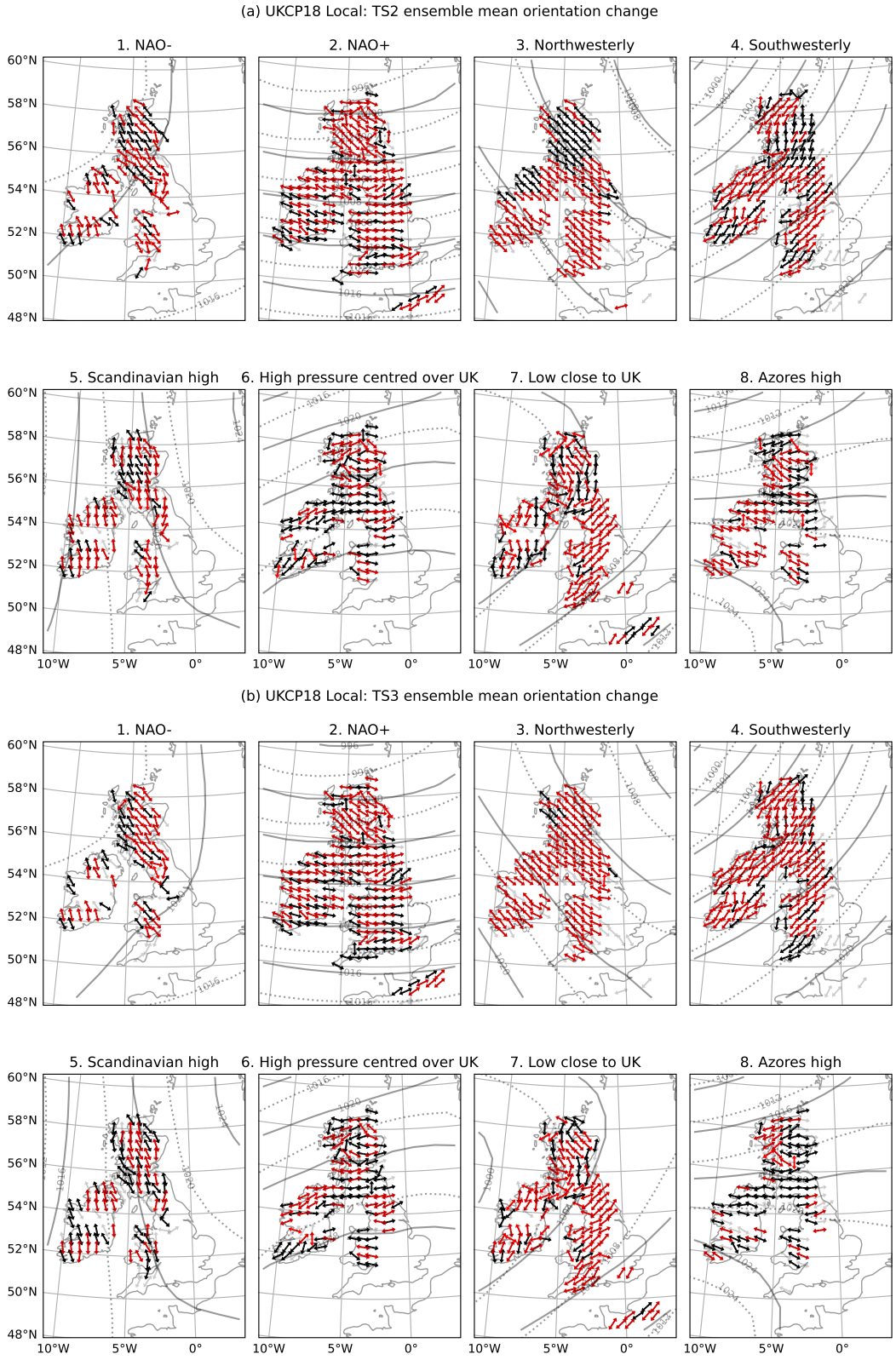


Figure 6.14: Maps showing mean lee wave orientation (the direction of wave propagation, which is perpendicular to the wave fronts) by weather pattern for TS2 (a) and TS3 (b). Red arrows indicate a significant change in orientation from present-day to future climate at the 95% confidence level. The semi-transparent overlay are the mean lee wave orientations by weather pattern from TS1 (as in Figure 5.10). Similarly to Figure 6.9, the solid contours show the MSLP pattern for TS2/3, and dotted for TS1.

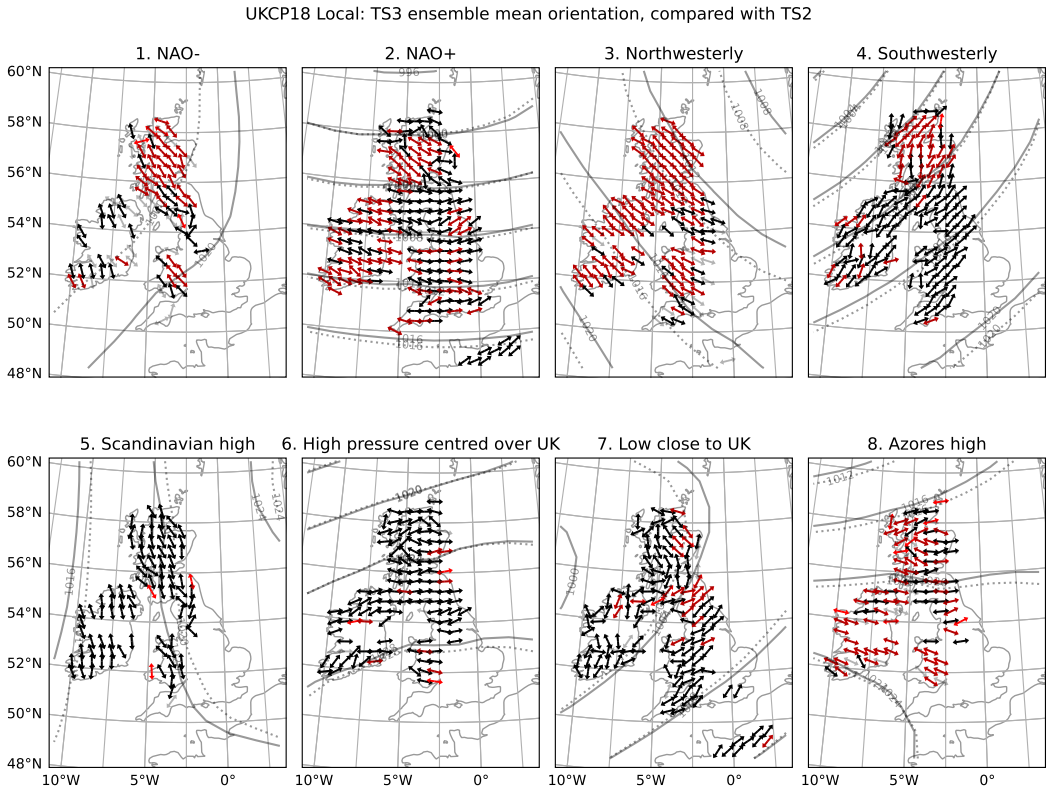


Figure 6.15: Maps showing the mean orientation (the direction of wave propagation, which is perpendicular to the wave fronts) by weather pattern for TS3, with TS2 shown as a semi-transparent underlay. Red arrows indicate a significant change in orientation from TS2 to TS3 at the 95% confidence level. The solid line contours show the MSLP pattern for TS3, and dotted TS2.

where this change from the present-day climate is significant at the 95% confidence level are shown as red arrows. As in TS1, the orientation of lee waves broadly follows the MSLP contours for that case. For example lee waves generated under the Southwesterly weather pattern generally propagate in a north east/south west direction. The mean lee wave orientation has not changed in any great measure within weather patterns from TS1 to TS2 and TS3: the greatest deviations are where there is a large standard error between ensemble members. There is also no clear large change in lee wave orientation within weather patterns from TS2 to TS3 (Figure 6.15). However, within the High pressure centred over UK and Azores high patterns, the mean wind direction (from the MSLP isobars) has shifted in the future simulations, and this is reflected in the mean orientations shown in Figure 6.14. There are small, but significant changes in mean lee wave orientation, even from TS2 to TS3 as shown in Figure 6.15, for example under the NAO+ and Southwesterly patterns in the Scottish Highlands, and the Northwesterly weather pattern over Ireland, Scotland, Wales and northern England. There is a change in the MSLP contours from TS2 TS3 under the Northwesterly weather pattern, perhaps explaining the significant changes in lee wave orientation under this pattern.

6.4 Conclusions

The ensemble mean frequency of occurrence of lee waves is modelled as decreasing slightly under RCP 8.5, however, the TS1 frequency of occurrence is well within the 95% confidence interval of the ensemble mean for TS2 and TS3 in the Scottish Highlands, and there were ensemble members with a modelled increase in lee wave occurrence over the Scottish Highlands (7 in TS2, and 5 in TS3 with an increased frequency of occurrence compared to TS1). From Figure 6.4, there is no significant change in the occurrence of lee waves from the present-day to the future climate. However, within different weather patterns, the occurrence and characteristics of lee waves are modelled as changing from the present-day to the future climate.

The occurrence of the 8 different weather patterns (from Neal et al., 2016) is modelled as changing from TS1 to TS2 and TS3. By using the geographical distribution of lee wave occurrence under these weather patterns from TS1, and applying them to the distribution of weather patterns under TS2 and TS3, the influence of weather

patterns alone (rather than the meteorology changing) was calculated. Figure 6.6 shows that the modelled variation between ensemble members of lee wave occurrence from TS1 to TS3 is greater than the variation of wave occurrence between ensemble members if the change in weather pattern distribution from the present-day to future climate was the only factor in the changing occurrence of lee waves from the present-day to future climate. In some cases the weather-pattern predicted change was the opposite sign to the UKCP18 future climate modelled change in lee waves, such as in the far north of Scotland. The modelled magnitude of change in lee wave occurrence in UKCP18 from TS1 to TS3 is greater than that attributable to the distribution of weather patterns changing. Hence, the meteorology within the weather patterns changes in the future simulations, resulting in different frequencies of lee wave occurrence, and as Figure 5.13 suggests this is likely to be because of changes in the Scorer parameter and the horizontal wind speed. In addition, as Figure 6.14 shows, the change in MSLP patterns results in a changing orientation of lee waves (for example under the High pressure centred over UK regime).

