The use of Gaussian process regression for wind forecasting in the UK

Victoria Ruth Hoolohan

Submitted in accordance with the requirements for the degree of

Doctor of Philosophy

as part of the integrated PhD & MSc in Low Carbon Technologies

University of Leeds

School of Chemical and Process Engineering

Doctoral Training Centre for Low Carbon Technologies

February 2018

Academic declaration

The candidate confirms that the work submitted is her own and that appropriate credit has been given where reference has been made to the work of others.

Victoria Ruth Hoolohan

Acknowledgements

I would like to thank my friends, family and colleagues for all their support throughout the course of writing this thesis, it certainly wouldn't have been possible without them. In particular, I would like to thank my supervisors Professors Alison Tomlin and Tim Cockerill. Their help, advice and ideas have been invaluable throughout the last 4 years. I am also grateful to Dr Tom Dunstan, Professor Simon Vosper and Alasdair Skea at the UK Met office for their guidance, expertise and access to the forecast data which has been crucial to my research.

I would also like to thank the staff and students of the University of Leeds Doctoral Training Centre in Low Carbon Technologies. Thank you for creating such an interesting and encouraging environment in which to work, I feel lucky to have been a part of the DTC. In particular, many thanks to my fellow students of the 2013 DTC cohort; Kate Palmer, Dougie Phillips, Aidan Smith, Katrina Adam, Rob Bloom, Kelly Marsh, James Hammerton and Niamh Ryall. Thank you for making the process of writing this thesis mostly enjoyable and for maintaining the endless supply of coffee. In addition to this, thank you to the previous DTC student (of whom there are too many to mention) for providing inspiration and encouragement. I am grateful to James McKay, Emily Bryan-Kinns and David Haynes for organising and supporting me throughout my time in the DTC, and to Professor Paul Williams for creating this opportunity to study in the first place.

Thank you to my friends outside of the DTC, to all of those who provided a welcome distraction from my research when it was needed most, for dragging me out and boosting productivity massively. I am also eternally grateful to my parents for encouraging me to learn and for supporting me every step of the way. To my sister, Dr. Claire Hoolohan for making me aim higher, for always listening and giving the best advice. And finally, to Carl, for being by my side from start to finish and for the many adventures along the way – thanks doesn't quite cut it!

Abstract

Wind energy has experienced remarkable growth in recent years, both globally and in the UK. As a low carbon source of electricity this progress has been, and continues to be, encouraged through legally binding targets and government policy. However, wind energy is non-dispatchable and difficult to predict in advance. In order to support continued development in the wind industry, increasingly accurate prediction techniques are sought to provide forecasts of wind speed and power output.

This thesis develops and tests a hybrid numerical weather prediction (NWP) and Gaussian process regression (GPR) model for the prediction of wind speed and power output from 3 hours to 72 hours in advance and considers the impact of incorporating atmospheric stability in the prediction model. In addition to this, the validity of the model as a probabilistic technique for wind power output forecasting is tested and the economic value of a forecast in the UK electricity market is discussed.

To begin with, the hybrid NWP and GPR model is developed and tested for prediction of 10 m wind speeds at 15 sites across the UK and hub height wind speeds at 1 site. Atmospheric stability is incorporated in the prediction model first by subdividing input data by Pasquill-Gifford-Turner (PGT) stability class, and then by using the predicted Obukhov length stability parameter as an input in the model. The model is developed further to provide wind power output predictions, both for a single turbine and for 22 wind farms distributed across the UK. This shows that the hybrid NWP and GPR model provide good predictions for wind power output in comparison to other methods. The hybrid NWP and GPR model for the prediction of near-surface wind speeds leads to a reduction in mean absolute percentage error (MAPE) of approximately 2% in comparison to the Met office NWP model. Furthermore, the use of the Obukhov length stability parameter as an input reduces wind power prediction errors in comparison to the same model without this parameter for the single turbine and for offshore wind farms but not for onshore wind farms. The inclusion of the Obukhov length stability parameter in the hub height wind speed prediction model leads to a reduction in MAPE of between 2 and

5%, dependent on the forecast horizon, over the model where Obukhov length is omitted. For the prediction of wind power at offshore wind farms, the inclusion of the Obukhov length stability parameter in the hybrid NWP and GPR model leads to a reduction in normalised mean absolute error (NMAE) of between 0.5 and 2%. The performance of the hybrid NWP and GPR model is also evaluated from a probabilistic perspective, with a particular focus on the appropriate likelihood function for the GPR model. The results suggest that using a beta likelihood function in the hybrid model for wind power prediction leads to better probabilistic predictions than implementing the same model with a Gaussian likelihood function. The results suggest an improvement of approximately 1% in continuous ranked probability score (CRPS) when the beta likelihood function is used rather than the Gaussian likelihood function.