As was the case in TS1, the most lee waves are generated under NAO+ conditions in TS2 and TS3. The mean frequency of occurrence increased over most of Britain and Ireland under NAO+, High pressure and Low close to UK regimes, and decreased under Northwesterly, Southwesterly, Scandinavian high and, most of all, Azores high regimes from the present-day climate. The decrease in the occurrence of wind speeds between 10 and $20\,\mathrm{m\,s^{-1}}$, and the increase in wind speeds $< 10\,\mathrm{m\,s^{-1}}$ during the Azores high weather pattern under the future scenario is likely to contribute to the decrease in occurrence of lee waves under the Azores high. In contrast, the increase in frequency of occurrence of lee waves under the High pressure centred over UK pattern seems to be linked to the changing distribution of 750 hPa horizontal wind speeds, with a decrease in the occurrence of 750 hPa horizontal wind speeds between $5\,\mathrm{m\,s^{-1}}$ and $15\,\mathrm{m\,s^{-1}}$ being compensated for by an increase in the occurrence of 750 hPa horizontal wind speeds faster than $20\,\mathrm{m\,s^{-1}}$. The modelled changes in amplitude, wavelength and orientation of lee waves are small, where they are significant at the 95% confidence level, under these two patterns compared to the changes under other weather patterns. Characteristics generally change less under the Azores high pattern in TS3 than in the other regimes in the future simulations

There were more cases of very high amplitude lee waves in TS2 and TS3 ($> 5\,\mathrm{m\,s^{-1}}$)

than in TS1, but these high amplitude lee waves occur very rarely. The 95th percentile lee wave amplitudes do not change systematically, even within weather patterns (Figure 6.11). Within the ensemble members, there was a great variation of cases of lee wave amplitude stronger than $1\,\mathrm{m\,s^{-1}}$ and $3\,\mathrm{m\,s^{-1}}$, with the TS1 climatology well within the 95% confidence interval of the ensemble in all cases. The ensemble mean change was for a slight but not significant increase in cases of lee wave amplitude stronger than $1\,\mathrm{m\,s^{-1}}$ and $3\,\mathrm{m\,s^{-1}}$ from TS1 to TS2 and TS3. From TS2 to TS3, the change in occurrence of lee waves exceeding amplitudes of $1\,\mathrm{m\,s^{-1}}$, $3\,\mathrm{m\,s^{-1}}$ and $5\,\mathrm{m\,s^{-1}}$ is not significant at the 95% confidence level. The projected ensemble mean change in lee wave wavelength under the Azores high pattern is less than 1 km across almost the entirety of Britain and Ireland, compared with changes in excess of 4 km in parts of the Pennines in the Low close to the UK pattern. While this is a small magnitude change, it is significant at the 95% confidence level in places. The mean orientation of lee waves shifts significantly to be more east/west under RCP 8.5 than in TS1. This change is most notable in the Low close to the UK and Southwesterly patterns. In general however, the mean orientation still follows the MSLP contours for the pattern. The ensemble MSLP patterns change from TS1 to TS2 and TS3, suggesting a shift in the storm track or similar from the present-day to the future climate.

While these results do not show anything particularly drastic occurring under the UKCP18 Local future climate scenario, and indeed disagreement between ensemble members in some cases on the sign of a change in lee wave occurrence from the present-day climate to the future under RCP 8.5, they still represent a worthy study into the occurrence and characteristics of lee waves under a future climate scenario. The increase in occurrence of high amplitude lee waves in RCP 8.5 suggests that there is a higher likelihood of rotors and strong turbulence in the future, however this as at the tail of the distribution, and there is uncertainty between ensemble members. Apart from the Azores high pattern, the occurrence of high amplitude lee waves significantly decreases from TS2 to TS3. However, there are significant increases in occurrence of high amplitude lee waves from TS1 to TS2 and TS3, for example under the NAO+ and Azores high patterns. The changes in the weather patterns, such as the MSLP contours changing in the future climate scenario suggests that the weather patterns are changing themselves (such as the shifting of the mean position of the low in the Low close to UK conditions), as seen by the changing

orientation of lee waves in the future climate scenarios.

There is still much uncertainty over what will change in a future climate, even within the case of models forced to one emissions scenario such as this one.

Chapter 7

Discussion and Conclusions

The work presented in this thesis has analysed whether machine learning models can be trained to detect and characterise trapped lee waves from NWP model data. The trained models were applied to large archives of NWP model output to develop a climatology of lee waves for both the present-day climate and a future climate scenario, to understand what affects the generation of lee waves over Britain and Ireland, and how that may change under a high emissions climate change scenario.

Chapter 4 describes the training and testing of several deep learning models designed to predict the location, wavelength, orientation and amplitude of trapped lee waves over Britain and Ireland, from 2D UKV vertical velocity data on the 700 hPa pressure surface, using the Python library `fastai` (Howard and Gugger, 2020). The model for segmenting (classifying each pixel in the input as either containing a lee wave or not) the input vertical velocity slice performed well against hand-labelled truth data, and the models trained to predict wave characteristics performed favourably against a spectral technique, the S-transform (Hindley et al., 2016; Stockwell et al., 1996), while also delivering a significant speed increase against the S-transform. Chapter 4 demonstrates that these models perform well at recognising and characterising lee waves, and this work has received interest from national meteorological agencies in potentially making these models operational as a post-processing forecasting tool.

Trapped lee waves form over Britain and Ireland on a regular basis, and can lead to the generation of strong turbulence under their crests. Chapter 5 examines the prevalence and characteristics of lee waves in different seasons, regions of Britain and

Ireland, diurnally and through synoptic weather patterns by the creation of a 31-year climatology of lee waves, from the ERA-Interim driven UKCP18 Local data set for 1982–2012. More lee waves are generated under weather patterns with generally faster synoptic wind speeds, such as the positive phase of the NAO, and these patterns have a higher likelihood of higher amplitude waves. The mean orientation of waves is broadly in line with the synoptic wind direction, though with a large spread in some cases. Figure 4.12 shows deviation from the UKV wind direction at 700 hPa and the S-transform and ML orientations in the test data set around 0°. This spread, which is systematically more easterly for northerly wind directions, could be caused by differences between the surface wind speed at the wind speed on the 700 hPa pressure surface, which may explain the spread seen between the geostrophic wind speeds inferred from the MSLP contours and the lee wave orientations in Figure 5.10. There is a correlation between the amplitude of lee waves and the horizontal winds experienced aloft, with higher amplitude waves implying stronger horizontal wind speeds, but not vice-versa. There is little evidence in the data of a diurnal cycle of lee wave prevalence. When other meteorological variables are used to predict the prevalence of lee waves in different regions using random forests and associated SHAP values, the Scorer parameter is the most important for predicting lee waves being generated in 4 out of 6 regions, and the range of the Scorer parameter where waves are present in the UKCP18 data correspond to plausible wavelengths.

Finally, Chapter 6 investigates how the occurrence and characteristics of lee waves may change under a high emissions future climate scenario (RCP 8.5), using the machine learning models applied to the future climate time slices (2021–2041 and 2060–2080) of the UKCP18 Local data. In the future climate projections, there was a slight decrease in the frequency of occurrence of lee waves over the Scottish Highlands from TS1 to TS2 and TS3, but with a large ensemble spread: some ensemble members showed an increase in occurrence of lee waves. The weather patterns from Neal et al. (2016) again proved useful at investigating different synoptic weather patterns on lee waves and their effects – with changes within weather patterns having more of an effect on lee wave occurrence than the changing distribution of weather pattern occurrence (Figure 6.6). There is evidence (e.g. Table 6.1) to suggest a slight but insignificant increase in the occurrence of high amplitude (amplitudes stronger than $3\,\mathrm{m\,s^{-1}}$) lee waves in the future climate scenario, but the occurrence from TS1 is within the 95% confidence interval of the ensemble for TS2 and TS3 for lee waves

of amplitudes exceeding $1\,\mathrm{m\,s^{-1}}$ and $3\,\mathrm{m\,s^{-1}}$. There were no lee waves of amplitude exceeding $5\,\mathrm{m\,s^{-1}}$ in the ERA-Interim driven TS1 simulation, but the overall change from TS2 to TS3 of lee waves of amplitude exceeding $5\,\mathrm{m\,s^{-1}}$ is not significant at the 95% confidence level. The mean orientation of lee waves is significantly more east/west in the future scenario than in TS1 in most weather patterns but orientations are still broadly in line with the synoptic wind directions as seen in the MSLP contours.

7.1 Broader implications

The machine learning models trained and discussed in Chapter 4 provide a computationally easy way to automatically derive information about lee waves from NWP model output. One benefit of using the segmentation and characteristics tools described there is that the perceived severity of lee waves can be understood by the output of amplitude from the neural network. In addition, the generation of lee waves can be understood in relation to the meteorology.

The outputs from the ML models perform well against the S-transform, with a smoother and more physically plausible field produced rather than the more discrete regions produced by the S-transform. One key advantage of the ML models is that, once trained, they can be significantly cheaper and more efficient to run than traditional spectral based methods such as the S-transform. For the 28 high spatial resolution images in the February UKV test set used in Chapter 4, the U-Net models can be up to 13 times faster than the S-transform to produce the wave mask, amplitude, wavelength and orientation measurements (as shown in Table 4.2).

This efficiency combined with their realistic output makes these machine learning models a powerful new tool for post-processing NWP model output. There is no reason why this can be limited to lee waves: any high impact phenomena recognised by eye that is resolved by NWP models can be used to create well-performing post-processing tools. The trained models have a potential benefit for forecasters in producing an automatic detection and characterisation of trapped lee waves directly from operational UKV model output. Indeed, there has been interest from national meteorological agencies, in potentially making these models operational as a post-processing tool to help forecasters diagnose regions of high amplitude lee waves more

easily. This shows the ease at which machine learning tools can be used to make interpreting model output much easier.

Another benefit of machine learning models applied to weather and climate data is the easy creation of climatology data of phenomena resolved within archives of meteorological data. For example, the climatology of lee waves described in Chapter 5, which were produced by applying the models from Chapter 4 to a 31 year archive of the UKCP18 Local data could not have been as easily generated without machine learning models trained to detect and characterise these phenomena.

There is no reason why the UKCP18 data could not be used to produce climatology information of other features (such as convection) identified via machine learning models. The climatology of lee waves created in Chapter 5 shows the ability to link together data from different sources to produce detailed insight into the drivers of lee waves. For example, by combining information on the weather patterns from Neal et al. (2016), the effects of different synoptic conditions on phenomena can be seen.

The use of tools such as feature attribution techniques (in this case SHAP), means that large swathes of data can be combined to probe the importance of different features in the result of a process - in Chapter 5 this was whether or not lee waves were generated. With the SHAP values indicating that the Scorer parameter and horizontal wind speed being the most important for lee wave generation, this is perhaps not surprising from simple linearised models of gravity waves. However, the ML approach in this thesis provides a way to test these assumptions in a more realistic 3-D model with all the real-world complexity of multi-scale orography, non-stationary waves and non idealised profiles.

The availability of the UKCP18 future climate simulations means that conclusions can be drawn about the changing nature of phenomena under a future climate scenario, and how the change in weather patterns has affected this. While there is a lot of uncertainty around future climate change and its effects on the meteorology experienced in Britain and Ireland, the ability to view potential changes to lee waves in this future climate scenario can aid planning. For example, the projected slight increase in high amplitude ($> 3\,\mathrm{m\,s^{-1}}$) lee wave events modelled in the future climate scenario (Table 6.1) means that rotor events may be more likely under the RCP 8.5

scenario, which may affect aviation and road transport.

7.2 Limitations

The hand-labelled training data used to train the lee wave segmentation model in Chapter 4 used a small subset of the available UKV archive of NWP model output, labelled by one human (the author). Since there is some subjectivity in deciding precisely where lee wave boundaries are, there are limitations on the effectiveness of this training. Section 4.3.4 attempted to quantify this subjectivity through the use of crowdsourced labelling using the online platform Zooniverse, albeit with a different labelling practice than the original hand labels, using the in-house Zooniverse labeller, rather than the one created using a Jupyter notebook (see Section 4.2.1). There was a small uptake from volunteers using the Zooniverse labelling tool, meaning that there was insufficient data to have a full idea of subjectivity between labellers, because there were data that had only been labelled twice. Since the edge of wave packets is not always easily defined, this subjectivity would be more reliably quantified with the addition of more crowdsourced labels.

Another approach which could have been fruitful was to train the segmentation model using a signed distance field (for example, values of 0 at pixels on the boundaries of the hand-drawn wave packets, and pixels with values increasing in magnitude the further away from the boundaries they are) – or similar – rather than a binary mask, which would not penalise the model as much for getting the "wrong" classification close to the boundaries of a wave packet, as labelled by the hand-labels. This problem was taken into account, to an extent, with the use of other metrics like the Jaccard Score to assess the trained segmentation model.

The characteristics ML models in Chapter 4 were trained on data containing synthetic "noise": normally distributed randomly generated data, in addition to synthetic wave packets. As discussed in Chapter 4, this was to train models to predict wave characteristics on data with other, non-wave, sources of vertical velocity. While the characteristics models were fine-tuned (so all the previous feature identification learning from the UKV data was retained), the data the characteristics models were trained on should reflect the "real-life" data as much as possible, which is not really the case for the synthetic data used – either in its regularly shaped oval wave pack-

ets, or the "noise". That said, on the UKV data, the ML characteristics models performed reasonably in line with the S-transform at predicting lee wave characteristics, suggesting that the synthetic data were sufficient enough for fine-tuning the segmentation model to produce wave characteristics.

The probing of the climatology of lee waves in Chapter 5 to understand which other variables were important for lee wave generation used just one feature attribution technique: SHAP. While SHAP attempts to account for cross correlation between variables, a more thorough investigation using different techniques (such as permutation feature importance) could compare these techniques and their metrics of "importance". In addition, this analysis could have been repeated, only focusing on cases of high amplitude lee waves to be able to distinguish between cases of more impactful lee waves, where there is a higher likelihood of rotor events.

Repeating the analysis from Section 5.3.4 for the future climate would have been a useful study, to see how the other meteorological variables affecting lee wave generation changed under the future climate scenarios. However, this would require a large volume of data for the 42 years to be pulled from MASS, a task that would take months, and require disk space on the order of terabytes - which was not feasible.

The ensemble within the future climate UKCP18 Local data covers a range of possible outcomes within the RCP 8.5 scenario. Probing why certain ensemble members show different changes in lee wave occurrence and characteristics would have been worthwhile. As Section 6.3.2 shows, changes in lee wave occurrence (and variation between ensemble members) in the future climate ensemble cannot be explained by the distribution of weather patterns alone. This suggests that there are differences in the meteorology and how conducive the meteorology is to lee wave activity, between individual ensemble members in the PPE. However, the precise distribution of changed parameters is not available, and the ensemble should be treated as covering the range of possible outcomes to the climate under RCP 8.5. Each ensemble member has a set of parameters perturbed, so being able to say precisely which parameters within the combination are responsible for the change in lee wave occurrence and characteristics would not be straightforward. Instead, the ensemble is to be treated as representing a range of outcomes expected under RCP 8.5.

7.3 Recommendations for future work

The models described in Chapter 4 have been trained on and applied to 2D vertical velocity NWP data over Britain and Ireland. Applying the same technique to other regions of the world would likely require retraining the models on the new data, as although the wave patterns would look similar, the other sources of vertical velocity in the data may not. Adding synthetic noise to the synthetic data to train characteristics models did produce better performing models than training on data without noise, and the trained models performed favourably compared to the S-transform. In addition, this approach used 2D slices of vertical velocities, at 700 hPa since lee waves are coherent in UKV model output over Britain and Ireland at 700 hPa. Using data on other pressure surfaces, or training a model on 3D data, or with other variables such as vertical profiles of the Scorer parameter in addition to the vertical velocities, may produce models that are more skilful than the ones presented in Chapter 4.

This thesis restricted itself to lee waves that form over Britain and Ireland, mainly because the majority of the data immediately available was from the MetUM, over Britain and Ireland. This work could be broadened beyond Britain and Ireland, but for a close as possible recreation of the work here, would require high resolution model simulations where lee waves have been resolved. In addition, the different meteorology and other non-wave sources of vertical velocities may produce different features in other parts of the world, so the fine-tuning of the ML models on data from the other parts of the world may result in better results (as opposed to only training on labelled data of lee waves from Britain and Ireland), and a more generalisable set of models. The use of satellite imagery was originally considered, but is not a like-for-like replacement: for example the visible satellite imagery requires presence of cloud cover and daylight for lee waves to be visible, restricting the cases that can be observed through visible satellite imagery.

There are high resolution MetUM simulations available from the Rockies and South Georgia where waves have been resolved. It would be well worth fully testing the performance of the ML models trained on data over Britain and Ireland on data where other sources of vertical velocity may look different to those the models were trained on. Figure 7.1 shows one such segmentation on an example from the Rockies,

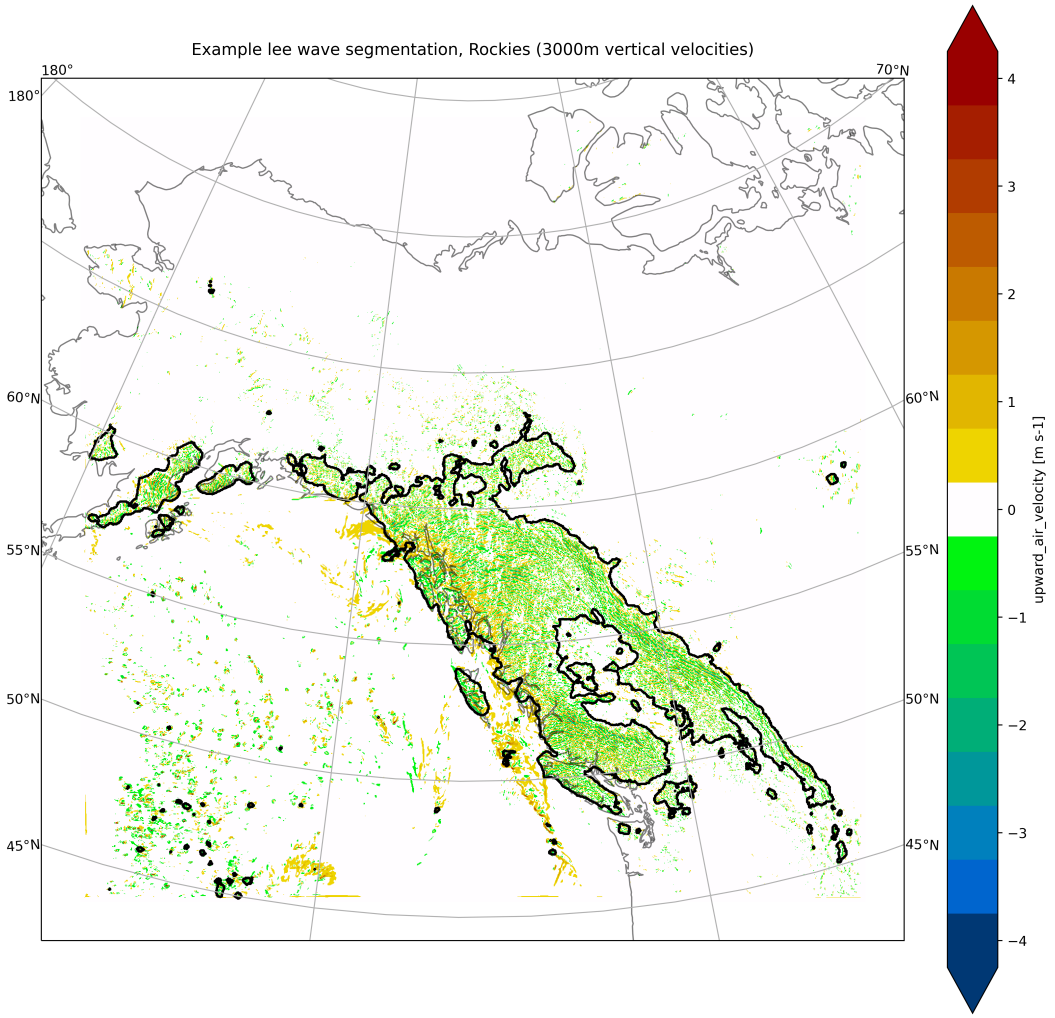


Figure 7.1: Example lee wave segmentation on data available from the Rockies (vertical velocities at 3000 m). Some regions of convection have been incorrectly recognised as lee waves by the model, suggesting some fine tuning to the region may be needed.

using the trained segmentation model from Chapter 4. Some regions of convection to the west have been incorrectly recognised as waves, suggesting that some fine-tuning to the model output over the Rockies may be needed to produce the best-performing model on data from the Rockies. However, qualitatively, the segmentation model performs well given it has not been exposed to data from the Rockies during training.

There is also the option to investigate whether models can be trained from satellite data to detect and characterise gravity waves. With lee waves, there may be more of a binary difference between peaks and troughs of the waves rather than a sinusoidal waveform as seen in the vertical velocity data. Lee waves do not tend to show up in satellite imagery if there is insufficient moisture, or if there is high level cloud cover obscuring the view of the waves.

There are also avenues worth exploring within ML model development, such as investigating the effects of feeding additional data to the ML models. These could include vertical velocities at different heights, surface winds, orography or the Scorer parameter. Incorporating known physical relationships between different variables for lee waves into a deep learning model may help a tool such as those described in Chapter 4 better diagnose waves, and address some nervousness about using "black box" machine learned models such as these operationally. The use of a U-Net architecture was chosen due to its performance at segmentation tasks in previous literature (both inside and outside atmospheric science: Dev et al., 2019; Ronneberger et al., 2015), but there could well be more effective model architectures for this task: for example, transformer based models have been shown to perform well at segmentation (Li et al., 2023).

The machine learning-derived climatology described in Chapter 5 presents an opportunity to investigate the correlation between the physics of lee waves and other meteorological phenomena, for example, for parametrisation schemes: lee waves are important for momentum transport, as well as surface impacts such as rotors (Bretherton, 1969). This could lead to the creation of a parametrisation scheme for global models where lee waves and their effects are currently not represented, if there is sufficient correlation between other (well resolved) variables in global models.

The 8 more broad weather patterns from Neal et al. (2016) were used here, but there is an option to use the more specific 30 weather patterns to further hone in on which specific regimes within the 8 patterns are more conducive to lee waves over Britain and Ireland. However, comparing PDFs or maps is harder for more patterns. A version of Figure 5.6 with 30 subplots instead of 8 would be much harder to interpret, especially when some patterns are rather similar and differences between the regimes could be more subtle than for the 8 broad patterns. As shown in Chapter 6, the MSLP patterns within the 8 weather patterns are modelled as changing from the present-day to the future. Work could be done to investigate which of the 30 regimes change in dominance from the present-day to the future, and how that affects meteorology over Britain and Ireland, not just for lee waves as investigated here.

The climatology of lee waves produced in Chapter 5 only explored the influence of horizontal wind speed on lee wave amplitude: further work investigating the

distribution of lee wave characteristics and different meteorological variables may help clarify the findings from using the SHAP importance to find which variables had the most influence on the generation of lee waves, especially the highest impact events with high amplitude waves.

In addition, a more thorough investigation on the difference between the ERA-Interim driven climatology and the ensemble for the present-day climate in TS1 may be useful to properly quantify the differences between the future climate projections and the present-day. The climatology data in Chapter 6 used the ensemble data from the UKCP18 Local data set to probe how lee waves may change in the high emissions RCP 8.5 scenario, while the present-day climate data were from the UKCP18 Local data driven by ERA-Interim reanalysis. It may have been more appropriate to use the same ensemble data to better track the changes from the present-day to future climate. It is highly possible that the differences between the ERA-Interim model data and the ensemble data for TS1 are small, as the analysis in Section 6.2.1 suggests. However, the reanalysis-driven model data is a better indication of the present-day climate than the present-day data from the PPE. Despite this, the reanalysis-driven model data used for the analysis of lee waves in TS1 was produced using the same configuration of the CPM (MetUM version 10.6) as for the future climate simulations. This configuration is similar to the UKV configuration of the MetUM that the machine learning models were originally trained on. Hence, resolved lee waves look similar in the CPM output regardless of whether it was driven by ERA-Interim reanalysis or the PPE.