After considering new techniques for the prediction of wind speed and power output, the final chapter in this thesis considers the economic benefit of implementing a forecast. The economic value of a wind power forecast is evaluated from the perspective of a wind generator participating in the UK electricity market. The impact of forecast accuracy and the change from a dual imbalance price to a single imbalance price is investigated. The results show that a reduction in random error in a wind power forecast does not have a large impact on the average price per MWh generated. However, it has a more significant impact on the variation in price received on an hourly basis. When the systematic bias in a forecast was zero, a forecast with NMAE of 20% of capacity results in less than £0.05 deviation in mean price per MWh in comparison with a perfect forecast. However, the same forecast leads to an increase in standard deviation of up to £21/MWh. This indicates that whilst a reduction in random error in a forecast might not lead to an improvement in mean price per MWh, it can lead to a more stable income stream. In addition to this, Chapter 6 considers the use of the probabilistic and deterministic forecasts developed throughout this thesis to choose an appropriate value to bid in the UK electricity market. This shows that using a probabilistic forecast can limit a generator's exposure to variable prices and decrease the standard deviation in hourly prices.

Contents

List of Figures

List of Tables

Glossary of Terms and Symbols

Symbols:

x	Scalars are represented by italic text
$\mathbf{x}$	Vectors are represented by a bold lowercase letter
$\mathbf{X}$	Matrices are represented by a bold uppercase letter
x_i	Elements of a vector are represented by a single subscript
x_{ij}	Elements of a matrix are represented by a subscript for the column and row
y_t	Observed wind speed at time *t*
$\widehat{y_t}$	Predicted wind speed at time *t*
ϵ_t	Error in predicted wind speed at time *t*
$\mathbf{x}_{ij}$	A row of a matrix is represented by a bold lower case letter with a subscript.
$\mathbf{v}$	Velocity vector (*u,v,w*) for a flow
$\mathbf{k}$	Vertical unit vector
t	Time
f	Coriolis parameter
p	Pressure of gas in a flow (Pa)
ρ	Density of air (kg m^{-3})
F	Force of friction (N)
R	Specific gas constant (J $kg^{-1}K^{-1}$)
T	Temperature (K)
θ	Potential temperature (K)
u_*	Friction velocity (ms^{-1})
z_0	Surface roughness (m)
C_{power}	Power coefficient
C	Wind farm capacity
g	Acceleration due to gravity = 9.81 ms^{-2}
H_{sen}	Sensible surface heat flux (Wm^{-2})

C_p	Specific heat of air (kJ $kg^{-1}K^{-1}$)
μ	Mean value
σ	Standard deviation
N(μ,σ)	Normal distribution with mean, μ and standard deviation σ.
$\sim$	Has distribution e.g. X ~ N(0,1) means "random variable X is normally distributed with mean 0 and standard deviation 1".
$[a,b]$	The set of numbers between and inclusive of a and b
$\in$	Belongs within the set e.g. $x \in [0,100]$ means "x belongs within the set [0,100]" or "$0 \leq x \leq 100$".
Φ	Cumulative distribution function of standard normal distribution (N(0,1))
φ	Probability density function of standard normal distribution (N(0,1))
B(α,β)	Beta distribution with shape parameters α,β.
$\Gamma(n)$	Gamma function $= (n-1)!$
H(x)	Heaviside step function $= \frac{d}{dx}\max\{x, 0\}$
L	Obukhov length (m)
K(X,X)	Covariance matrix
m(X)	Mean function
F_t	Cumulative distribution function at time t
F_t^{-1}	Inverse cumulative distribution function at time t
$\hat{F}_t$	Predicted cumulative distribution function at time t
R_t	Price received for electricity generated at time t (£/MWh)
$\overline{R}_{MWh}$	Mean price received for electricity generated (£/MWh)
B_t	Bid volume at time t (MWh)
MIP_t	Market index price at time t (£/MWh)
SIP_t	Single imbalance price at time t (£/MWh)
TP_t	Trade price at time t (£/MWh)