7.4 Concluding remarks

Machine learning tools continue to complete tasks with skill previously thought to be near on impossible for computers to complete, and at a fast rate of progress. This thesis has shown the ability for machine learning to potentially aid forecasters by the production of accurate machine learning models which can detect and characterise trapped lee waves, and also inform meteorology and climate research through an investigation of the drivers behind lee waves over Britain and Ireland, as well as a study of the potential for lee waves to change under a high emissions future climate scenario.

Bibliography

Ágústsson, H. and Ólafsson, H. 2014. Simulations of Observed Lee Waves and Rotor Turbulence. *Monthly Weather Review*. **142**(2), pp. 832–849. DOI: 10.1175/MWR-D-13-00212.1.

Alexander, M. J., Gille, J., Cavanaugh, C., Coffey, M., Craig, C., Eden, T., Francis, G., Halvorson, C., Hannigan, J., Khosravi, R., Kinnison, D., Lee, H., Massie, S., Nardi, B., Barnett, J., Hepplewhite, C., Lambert, A. and Dean, V. 2008. Global estimates of gravity wave momentum flux from High Resolution Dynamics Limb Sounder observations. *Journal of Geophysical Research: Atmospheres*. **113**(D15), pp. 1–11. DOI: 10.1029/2007JD008807.

American Meteorological Society. 2012. *Mountain Wave*. [Online]. Glossary of Meteorology. URL: https://glossary.ametsoc.org/wiki/Mountain_wave.

Andrade, C. W. L., Montenegro, S. M. G. L., Montenegro, A. A. A., Lima, J. R. d. S., Srinivasan, R. and Jones, C. A. 2021. Climate change impact assessment on water resources under RCP scenarios: A case study in Mundaú River Basin, Northeastern Brazil. *International Journal of Climatology*. **41**(S1), E1045–E1061. DOI: 10.1002/joc.6751.

Bell, T. H. 1975. Lee waves in stratified flows with simple harmonic time dependence. *Journal of Fluid Mechanics*. **67**(4), pp. 705–722. DOI: 10.1017/S0022112075000560.

Ben-Bouallegue, Z., Clare, M. C. A., Magnusson, L., Gascon, E., Maier-Gerber, M., Janousek, M., Rodwell, M., Pinault, F., Dramsch, J. S., Lang, S. T. K., Raoult, B., Rabier, F., Chevallier, M., Sandu, I., Dueben, P., Chantry, M. and Pappenberger, F. 2023. The rise of data-driven weather forecasting [Preprint]. URL: http://arxiv.org/abs/2307.10128.

Benedict, J. J., Lee, S. and Feldstein, S. B. 2004. Synoptic View of the North Atlantic Oscillation. *Journal of the Atmospheric Sciences*. **61**(2), pp. 121–144. DOI: 10.1175/1520-0469(2004)061<0121:SVOTNA>2.0.CO;2.

Bi, K., Xie, L., Zhang, H., Chen, X., Gu, X. and Tian, Q. 2023. Accurate medium-range global weather forecasting with 3D neural networks. *Nature*. **619**(7970), pp. 533–538. DOI: 10.1038/s41586-023-06185-3.

Biau, G. and Scornet, E. 2016. A random forest guided tour. *TEST*. **25**(2), pp. 197–227. DOI: 10.1007/s11749-016-0481-7.

Bishop, C. M. 2006. *Pattern Recognition and Machine Learning*. New York: Springer New York.

Blockley, J. A. and Lyons, T. J. 1994. Airflow over a two-dimensional escarpment. III: Nonhydrostatic flow. *Quarterly Journal of the Royal Meteorological Society*. **120**(515), pp. 79–109. DOI: 10.1002/qj.49712051507.

Bloomfield, H. C., Bates, P., Shaffrey, L. C., Hillier, J., Champion, A., Cotterill, D., Pope, J. O. and Kumar, D. 2024. Synoptic conditions conducive for compound wind-flood events in Great Britain in present and future climates. *Environmental Research Letters*. **19**(2). DOI: 10.1088/1748-9326/ad1cb7.

Böhm, C., Schween, J. H., Reyers, M., Maier, B., Löhnert, U. and Crewell, S. 2021. Towards a climatology of fog frequency in the Atacama Desert via multi-spectral satellite data and machine learning techniques. *Journal of Applied Meteorology and Climatology*. **60**(8), pp. 1149–1169. DOI: 10.1175/JAMC-D-20-0208.1.

Bowler, N. E., Pierce, C. E. and Seed, A. 2004. Development of a precipitation nowcasting algorithm based upon optical flow techniques. *Journal of Hydrology*. **288**(1-2), pp. 74–91. DOI: 10.1016/j.jhydrol.2003.11.011.

Bramberger, M., Dörnbrack, A., Wilms, H., Gemsa, S., Raynor, K. and Sharman, R. 2018. Vertically propagating Mountain Waves-A hazard for high-flying aircraft? *Journal of Applied Meteorology and Climatology*. **57**(9), pp. 1957–1975. DOI: 10.1175/JAMC-D-17-0340.1.

Breiman, L. 2001. Random Forests. *Machine Learning*. **45**, pp. 5–32. DOI: 10.1023/A:1010933404324.

Bretherton, F. P. 1969. Momentum transport by gravity waves. *Quarterly Journal of the Royal Meteorological Society*. **95**(404), pp. 213–243. DOI: 10.1002/qj.49709540402.

Brown, A., Milton, S., Cullen, M., Golding, B., Mitchell, J. and Shelly, A. 2012. Unified Modeling and Prediction of Weather and Climate: A 25-Year Journey. *Bulletin of the American Meteorological Society*. **93**(12), pp. 1865–1877. DOI: 10.1175/BAMS-D-12-00018.1.

Bush, M., Boutle, I., Edwards, J., Finnenkoetter, A., Franklin, C., Hanley, K., Jayakumar, A., Lewis, H., Lock, A., Mittermaier, M., Mohandas, S., North, R., Porson, A., Roux, B., Webster, S. and Weeks, M. 2023. The second Met Office Unified Model–JULES Regional Atmosphere and Land configuration, RAL2. *Geoscientific Model Development*. **16**(6), pp. 1713–1734. DOI: 10.5194/gmd-16-1713-2023.

Chantry, M., Alexe, M., Lang, S., Raoult, B., Dramsch, J., Pinault, F., Bouallegue, Z. B., Clare, M., Lessig, C., Magnusson, L., Dueben, P., Brown, A., Pappenberger,

F. and Rabier, F. 2024. 'AIFS–ECMWF's Data-Driven Probabilistic Forecasting System'. *104th AMS Annual Meeting*. AMS.

Charlton-Perez, A. J., Dacre, H. F., Driscoll, S., Gray, S. L., Harvey, B., Harvey, N. J., Hunt, K. M. R., Lee, R. W., Swaminathan, R., Vandaele, R. and Volonté, A. 2024. Do AI models produce better weather forecasts than physics-based models? A quantitative evaluation case study of Storm Ciarán. *npj Climate and Atmospheric Science*. **7**(1), p. 93. DOI: 10.1038/s41612-024-00638-w.

Chen, H., Covert, I. C., Lundberg, S. M. and Lee, S.-I. 2023. Algorithms to estimate Shapley value feature attributions. *Nature Machine Intelligence*. **5**(6), pp. 590–601. DOI: 10.1038/s42256-023-00657-x.

Cohn, S. A., Grubišić, V. and Brown, W. O. J. 2011. Wind Profiler Observations of Mountain Waves and Rotors during T-REX. *Journal of Applied Meteorology and Climatology*. **50**(4), pp. 826–843. DOI: 10.1175/2010JAMC2611.1.

Colfescu, I., Klemp, J. B., Bollasina, M. A., Mobbs, S. D. and Burton, R. R. 2021. The Dynamics of Observed Lee Waves over the Snæfellsnes Peninsula in Iceland. *Monthly Weather Review*. **149**(5), pp. 1559–1575. DOI: 10.1175/MWR-D-20-0288.1.

Coney, J., Denby, L., Ross, A. N., Wang, H., Vosper, S., van Niekerk, A., Dunstan, T. and Hindley, N. 2024. Identifying and characterising trapped lee waves using deep learning techniques. *Quarterly Journal of the Royal Meteorological Society*. **150**(758), pp. 213–231. DOI: 10.1002/qj.4592.

Copeland, B. J., ed. 2004. *The Essential Turing*. Oxford University Press. DOI: 10.1093/oso/9780198250791.001.0001. URL: https://academic.oup.com/book/42030.

Cotterill, D. F., Pope, J. O. and Stott, P. A. 2023. Future extension of the UK summer and its impact on autumn precipitation. *Climate Dynamics*. **60**(5-6), pp. 1801–1814. DOI: 10.1007/s00382-022-06403-0.

Darby, L. S. and Poulos, G. S. 2006. The evolution of lee-wave-rotor activity in the lee of Pike's Peak under the influence of a cold frontal passage: Implications for aircraft safety. *Monthly Weather Review*. **134**(10), pp. 2857–2876. DOI: 10.1175/MWR3208.1.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., Mcnally, A. P., Monge-Sanz, B. M., Morcrette, J. J., Park, B. K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J. N. and Vitart, F. 2011. The ERA-Interim reanalysis: Configuration

and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*. **137**(656), pp. 553–597. DOI: 10.1002/qj.828.

Denby, L. 2020. Discovering the Importance of Mesoscale Cloud Organization Through Unsupervised Classification. *Geophysical Research Letters*. **47**(1), pp. 1–10. DOI: 10.1029/2019GL085190.

Denby, L. 2023. *Synthetic gravity waves*. [Online]. Last accessed: 2023-04-03. URL: https://doi.org/10.5281/zenodo.7576811. DOI: 10.5281/zenodo.7576811.

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K. and Fei-Fei, L. 2009. 'ImageNet: A large-scale hierarchical image database'. *2009 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, pp. 248–255. DOI: 10.1109/CVPRW.2009.5206848.

Dev, S., Manandhar, S., Lee, Y. H. and Winkler, S. 2019. Multi-label Cloud Segmentation Using a Deep Network. *2019 USNC-URSI Radio Science Meeting (Joint with AP-S Symposium), USNC-URSI 2019 - Proceedings*, pp. 113–114. DOI: 10.1109/USNC-URSI.2019.8861850.

Diamantakis, M. 2013. *The semi-Lagrangian technique in atmospheric modelling : current status and future challenges*. Tech. rep. September. ECMWF, pp. 183–200. URL: http://www.ecmwf.int/sites/default/files/elibrary/2014/9054-semi-lagrangian-technique-atmospheric-modelling-current-status-and-future-challenges.pdf.

Doger de Speville, C., Seviour, W. J. M. and Lo, Y. T. E. 2023. Predicting future UK nighttime urban heat islands using observed short-term variability and regional climate projections. *Environmental Research Letters*. **18**(10), p. 104044. DOI: 10.1088/1748-9326/acf94c.

Doyle, J. D. and Durran, D. R. 2002. The dynamics of mountain-wave-induced rotors. *Journal of the Atmospheric Sciences*. **59**(2), pp. 186–201. DOI: 10.1175/1520-0469(2002)059<0186:TDOMWI>2.0.CO;2.

Durran, D. 2003. 'Lee Waves and Mountain Waves'. *Encyclopedia of Atmospheric Sciences*. Ed. by J. R. Holton, J. A. Curry and J. A. Pyle. Burlington: Elsevier, pp. 1161–1169. DOI: 10.1016/B0-12-227090-8/00202-5.

ECMWF. 2024. *ECMWF Charts*. [Online]. Last accessed: 2024-03-18. URL: https://charts.ecmwf.int/?facets=%7B%22Producttype%22%3A%5B%22Experimental%3AMachinelearningmodels%22%5D%7D.

Elliott, S. 2001. 'Active Control of Nonlinear Systems'. *Signal Processing for Active Control*. Vol. 49. 1. Elsevier, pp. 367–409. DOI: 10.1016/B978-012237085-4/50010-7.

Elvidge, A. D., Vosper, S. B., Wells, H., Cheung, J. C. H., Derbyshire, S. H. and Turp, D. 2017. Moving towards a wave-resolved approach to forecasting mountain

wave induced clear air turbulence. *Meteorological Applications*. **24**(3), pp. 540–550. DOI: 10.1002/met.1656.

Fortun, D., Bouthemy, P. and Kervrann, C. 2015. Optical flow modeling and computation: A survey. *Computer Vision and Image Understanding*. **134**, pp. 1–21. DOI: 10.1016/j.cviu.2015.02.008.

Gagne, D. J., Haupt, S. E., Nychka, D. W. and Thompson, G. 2019. Interpretable deep learning for spatial analysis of severe hailstorms. *Monthly Weather Review*. **147**(8), pp. 2827–2845. DOI: 10.1175/MWR-D-18-0316.1.

Goh, A. 1995. Back-propagation neural networks for modeling complex systems. *Artificial Intelligence in Engineering*. **9**(3), pp. 143–151. DOI: 10.1016/0954-1810(94)00011-S.

Govender, P. and Sivakumar, V. 2020. Application of k-means and hierarchical clustering techniques for analysis of air pollution: A review (1980–2019). *Atmospheric Pollution Research*. **11**(1), pp. 40–56. DOI: 10.1016/j.apr.2019.09.009.

Grange, S. K., Carslaw, D. C., Lewis, A. C., Boleti, E. and Hueglin, C. 2018. Random forest meteorological normalisation models for Swiss PM10 trend analysis. *Atmospheric Chemistry and Physics*. **18**(9), pp. 6223–6239. DOI: 10.5194/acp-18-6223-2018.

Grazzini, F., Craig, G. C., Keil, C., Antolini, G. and Pavan, V. 2020. Extreme precipitation events over northern Italy. Part I: A systematic classification with machine-learning techniques. *Quarterly Journal of the Royal Meteorological Society*. **146**(726), pp. 69–85. DOI: 10.1002/qj.3635.

Grubišić, V. 1997. 'Critical levels and wavebreaking in 3D orographic flows'. *Workshop on Orography*. Reading: ECMWF, pp. 67–97.

Grubišić, V. and Billings, B. J. 2008. Climatology of the Sierra Nevada Mountain-Wave Events. *Monthly Weather Review*. **136**(2), pp. 757–768. DOI: 10.1175/2007MWR1902.1.

Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, G., Cai, J. and Chen, T. 2018. Recent advances in convolutional neural networks. *Pattern Recognition*. **77**, pp. 354–377. DOI: 10.1016/j.patcog.2017.10.013.

Guastavino, S., Piana, M., Tizzi, M., Cassola, F., Iengo, A., Sacchetti, D., Solazzo, E. and Benvenuto, F. 2022. Prediction of severe thunderstorm events with ensemble deep learning and radar data. *Scientific Reports*. **12**(1), p. 20049. DOI: 10.1038/s41598-022-23306-6.

Haenlein, M. and Kaplan, A. 2019. A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence. *California Management Review*. **61**(4), pp. 5–14. DOI: 10.1177/0008125619864925.

Hawkins, E., Burt, S., Brohan, P., Lockwood, M., Richardson, H., Roy, M. and Thomas, S. 2019. Hourly weather observations from the Scottish Highlands (1883–1904) rescued by volunteer citizen scientists. *Geoscience Data Journal*. **6**(2), pp. 160–173. DOI: 10.1002/gdj3.79.

He, K., Zhang, X., Ren, S. and Sun, J. 2016. 'Deep Residual Learning for Image Recognition'. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Vol. 2016-Decem. IEEE, pp. 770–778. DOI: 10.1109/CVPR.2016.90.

Hendry, A., Haigh, I. D., Nicholls, R. J., Winter, H., Neal, R., Wahl, T., Joly-Laugel, A. and Darby, S. E. 2019. Assessing the characteristics and drivers of compound flooding events around the UK coast. *Hydrology and Earth System Sciences*. **23**(7), pp. 3117–3139. DOI: 10.5194/hess-23-3117-2019.

Hertenstein, R. F. and Kuettner, J. P. 2005. Rotor types associated with steep lee topography: influence of the wind profile. *Tellus A: Dynamic Meteorology and Oceanography*. **57**(2), p. 117. DOI: 10.3402/tellusa.v57i2.14625.

Hill, A. J., Herman, G. R. and Schumacher, R. S. 2020. Forecasting Severe Weather with Random Forests. *Monthly Weather Review*. **148**(5), pp. 2135–2161. DOI: 10.1175/MWR-D-19-0344.1.

Hindley, N. P., Wright, C. J., Smith, N. D. and Mitchell, N. J. 2015. The southern stratospheric gravity wave hot spot: individual waves and their momentum fluxes measured by COSMIC GPS-RO. *Atmospheric Chemistry and Physics*. **15**(14), pp. 7797–7818. DOI: 10.5194/acp-15-7797-2015.

Hindley, N. P., Smith, N. D., Wright, C. J., Rees, D. A. S. and Mitchell, N. J. 2016. A two-dimensional Stockwell transform for gravity wave analysis of AIRS measurements. *Atmospheric Measurement Techniques*. **9**(6), pp. 2545–2565. DOI: 10.5194/amt-9-2545-2016.

Hindley, N. P., Wright, C. J., Gadian, A. M., Hoffmann, L., Hughes, J. K., Jackson, D. R., King, J. C., Mitchell, N. J., Moffat-Griffin, T., Moss, A. C., Vosper, S. B. and Ross, A. N. 2021. Stratospheric gravity waves over the mountainous island of South Georgia: testing a high-resolution dynamical model with 3-D satellite observations and radiosondes. *Atmospheric Chemistry and Physics*. **21**(10), pp. 7695–7722. DOI: 10.5194/acp-21-7695-2021.

Hindley, N. P., Wright, C. J., Smith, N. D., Hoffmann, L., Holt, L. A., Alexander, M. J., Moffat-Griffin, T. and Mitchell, N. J. 2019. Gravity waves in the winter stratosphere over the Southern Ocean: high-resolution satellite observations and 3-D spectral analysis. *Atmospheric Chemistry and Physics*. **19**(24), pp. 15377–15414. DOI: 10.5194/acp-19-15377-2019.

Holmboe, J. and Klieforth, H. 1957. *Investigation of mountain lee waves and the air flow over the Sierra Nevada*. Tech. rep. Los Angeles: University of California. URL: https://apps.dtic.mil/sti/tr/pdf/AD0133606.pdf.

Holton, J. R. and Hakim, G. J. 2013. *An Introduction to Dynamic Meteorology*. 5th. Vol. 9780123848. Oxford: Elsevier, pp. 1–532. DOI: 10.1016/C2009-0-63394-8.

Howard, J. and Gugger, S. 2020. Fastai: A Layered API for Deep Learning. *Information*. **11**(2), p. 108. DOI: 10.3390/info11020108.

Ideami, J. 2024. *Loss Landscape*. [Online]. Last accessed: 2024-08-10. URL: https://losslandscape.com/.

International Civil Aviation Organization. 1968. *BOAC, Boeing 707, G-APFE, accident at the foot of Mount Fuji, Japan, on 5 March 1966*. Tech. rep. 7. United Kingdom: Board of Trade, pp. 35–55.

IPCC. 2023. *Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)]. IPCC, Geneva, Switzerland.* Tech. rep. Intergovernmental Panel on Climate Change. DOI: 10.59327/IPCC/AR6-9789291691647.

Jadon, S. 2020. 'A survey of loss functions for semantic segmentation'. *2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*. IEEE, pp. 1–7. DOI: 10.1109/CIBCB48159.2020.9277638.

Jiang, Q. and Doyle, J. D. 2008. Diurnal Variation of Downslope Winds in Owens Valley during the Sierra Rotor Experiment. *Monthly Weather Review*. **136**(10), pp. 3760–3780. DOI: 10.1175/2008MWR2469.1.

Johns, R. H. and Doswell, C. A. 1992. Severe Local Storms Forecasting. *Weather and Forecasting*. **7**(4), pp. 588–612.

Justin, A. D., Willingham, C., McGovern, A. and Allen, J. T. 2023. Toward Operational Real-time Identification of Frontal Boundaries Using Machine Learning. *Artificial Intelligence for the Earth Systems*, pp. 1–56. DOI: 10.1175/aies-d-22-0052.1.

Keisler, R. 2022. Forecasting Global Weather with Graph Neural Networks [Preprint]. URL: http://arxiv.org/abs/2202.07575.

Kendon, E., Short, C., Pope, J., Chan, S., Wilkinson, J., Tucker, S., Bett, P. and Harris, G. 2021. *Update to UKCP Local (2.2km) projections, Science Report, Met Office Hadley Centre*. Tech. rep. July. Exeter, Devon: Met Office, pp. 1–114. URL: https://www.metoffice.gov.uk/pub/data/weather/uk/ukcp18/science-reports/ukcp18_local_update_report_2021.pdf.

Kendon, E. J., Roberts, N. M., Fosser, G., Martin, G. M., Lock, A. P., Murphy, J. M., Senior, C. A. and Tucker, S. O. 2020. Greater future U.K. winter precipitation increase in new convection-permitting scenarios. *Journal of Climate*. **33**(17), pp. 7303–7318. DOI: 10.1175/JCLI-D-20-0089.1.

Kennedy-Asser, A. T., Andrews, O., Mitchell, D. M. and Warren, R. F. 2021. Evaluating heat extremes in the UK Climate Projections (UKCP18). *Environmental Research Letters*. **16**(1), p. 014039. DOI: 10.1088/1748-9326/abc4ad.

Kotsiantis, S. B. 2013. Decision trees: a recent overview. *Artificial Intelligence Review*. **39**(4), pp. 261–283. DOI: 10.1007/s10462-011-9272-4.

Krizhevsky, A., Sutskever, I. and Hinton, G. E. 2012. 'ImageNet Classification with Deep Convolutional Neural Networks'. *Advances in Neural Information Processing Systems*. Ed. by F Pereira, C. J. Burges, L Bottou and K. Q. Weinberger. Vol. 25. Curran Associates, Inc. URL: https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf.

Kühnlein, C., Dörnbrack, A. and Weissmann, M. 2013. High-Resolution Doppler Lidar Observations of Transient Downslope Flows and Rotors. *Monthly Weather Review*. **141**(10), pp. 3257–3272. DOI: 10.1175/MWR-D-12-00260.1.

Lam, R., Sanchez-Gonzalez, A., Willson, M., Wirnsberger, P., Fortunato, M., Alet, F., Ravuri, S., Ewalds, T., Eaton-Rosen, Z., Hu, W., Merose, A., Hoyer, S., Holland, G., Vinyals, O., Stott, J., Pritzel, A., Mohamed, S. and Battaglia, P. 2023. Learning skillful medium-range global weather forecasting. *Science*. **382**(6677), pp. 1416–1421. DOI: 10.1126/science.adi2336.

Laprise, R. 2008. Regional climate modelling. *Journal of Computational Physics*. **227**(7), pp. 3641–3666. DOI: 10.1016/j.jcp.2006.10.024.

LeCun, Y., Bengio, Y. and Hinton, G. 2015. Deep learning. *Nature*. **521**(7553), pp. 436–444. DOI: 10.1038/nature14539.