Abbreviations:

ANN	Artificial neural network
BEIS	(Department for) Business, Energy and Industrial Strategy
BSC	Balancing and settlement code
CDF	Cumulative distribution function
CFD	Contracts for difference
CRPS	Continuous ranked probability score
DECC	Department of Energy and Climate Change
EMR	Electricity market reform
EPS	Emissions trading scheme
ETS	Electricity market reform
GHG	Greenhouse gas
GPML	Gaussian processes for machine learning
GPR	Gaussian process regression
IPCC	Intergovernmental Panel on Climate Change
MAE (NMAE)	Mean absolute error (normalised)
MAPE	Mean absolute percentage error
MBE (NMBE)	Mean bias error (normalised)
MIDAS	Met Office integrated data archive system
MIP	Market index price
NGET	National Grid electricity transmission
NRI	Net radiation index
NWP	Numerical weather prediction
PAR	Price average reference
PBL	Planetary boundary layer
pdf	Probability density function
PGT	Pasquill Gifford Turner (stability classification)
RMSE (NRMSE)	Root mean squared error (normalised)
ROC	Renewable obligation certificate
SIP	Single imbalance price
SSB	System buy price
SSP	System sell price

Chapter 1. Introduction

1.1 Background and policy

Extensive scientific evidence over many years has shown that human activity is playing a key role in altering the global climate [1]. In the last 60 years, average sea surface temperatures have exceeded all previous recordings, global sea levels have risen and sea ice has receded [2]. Increased consumption of fossil fuels since the industrial revolution has resulted in increased concentration of CO_2 and other greenhouse gases (GHG) in the atmosphere. If nothing is done to change current trends, it is likely that global warming will have a significant effect on ecological systems and human life, through increases in extreme weather events, effects on food and water supplies and rising sea levels. Whilst exact consequences of climate change cannot be predicted it is important that GHG emissions are cut to reduce long-term impacts [3].

The potential impact of climate change, socially, politically and economically, has been acknowledged by international leaders for many years. In 1992 the United Nations Framework Convention on Climate change (UNFCCC) was implemented in order to establish a framework for stabilising GHG emissions to avoid potentially dangerous climate change [4]. This was extended in 1997 with the Kyoto agreement, which took effect in 2005 and gave individual limits on GHG emissions for countries within the agreement [5]. In 2015 the Paris agreement was negotiated, aimed at strengthening current actions to combat climate change [6].

Following the agreements outlined above targets for GHG emission reductions have been established by some parties. Within the European Union (EU) the 2020 package was legislated in 2009 in order to ensure targets were met for 2020 [7]. The three key targets were:

(i) 20% reduction in GHG emissions from 1990 levels;
(ii) 20% of EU energy to be generated from renewable sources;

(iii) 20% improvements in energy efficiency.

These were updated in 2014 [8] to extend targets for 2030 with new targets of:

(i) At least a 40% cut in GHG emissions from 1990 levels;
(ii) At least a 27% of EU energy to be generated from renewable sources;
(iii) At least a 27% improvement in energy efficiency.

The UK government has outlined measures to limit and reduce UK emissions of GHG, through objectives protected by law. The 2008 climate change act includes a system for carbon budgeting and outlines a number of targets [9]. Of these the most notable being the aim to reduce GHG emissions by 80% in 2050 based on 1990 levels. In the short term the act outlines 'carbon budgets' which provide a framework for five-year limits on GHG emissions. The fifth carbon budget outlined in 2016 covering 2028 – 2032 requires a reduction in GHG emissions of 57% compared to 1990 levels [10]. Furthermore, the UK government defined a target of 15% of total energy demand being met by renewables by 2020. This objective included three sub-targets, one of which being to generate 30% of electricity from renewable sources which would be achieved by increasing renewable energy operational capacity to 29 GW [11]. Reaching these targets will not be an easy task, and significant changes will need to be made to current status quo in order to reduce emissions. Increasing the use of renewable electricity and reducing the use of gas and coal to generate electricity is fundamental to achieving these targets.