LeCun, Y., Boser, B., Denker, J., Henderson, D., Howard, R, Hubbard, W. and Jackel, L. 1989. 'Handwritten Digit Recognition with a Back-Propagation Network'. *Advances in Neural Information Processing Systems*. Ed. by D Touretzky. Vol. 2. Morgan-Kaufmann.

LeCun, Y., Haffner, P., Bottou, L. and Bengio, Y. 1999. 'Object Recognition with Gradient-Based Learning'. *Shape, Contour and Grouping in Computer Vision*. Ed. by D. A. Forsyth, J. L. Mundy, V. di Gesù and R. Cipolla. Springer. Chap. 6, pp. 319–345.

Lemos, M. C. and Rood, R. B. 2010. Climate projections and their impact on policy and practice. *WIREs Climate Change*. **1**(5), pp. 670–682. DOI: 10.1002/wcc.71.

LeNail, A. 2019. NN-SVG: Publication-Ready Neural Network Architecture Schematics. *Journal of Open Source Software*. **4**(33), p. 747. DOI: 10.21105/joss.00747.

Lester, P. F. 1978. A Lee wave cloud climatology for pincher creek, Alberta. *Atmosphere-Ocean*. **16**(2), pp. 157–168. DOI: 10.1080/07055900.1978.9649023.

Li, H., Xu, Z., Taylor, G., Studer, C. and Goldstein, T. 2017. Visualizing the Loss Landscape of Neural Nets [Preprint]. URL: http://arxiv.org/abs/1712.09913.

Li, X., Ding, H., Yuan, H., Zhang, W., Pang, J., Cheng, G., Chen, K., Liu, Z. and Loy, C. C. 2023. Transformer-Based Visual Segmentation: A Survey [Preprint]. URL: http://arxiv.org/abs/2304.09854.

Lorenz, T. and Barstad, I. 2016. A dynamical downscaling of ERA-Interim in the North Sea using WRF with a 3 km grid-for wind resource applications. *Wind Energy*. **19**(10), pp. 1945–1959. DOI: 10.1002/we.1961.

Lowe, J. A., Bernie, D., Bett, P. E., Bricheno, L., Brown, S. J., Calvert, D., Clark, R. T., Eagle, K. E., Edwards, T. and Fosser, G. 2018. UKCP18 science overview report. *Met Office Hadley Centre: Exeter, UK*, pp. 1–73. URL: https://www.metoffice.gov.uk/pub/data/weather/uk/ukcp18/science-reports/UKCP18-Overview-report.pdf.

Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N. and Lee, S. I. 2020. From local explanations to global understanding with explainable AI for trees. *Nature Machine Intelligence*. **2**(1), pp. 56–67. DOI: 10.1038/s42256-019-0138-9.

Lundberg, S. M. and Lee, S.-I. 2017. 'A Unified Approach to Interpreting Model Predictions'. *Advances in Neural Information Processing Systems 30*. Ed. by I Guyon, U. V. Luxburg, S Bengio, H Wallach, R Fergus, S Vishwanathan and R Garnett. Curran Associates, Inc., pp. 4765–4774.

Manley, G. 1945. The Helm Wind of Crossfell, 1937-1939. *Quarterly Journal of the Royal Meteorological Society*. **71**(309-310), pp. 197–219. DOI: 10.1002/qj.49707130901.

Manning, C., Kendon, E. J., Fowler, H. J. and Roberts, N. M. 2023. Projected increase in windstorm severity and contribution from sting jets over the UK and Ireland. *Weather and Climate Extremes*. **40**, p. 100562. DOI: 10.1016/j.wace.2023.100562.

Markowski, P. and Richardson, Y. 2010. *Mesoscale Meteorology in Midlatitudes*. Chichester, UK: John Wiley & Sons, Ltd. DOI: 10.1002/9780470682104.

Marriott, W. 1886. The helm wind of August 19th, 1885. *Quarterly Journal of the Royal Meteorological Society*. **12**(57), pp. 1–10. DOI: 10.1002/qj.4970125701.

Marriott, W. 1889. Report on the Helm Wind inquiry. *Quarterly Journal of the Royal Meteorological Society*. **15**(70), pp. 103–118. DOI: 10.1002/qj.4970157010.

Matsuoka, D., Watanabe, S., Sato, K., Kawazoe, S., Yu, W. and Easterbrook, S. 2020. Application of Deep Learning to Estimate Atmospheric Gravity Wave Parameters in Reanalysis Data Sets. *Geophysical Research Letters*. **47**(19). DOI: 10.1029/2020GL089436.

Met Office. 1997. *Forecasters' Reference Book*. 3rd. Bracknell: Meteorological Office.

Met Office. 2018. *UKCP18 guidance: data availability, access and formats*. Tech. rep. Exeter: Met Office, pp. 1–15. URL: https://www.metoffice.gov.uk/binaries/content/assets/metofficegovuk/pdf/research/ukcp/ukcp18-guidance-data-availability-access-and-formats.pdf.

Met Office. 2019a. *Numerical weather prediction models*. [Online]. Last accessed: 2024-03-12. URL: https://www.metoffice.gov.uk/research/approach/modelling-systems/unified-model/weather-forecasting.

Met Office. 2019b. *UKCP18 Factsheet: UKCP Local (2.2km) Projections*. Tech. rep. Exeter: Met Office, pp. 1–21. URL: https://www.metoffice.gov.uk/binaries/content/assets/metofficegovuk/pdf/research/ukcp/ukcp18-factsheet-local-2.2km.pdf.

Met Office. 2019c. *Unified Model*. [Online]. Last accessed: 2024-04-22. URL: https://www.metoffice.gov.uk/research/approach/modelling-systems/unified-model.

Met Office Hadley Centre. 2020. *UKCP Global (60km) - European Circulation Indices*. [Online]. Centre for Environmental Data Analysis. Last accessed: 2024-03-19. URL: https://catalogue.ceda.ac.uk/uuid/6e61f79cb6b0457eb84edaffcf0aab3a.

Mobbs, S., Vosper, S., Sheridan, P., Cardoso, R., Burton, R., Arnold, S., Hill, M., Horlacher, V. and Gadian, A. 2005. Observations of downslope winds and rotors in the Falkland Islands. *Quarterly Journal of the Royal Meteorological Society*. **131**(605), pp. 329–351. DOI: 10.1256/qj.04.51.

Mohri, M., Rostamizadeh, A. and Talwalkar, A. 2018. *Foundations of Machine Learning*. 2nd. London: MIT Press.

Morcrette, C., Brown, K., Bowyer, R., Gill, P. and Suri, D. 2019. Development and Evaluation of In-Flight Icing Index Forecast for Aviation. *Weather and Forecasting*. **34**(3), pp. 731–750. DOI: 10.1175/WAF-D-18-0177.1.

Morris, R. M. 1986. The Spanish plume – testing the forecasters nerve. *Meteorological Magazine*. **115**(1372), pp. 349–357.

Mostajabi, A., Finney, D. L., Rubinstein, M. and Rachidi, F. 2019. Nowcasting lightning occurrence from commonly available meteorological parameters using machine learning techniques. *npj Climate and Atmospheric Science*. **2**(1), pp. 1–15. DOI: 10.1038/s41612-019-0098-0.

Murphy, J., Harris, G., Sexton, D., Kendon, E., Bett, P., Clark, R., Eagle, K., Fosser, G, Fung, F, Lowe, J., McDonald, R., McInnes, R., McSweeney, C., Mitchell, J., Rostron, J., Thornton, H., Tucker, S and Yamazaki, K. 2019. *UKCP18 Land Projections: Science Report*. Tech. rep. March 2019. Exeter: Met Office. URL: https://www.metoffice.gov.uk/pub/data/weather/uk/ukcp18/science-reports/UKCP18-Land-report.pdf.

Nance, L. B. and Durran, D. R. 1998. A Modeling Study of Nonstationary Trapped Mountain Lee Waves. Part II: Nonlinearity. *Journal of the Atmospheric Sciences*. **55**(8), pp. 1429–1445. DOI: 10.1175/1520-0469(1998)055<1429:AMSONT>2.0.CO;2. URL: http://journals.ametsoc.org/doi/10.1175/1520-0469(1998)055%3C1429:AMSONT%3E2.0.CO;2.

Nappo, C. J. 2002. *An Introduction to Atmospheric Gravity Waves*. San Diego: Academic Press.

Neal, R., Fereday, D., Crocker, R. and Comer, R. E. 2016. A flexible approach to defining weather patterns and their application in weather forecasting over Europe. *Meteorological Applications*. **23**(3), pp. 389–400. DOI: 10.1002/met.1563.

Niebler, S., Miltenberger, A., Schmidt, B. and Spichtinger, P. 2022. Automated detection and classification of synoptic-scale fronts from atmospheric data grids. *Weather and Climate Dynamics*. **3**(1), pp. 113–137. DOI: 10.5194/wcd-3-113-2022.

O'Shea, K. and Nash, R. 2015. An Introduction to Convolutional Neural Networks [Preprint]. URL: http://arxiv.org/abs/1511.08458.

Pan, S. J. and Yang, Q. 2010. A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering*. **22**(10), pp. 1345–1359. DOI: 10.1109/TKDE.2009.191.

Pandey, M., Fernandez, M., Gentile, F., Isayev, O., Tropsha, A., Stern, A. C. and Cherkasov, A. 2022. The transformational role of GPU computing and deep learning in drug discovery. *Nature Machine Intelligence*. **4**(3), pp. 211–221. DOI: 10.1038/s42256-022-00463-x.

Parr, T. and Howard, J. 2018. The Matrix Calculus You Need For Deep Learning [Preprint]. URL: http://arxiv.org/abs/1802.01528.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Köpf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J. and Chintala, S. 2019. PyTorch: An Imperative Style, High-Performance Deep Learning Library. *Advances in Neural Information Processing Systems*. **32**. URL: http://arxiv.org/abs/1912.01703.

Pathak, J., Subramanian, S., Harrington, P., Raja, S., Chattopadhyay, A., Mardani, M., Kurth, T., Hall, D., Li, Z., Azizzadenesheli, K., Hassanzadeh, P., Kashinath, K. and Anandkumar, A. 2022. FourCastNet: A Global Data-driven High-resolution Weather Model using Adaptive Fourier Neural Operators [Preprint]. URL: http://arxiv.org/abs/2202.11214.

Peak, J. E. and Tag, P. M. 1992. Toward Automated Interpretation of Satellite Imagery for Navy Shipboard Applications. *Bulletin of the American Meteorological*

Society. **73**(7), pp. 995–1008. DOI: 10.1175/1520-0477(1992)073<0995:TAIOSI>2.0.CO;2.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Müller, A., Nothman, J., Louppe, G., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M. and Duchesnay, E. 2012. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*. **12**(85), pp. 2825–2830. DOI: 10.48550/arXiv.1201.0490.

Perks, R. J., Bernie, D., Lowe, J. and Neal, R. 2023. The influence of future weather pattern changes and projected sea-level rise on coastal flood impacts around the UK. *Climatic Change*. **176**(3), pp. 1–21. DOI: 10.1007/s10584-023-03496-2.

Philipp, A., Della-Marta, P. M., Jacobeit, J., Fereday, D. R., Jones, P. D., Moberg, A. and Wanner, H. 2007. Long-Term Variability of Daily North Atlantic–European Pressure Patterns since 1850 Classified by Simulated Annealing Clustering. *Journal of Climate*. **20**(16), pp. 4065–4095. DOI: 10.1175/JCLI4175.1.

Pope, J. O., Brown, K., Fung, F., Hanlon, H. M., Neal, R., Palin, E. J. and Reid, A. 2022. Investigation of future climate change over the British Isles using weather patterns. *Climate Dynamics*. **58**(9-10), pp. 2405–2419. DOI: 10.1007/s00382-021-06031-0.