1.2 Development and deployment of renewable electricity

Globally, 21.6% of electricity produced came from renewables in 2015 [12]. In Europe, this figure was 30% and in the UK 17.6%. The main source of renewable energy worldwide is hydropower, with over 60% of renewable electricity generated by hydropower in 2015 [12]. However, wind energy is a rapidly growing source of renewable energy. Wind energy is an abundant source of renewable energy and the cost of wind energy projects is decreasing as technology advances. This suggests that wind energy will be a significant part of delivering a low cost, low emissions electricity system in the future. The importance of wind energy in the meeting of

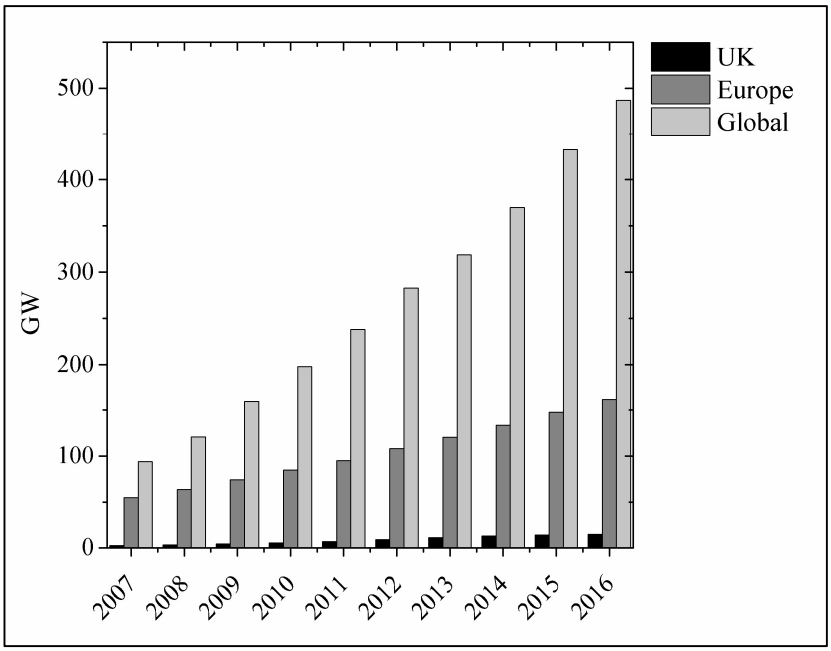
UK renewable energy targets is highlighted by the UK renewable energy roadmap, a document produced by the former UK Department of Energy and Climate change (DECC) to outline how the targets in Section 1.1 will be achieved. From this, it is expected that between 57 and 90 TWh of renewable electricity will come from wind energy by 2020, both onshore and offshore, as shown in Table 1-1. This is larger than the contribution of any other source.

Table 1-1: UK government renewable energy output targets for 2020 [11]

Renewable energy type	Annual output target for 2020 (TWh)
Onshore wind	24-32
Offshore wind	33-58
Biomass electricity	32-50
Marine electricity	1
Biomass heat (non-domestic)	36-50
Air source and ground source heat pumps (non-domestic)	16-22
Others (including hydro, geothermal, solar and domestic heat)	14
Renewable transport	Up to 48
Total to meet 15% target	234

Since 2007, progress has been made in increasing wind energy capacity globally. Figure 1-1 shows wind energy capacity from 2007 to 2016, showing that capacity has significantly increased worldwide.

Figure 1-1: Installed wind energy capacity from 2007-2016 [13].



In addition to the global impact of wind energy, this form of renewable energy also contributes significantly to the UK electricity supply. It can be seen in Table 1-2 that in 2012 around 11% of electricity in the UK was generated from renewable sources, which rose to nearly 25% by 2015 [14]. Of this, nearly 50% came from wind energy, as highlighted in Figure 1-2.

Table 1-2: Renewable energy contributions in the UK [14, 15]

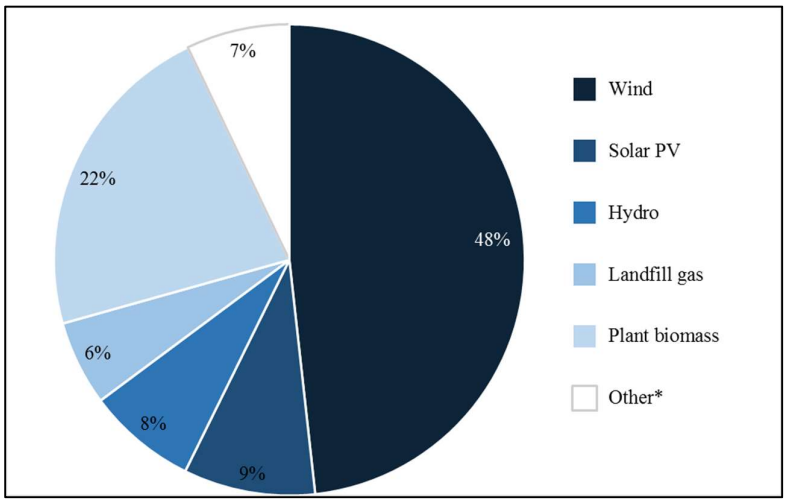
Year	Total renewables (% of electricity generation)
2012	11.3%
2013	14.9%
2014	19.1%
2015	24.6%
2016	24.4%

Despite increases in the amount of electricity generated from renewable sources between 2012 and 2015, 2016 saw a plateau in renewable electricity generation. Renewables' share of electricity generation in 2016 fell by 0.2% in comparison to 2015 despite increases in installed capacity [15]. Furthermore, total electricity generated from wind fell by 7% between 2015 and 2016 [15]. According to a

summary report by the UK Department for business, energy and industrial strategy (BEIS) this was due to an 11% fall in average wind speed [16]. This indicates that simply increasing installed capacity of renewable energy is insufficient to reach renewable energy targets in terms of electricity generated from renewable sources.

Looking to the future, it is anticipated the increase in installed capacity of wind energy will continue in the UK. As of September 2017 there were nearly 7600 turbines operational in the UK, providing 17.1 GW of installed capacity [17]. A further 4.6 GW was under construction and 16.4 GW currently had planning permission, which if all projects are completed will deliver targets for 2020, in terms of capacity installed. Of the wind farm capacity currently operational in the UK, 35% is onshore and 65% is offshore. Of the 16.4 GW with planning permission for the future, 14% is onshore and 86% is offshore. This demonstrates a clear move towards offshore wind power production in the future of the UK electricity supply.

Figure 1-2: UK Renewable energy generation in 2015. [18]



*Other includes electricity from wave, tidal, sewerage, co-firing with fossil fuels, anaerobic digestion, and animal biomass.

Alongside the increase in installed capacity of wind energy, technical advancement in size and efficiency of wind turbines has helped increase wind power output. The size and height of turbines have grown significantly in the past 30 years. Commercial offshore turbines are now normally 3 MW or larger [17]. The largest commercially available turbine is currently the Vestas V164 9.5 MW turbine, and a

number of turbines are available with rated capacity exceeding 7 MW [19]. Onshore turbines are now normally between 1 and 3.5 MW [17]. The increase in size increases the power which can be generated from individual turbines; however also limits where they can be placed. Large wind developments are crucial to reaching 2020 government targets. As of September 2017, the largest operational wind farm in the UK was the London Array, a 630 MW development of 175 turbines. A further 5 offshore wind farms with rated capacity over 1GW currently have planning permission, as detailed in Table 1-3, offering 9 GW of further capacity [17]. This increase in size of wind farm projects has helped deliver sufficient capacity to reach renewable energy targets.

Table 1-3: Wind farms over 1 GW with planning consent [17]

Project	Location	Installed capacity (MW)	Turbine capacity (MW)	No. of turbines
Dogger Bank Creyke Beck A & B	125km off coast of Yorkshire	2,400	5.0	480
Dogger Bank Teesside A & B	125km off coast of Yorkshire	2,400	5.0	480
Hornsea Project Two – Optimus and Breesea	Off coast of Yorkshire	1,800	5	360
East Anglia 3	Offshore, East of England	1,200	7	172
Hornsea Project One - Heron & Njord	Off coast of Yorkshire	1,200	7	174

Despite significant progress towards achieving targets in producing electricity from renewable sources, there is much more work required than simply building sufficient renewable energy capacity. As seen in Table 1-2 and Figure 1-1, despite an increase in installed wind energy capacity in the UK between 2015 and 2016 there was a decrease in the proportion of electricity generated from renewable sources. Further work is required to ensure that a low carbon electricity system is achieved whilst maintaining the security of supply and providing electricity at an affordable rate. In addition to this, large wind farms and large individual wind turbines present new problems for the electricity grid in terms of power balancing and planning, such as

the potential for significant highs and lows in power output dependent on regional weather events. In order to ensure increased wind energy capacity can be effectively used to provide low carbon electricity, the integration of wind energy in the electricity system needs to be effectively managed. In the next sections, some of the issues in integration of wind energy are discussed.