Prechelt, L. 1998. 'Early Stopping – But When?' *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, pp. 55–69.

Ralph, F. M., Neiman, P. J., Keller, T. L., Levinson, D. and Fedor, L. 1997. Observations, simulations, and analysis of nonstationary trapped lee waves. *Journal of the Atmospheric Sciences*. **54**(10), pp. 1308–1333. DOI: 10.1175/1520-0469(1997)054<1308:OSAAON>2.0.CO;2.

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N. and Prabhat. 2019. Deep learning and process understanding for data-driven Earth system science. *Nature*. **566**(7743), pp. 195–204. DOI: 10.1038/s41586-019-0912-1.

Riahi, K., Rao, S., Krey, V., Cho, C., Chirkov, V., Fischer, G., Kindermann, G., Nakicenovic, N. and Rafaj, P. 2011. RCP 8.5—A scenario of comparatively high greenhouse gas emissions. *Climatic Change*. **109**(1-2), pp. 33–57. DOI: 10.1007/s10584-011-0149-y.

Richardson, D., Fowler, H. J., Kilsby, C. G. and Neal, R. 2018. A new precipitation and drought climatology based on weather patterns. *International Journal of Climatology*. **38**(2), pp. 630–648. DOI: 10.1002/joc.5199.

Ronneberger, O., Fischer, P. and Brox, T. 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation [Preprint]. URL: http://arxiv.org/abs/1505.04597.

Ruff, F. and Ólafsson, H. 2019. Analysis of observed rapid increases in surface wind speed. *Quarterly Journal of the Royal Meteorological Society.* **145**(718), pp. 28–39. DOI: 10.1002/qj.3377.

Scorer, R. S. 1949. Theory of waves in the lee of mountains. *Quarterly Journal of the Royal Meteorological Society.* **75**(323), pp. 41–56. DOI: 10.1002/qj.49707532308.

Scorer, R. S. 1961. Lee Waves in the Atmosphere. *Scientific American.* **204**(3), pp. 124–137. DOI: 10.2307/24937396.

Sexton, D. M. H., McSweeney, C. F., Rostron, J. W., Yamazaki, K., Booth, B. B. B., Murphy, J. M., Regayre, L., Johnson, J. S. and Karmalkar, A. V. 2021. A perturbed parameter ensemble of HadGEM3-GC3.05 coupled model projections: part 1: selecting the parameter combinations. *Climate Dynamics.* **56**(11-12), pp. 3395–3436. DOI: 10.1007/s00382-021-05709-9.

Shapley, L. S. 1953. '17. A Value for n-Person Games'. *Contributions to the Theory of Games (AM-28), Volume II.* Vol. 114. 4. Princeton University Press, pp. 307–318. DOI: 10.1515/9781400881970-018.

Sheridan, P. F., Horlacher, V., Rooney, G. G., Hignett, P., Mobbs, S. D. and Vosper, S. B. 2007. Influence of lee waves on the near-surface flow downwind of the Pennines. *Quarterly Journal of the Royal Meteorological Society.* **133**(627), pp. 1353–1369. DOI: 10.1002/qj.110.

Sheridan, P. and Vosper, S. 2012. High-Resolution Simulations of Lee Waves and Downslope Winds over the Sierra Nevada during T-REX IOP 6. *Journal of Applied Meteorology and Climatology.* **51**(7), pp. 1333–1352. DOI: 10.1175/JAMC-D-11-0207.1.

Sheridan, P., Vosper, S. and Brown, P. 2017. Mountain Waves in High Resolution Forecast Models: Automated Diagnostics of Wave Severity and Impact on Surface Winds. *Atmosphere.* **8**(12), p. 24. DOI: 10.3390/atmos8010024.

Shi, X., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W.-k. and Woo, W.-c. 2015. 'Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting'. *Advances in Neural Information Processing Systems.* Ed. by C Cortes, N Lawrence, D Lee, M Sugiyama and R Garnett. Vol. 28. Curran Associates, Inc.

Shorten, C. and Khoshgoftaar, T. M. 2019. A survey on Image Data Augmentation for Deep Learning. *Journal of Big Data.* **6**(1), p. 60. DOI: 10.1186/s40537-019-0197-0.

Shutts, G. J. and Vosper, S. B. 2011. Stratospheric gravity waves revealed in NWP model forecasts. *Quarterly Journal of the Royal Meteorological Society.* **137**(655), pp. 303–317. DOI: 10.1002/qj.763.

Shutts, G. 1992. Observations and Numerical Model Simulation of a Partialy Trapped Lee Wave over the Welsh Mountains. *Monthly Weather Review.* **120**(9), pp. 2056–2066. DOI: 10.1175/1520-0493(1992)120<2056:OANMSO>2.0.CO;2.

Shutts, G. 1997. Operational lee wave forecasting. *Meteorological Applications.* **4**(1), pp. 23–35. DOI: 10.1017/S1350482797000340.

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., van den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T. and Hassabis, D. 2016. Mastering the game of Go with deep neural networks and tree search. *Nature.* **529**(7587), pp. 484–489. DOI: 10.1038/nature16961.

Simpson, R., Page, K. R. and De Roure, D. 2014. 'Zooniverse'. *Proceedings of the 23rd International Conference on World Wide Web.* New York, NY, USA: ACM, pp. 1049–1054. DOI: 10.1145/2567948.2579215.

Smith, R. 2003. 'Mountain Meteorology'. *Encyclopedia of Atmospheric Sciences.* Ed. by J. R. Holton, J. A. Curry and J. A. Pyle. Burlington: Elsevier, pp. 1400–1405. DOI: 10.1016/B0-12-227090-8/00239-6.

Staniforth, A. and Côté, J. 1991. Semi-Lagrangian Integration Schemes for Atmospheric Models—A Review. *Monthly Weather Review.* **119**(9), pp. 2206–2223. DOI: 10.1175/1520-0493(1991)119<2206:SLISFA>2.0.CO;2.

Steinley, D. 2006. K-means clustering: A half-century synthesis. *British Journal of Mathematical and Statistical Psychology.* **59**(1), pp. 1–34. DOI: 10.1348/000711005X48266.

Stockwell, R., Mansinha, L. and Lowe, R. 1996. Localization of the complex spectrum: the S transform. *IEEE Transactions on Signal Processing.* **44**(4), pp. 998–1001. DOI: 10.1109/78.492555.

Strauss, L., Serafin, S. and Grubišić, V. 2016. Atmospheric Rotors and Severe Turbulence in a Long Deep Valley. *Journal of the Atmospheric Sciences.* **73**(4), pp. 1481–1506. DOI: 10.1175/JAS-D-15-0192.1.

Strauss, L., Serafin, S., Haimov, S. and Grubišić, V. 2015. Turbulence in breaking mountain waves and atmospheric rotors estimated from airborne in situ and Doppler radar measurements. *Quarterly Journal of the Royal Meteorological Society.* **141**(693), pp. 3207–3225. DOI: 10.1002/qj.2604.

Tang, Y., Lean, H. W. and Bornemann, J. 2013. The benefits of the Met Office variable resolution NWP model for forecasting convection. *Meteorological Applications*. **20**(4), pp. 417–426. DOI: 10.1002/met.1300.

Taylor, K. E., Stouffer, R. J. and Meehl, G. A. 2012. An Overview of CMIP5 and the Experiment Design. *Bulletin of the American Meteorological Society*. **93**(4), pp. 485–498. DOI: 10.1175/BAMS-D-11-00094.1.

Teixeira, M. A. 2014. The physics of orographic gravity wave drag. *Frontiers in Physics*. **2**(July), pp. 1–24. DOI: 10.3389/fphy.2014.00043.

Udina, M., Bech, J., Gonzalez, S., Soler, M. R., Paci, A., Miró, J. R., Trapero, L., Donier, J. M., Douffet, T., Codina, B. and Pineda, N. 2020. Multi-sensor observations of an elevated rotor during a mountain wave event in the Eastern Pyrenees. *Atmospheric Research*. **234**. DOI: 10.1016/j.atmosres.2019.104698.

van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G. C., Kram, T., Krey, V., Lamarque, J.-F., Masui, T., Meinshausen, M., Nakicenovic, N., Smith, S. J. and Rose, S. K. 2011. The representative concentration pathways: an overview. *Climatic Change*. **109**(1-2), pp. 5–31. DOI: 10.1007/s10584-011-0148-z.

Vicari, R., Stephan, C. C., Lane, T. P. and Huang, Y. 2024. Analysis of Trapped Small-Scale Internal Gravity Waves Automatically Detected in Satellite Imagery. *Journal of Geophysical Research: Atmospheres*. **129**(4), pp. 1–17. DOI: 10.1029/2023JD038956.

Vosper, S. B. 2004. Inversion effects on mountain lee waves. *Quarterly Journal of the Royal Meteorological Society*. **130**(600), pp. 1723–1748. DOI: 10.1256/qj.03.63.

Vosper, S. B., Wells, H., Sinclair, J. A. and Sheridan, P. F. 2013. A climatology of lee waves over the UK derived from model forecasts. *Meteorological Applications*. **20**(4), pp. 466–481. DOI: 10.1002/met.1311.

Vosper, S. 2003. Development and testing of a high resolution mountain-wave forecasting system. *Meteorological Applications*. **10**(1), pp. 75–86. DOI: 10.1017/S1350482703005085.

Vosper, S., Ross, A., Renfrew, I., Sheridan, P., Elvidge, A. and Grubišić, V. 2018. Current Challenges in Orographic Flow Dynamics: Turbulent Exchange Due to Low-Level Gravity-Wave Processes. *Atmosphere*. **9**(9), p. 361. DOI: 10.3390/atmos9090361.

Vosper, S. B., Sheridan, P. F. and Brown, A. R. 2006. Flow separation and rotor formation beneath two-dimensional trapped lee waves. *Quarterly Journal of the Royal Meteorological Society*. **132**(620), pp. 2415–2438. DOI: 10.1256/qj.05.174.

Vosper, S. B., van Niekerk, A., Elvidge, A., Sandu, I. and Beljaars, A. 2020. What can we learn about orographic drag parametrisation from high-resolution models?

A case study over the Rocky Mountains. *Quarterly Journal of the Royal Meteorological Society*. **146**(727), pp. 979–995. DOI: 10.1002/qj.3720.

Walters, D., Baran, A. J., Boutle, I., Brooks, M., Earnshaw, P., Edwards, J., Furtado, K., Hill, P., Lock, A., Manners, J., Morcrette, C., Mulcahy, J., Sanchez, C., Smith, C., Stratton, R., Tennant, W., Tomassini, L., Van Weverberg, K., Vosper, S., Willett, M., Browse, J., Bushell, A., Carslaw, K., Dalvi, M., Essery, R., Gedney, N., Hardiman, S., Johnson, B., Johnson, C., Jones, A., Jones, C., Mann, G., Milton, S., Rumbold, H., Sellar, A., Ujiie, M., Whitall, M., Williams, K. and Zerroukat, M. 2019. The Met Office Unified Model Global Atmosphere 7.0/7.1 and JULES Global Land 7.0 configurations. *Geoscientific Model Development*. **12**(5), pp. 1909–1963. DOI: 10.5194/gmd-12-1909-2019.

Washington, R. and Palmer, M. 1999. The North Atlantic Oscillation. *Geography Review*. **13**(2), pp. 2–5. DOI: 10.1017/cbo9781139034135.008.

WCRP. 2024. *WCRP Coupled Model Intercomparison Project (CMIP)*. [Online]. Last accessed: 2024-03-15. URL: https://www.wcrp-climate.org/wgcm-cmip.

Weiss, K., Khoshgoftaar, T. M. and Wang, D. 2016. A survey of transfer learning. *Journal of Big Data*. **3**(1), p. 9. DOI: 10.1186/s40537-016-0043-6.

Wilkinson, J. M. and Neal, R. 2021. Exploring relationships between weather patterns and observed lightning activity for Britain and Ireland. *Quarterly Journal of the Royal Meteorological Society*. **147**(738), pp. 2772–2795. DOI: 10.1002/qj.4099.

Wilson, R., Chanin, M. L. and Hauchecorne, A. 1991. Gravity waves in the middle atmosphere observed by Rayleigh lidar: 2. Climatology. *Journal of Geophysical Research: Atmospheres*. **96**(D3), pp. 5169–5183. DOI: 10.1029/90JD02610.

Wold, S., Esbensen, K. and Geladi, P. 1987. Principal component analysis. *Chemometrics and Intelligent Laboratory Systems*. **2**(1-3), pp. 37–52. DOI: 10.1016/0169-7439(87)80084-9.

Wood, N., Staniforth, A., White, A., Allen, T., Diamantakis, M., Gross, M., Melvin, T., Smith, C., Vosper, S., Zerroukat, M. and Thuburn, J. 2014. An inherently mass-conserving semi-implicit semi-Lagrangian discretization of the deep-atmosphere global non-hydrostatic equations. *Quarterly Journal of the Royal Meteorological Society*. **140**(682), pp. 1505–1520. DOI: 10.1002/qj.2235.

World Meteorological Organization. 1993. *Handbook of Meteorological Forecasting for Soaring Flight*. Tech. rep. WMO - No. 495. Geneva: World Meteorological Organization, p. 83.

Worthington, R. M. 2006. Diurnal variation of mountain waves. *Annales Geophysicae*. **24**(11), pp. 2891–2900. DOI: 10.5194/angeo-24-2891-2006.

Wright, C. J., Hindley, N. P., Hoffmann, L., Alexander, M. J. and Mitchell, N. J. 2017. Exploring gravity wave characteristics in 3-D using a novel S-transform technique: AIRS/Aqua measurements over the Southern Andes and Drake Passage. *Atmospheric Chemistry and Physics*. **17**(13), pp. 8553–8575. DOI: 10.5194/acp-17-8553-2017.

Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C. and Yu, P. S. 2021. A Comprehensive Survey on Graph Neural Networks. *IEEE Transactions on Neural Networks and Learning Systems*. **32**(1), pp. 4–24. DOI: 10.1109/TNNLS.2020.2978386.

Wurtele, M. G., Sharman, R. D. and Datta, A. 1996. Atmospheric Lee Waves. *Annual Review of Fluid Mechanics*. **28**(1), pp. 429–476. DOI: 10.1146/annurev.fl.28.010196.002241.

Yamazaki, K., Sexton, D. M., Rostron, J. W., McSweeney, C. F., Murphy, J. M. and Harris, G. R. 2021. A perturbed parameter ensemble of HadGEM3-GC3.05 coupled model projections: part 2: global performance and future changes. *Climate Dynamics*. **56**(11-12), pp. 3437–3471. DOI: 10.1007/s00382-020-05608-5.

Yu, X.-H. and Chen, G.-A. 1997. Efficient Backpropagation Learning Using Optimal Learning Rate and Momentum. *Neural Networks*. **10**(3), pp. 517–527. DOI: 10.1016/S0893-6080(96)00102-5.

Zhang, P., Ke, Y., Zhang, Z., Wang, M., Li, P. and Zhang, S. 2018. Urban Land Use and Land Cover Classification Using Novel Deep Learning Models Based on High Spatial Resolution Satellite Imagery. *Sensors*. **18**(11), p. 3717. DOI: 10.3390/s18113717.

Appendix A

Distribution of lee wave occurrence and characteristics in the future climate scenario

As an addition to the PDFs shown in Chapter 6 (Figures 6.5 and 6.7), which show the change in distribution between the present-day climate time slice (TS1) and the future climate time slices (TS2 and TS3), these plots are just the PDFs of the occurrence and characteristics of lee waves (analogues of Figures 5.5 and 5.7, but for the future climate).

Figure A.1 shows PDFs for lee wave frequency of occurrence and characteristics by hour of the day, for TS3 (the plots for TS2 look very similar). There is no noticeable diurnal cycle in lee wave occurrence or characteristics. The very small changes in amplitude (b) are at very small probabilities ($< 10^{-7}$). The distributions of lee wave wavelength (c) and orientation (d) remain similar at every time during the day in the data.

Figure A.2 shows probability density functions (PDFs) of lee wave frequency of occurrence and characteristics (amplitude, wavelength and orientation) for the 8 different synoptic weather patterns for TS2 (a) and TS3 (b). The distributions are similar to those in TS1. Most lee waves occur under NAO+ and Southwesterly conditions, and NAO+ has conditions that produce the longest wavelength lee waves. Orientations under (for example) NAO+ and High pressure centred over UK con-

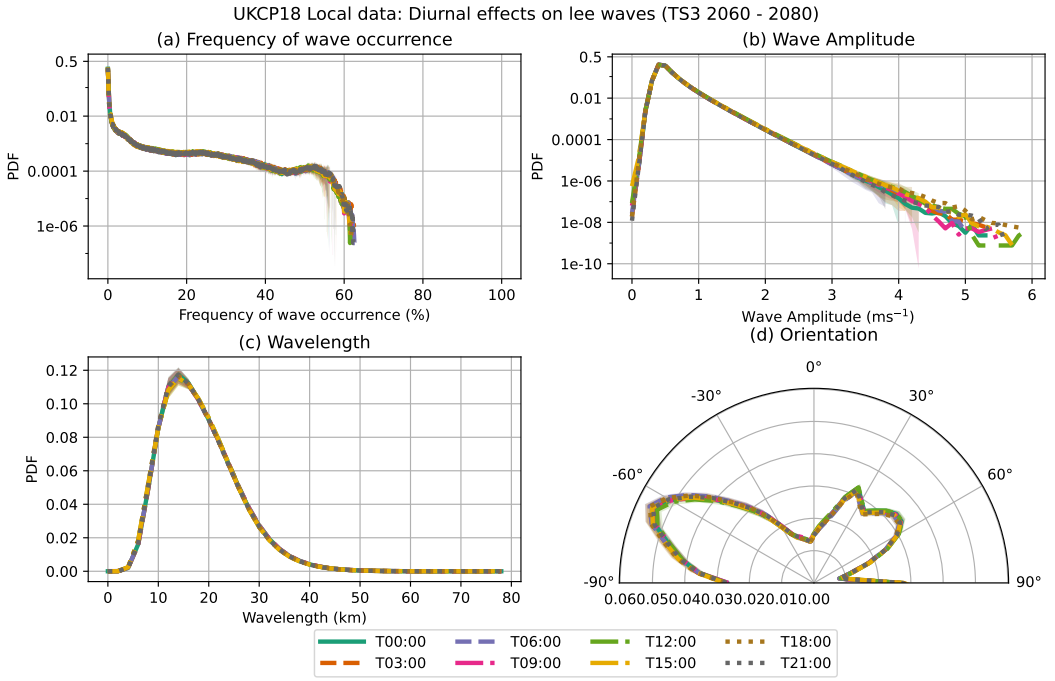


Figure A.1: PDFs of lee wave frequency of occurrence (a) and characteristics (b-d) for TS3, by hour of the day. The plots for TS2 are very similar, showing very little change in lee wave prevalence or characteristics between hours of the day.

ditions exhibit a bimodal distribution, while other patterns such as Northwesterly and Southwesterly are more unimodal and exhibit a distribution in line with their name.

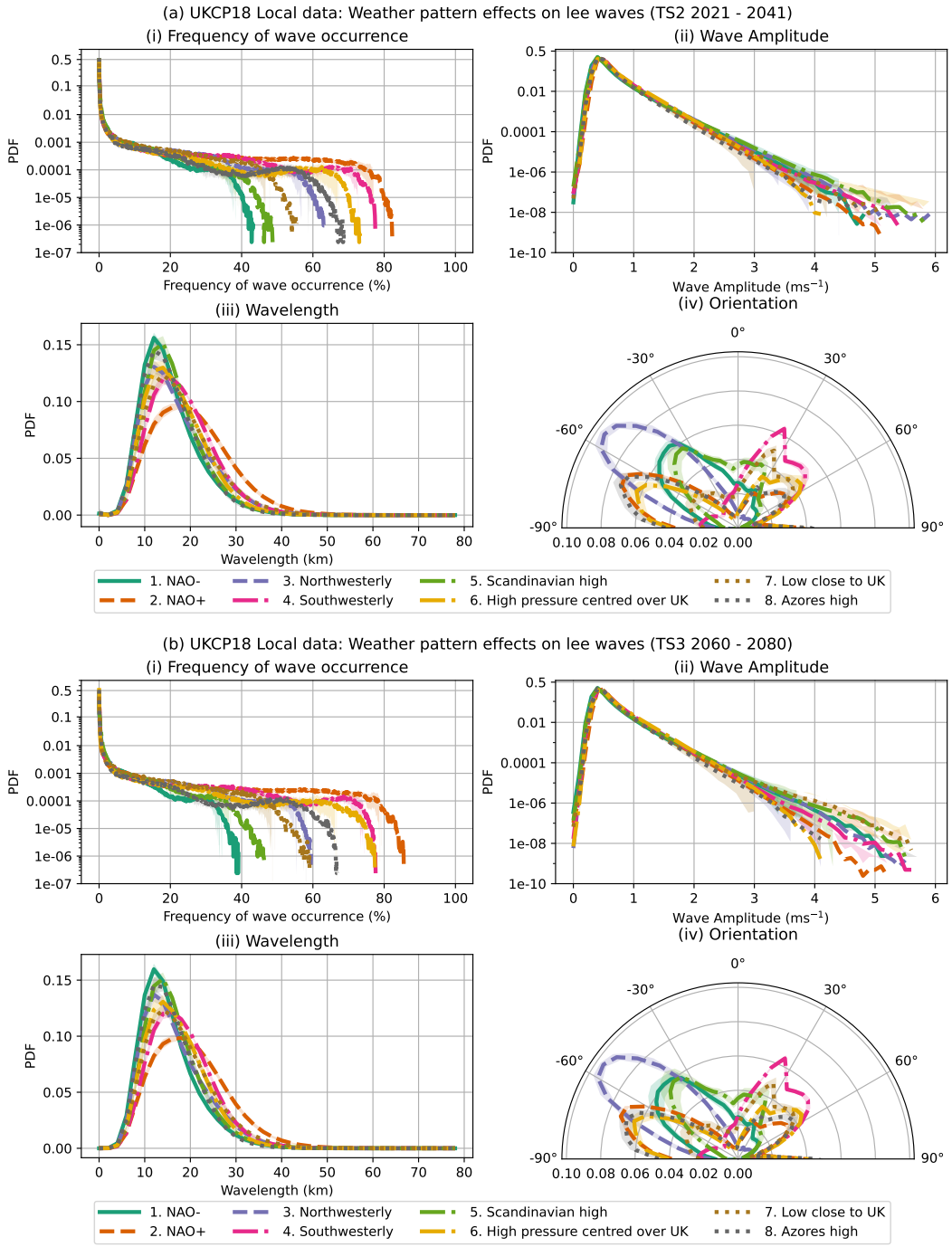


Figure A.2: The ensemble mean frequency of occurrence and characteristics of lee waves for the 8 weather patterns in TS2 (a) and TS3 (b). Shading either side of each line indicates 1 standard deviation away from the ensemble mean.