1.3 Challenges of integrating wind energy

Integration in electricity systems is the process of coordinating the operation of numerous individual power generation units to provide a reliable and cost-effective electricity supply. To do this, demand for electricity must be met by generation in real time. In contrast to traditional sources of power, renewable energy (particularly non-dispatchable renewable energy such as wind and solar power) cannot be produced on demand and can only be predicted with limited accuracy. This creates a number of issues. Firstly, the uncertain nature of wind power available leads to scheduling difficulties [20]. As electricity needs to be generated to meet demand, predictions of demand and generation are made in advance. This allows power generation units to be scheduled to supply sufficient power to meet demand. Electricity storage systems can be used to aid this process of balancing by storing electricity when production is greater than expected and discharging when electricity production is less than expected. However, currently electricity storage systems are used infrequently due to the costs associated. Therefore the issues related to the integration of wind energy considered here assume large scale electricity storage systems are unavailable. Prediction of electricity demand is relatively well understood but still contains some uncertainty. Uncertainty in prediction of renewable energy generation adds to the difficulty of scheduling power delivery. Secondly, uncertainty in renewable energy generation can lead to additional actions to balance the system being needed. This can increase costs. Finally, variations and uncertainty in renewable energy production require alternative load to be available to meet demand. For example, coal or gas plants may be required to change power output. These changes can increase wear and tear and decrease operating efficiency. This both increases costs and limits the emissions reductions achieved through the use of renewable electricity.

In the UK, the integration of a large amount of wind energy has been managed so far. However, the cost of balancing the system has been increasing over the last decade due to an increase in supply from intermittent generators [21]. If energy supply cannot be effectively balanced, some generators may be forced to stop producing electricity, known as curtailment. For wind energy this is significant as the wind farms can only produce power at certain times, if this power cannot be used, income cannot be maximised. System management and balancing are discussed further in Section 2.3.1. However, the examples shown here indicate some of the issues with increasing use of wind energy and other forms of variable renewable energy.

Despite the integration of wind energy being a complex task, significant progress has been made in other countries. For example in Denmark in 2015 42% of electricity consumed was produced by wind energy [22]. In Denmark, a high penetration of wind is managed using interconnections with Norway, Sweden, and Germany [23]. Given the relatively small size of Denmark's electricity system in comparison with connected countries, this allows the fluctuations of wind power to be balanced with the relatively reliable and stable power provided by hydropower in Norway. This is an example of the different natural resources available in neighbouring countries being used to complement one another. Alternatively, in Spain, substantial interconnections like those that exist in Denmark are not available. Spain has one connection with France which covers approximately 2% of installed capacity which is insufficient to cover potential voltage drops caused by renewable energy [23]. To combat balancing issues, large suppliers in Spain are mandated to forecast their energy production. The centre for control of renewable energy (CIEMAT) monitors the production of renewables over 10 MW in real time and uses this control to balance supply and demand [23]. These are two of the different ways of dealing with the balancing issues presented by the use of non-dispatchable renewable energy sources.

In the UK, large interconnections are not available. Large generators are required to submit their expected power output to the system operator, the National Grid, half an hour in advance of generation. The National Grid is then responsible for grid balancing. The cost of deviating from the expected electricity production is passed

on the generator through an imbalance price. If a substantial number of deviations from the predicted schedule need to be balanced this imbalance price may rise significantly. Therefore there is an incentive for generators to accurately predict power output, ultimately enabling the system operator to manage numerous variable electricity producers. This method relies on market forces to enable effective integration of variable electricity generators. This can make it challenging for them to be economically competitive in a system still dominated by well-established thermal generators such as gas power plants.

1.4 Need for research

So far, it has been identified that there is a need to continue increasing the use of renewable electricity and that wind energy can contribute to this. However, issues with the integration of wind energy in an electricity system have been acknowledged and these need to be managed if the use of wind energy is to be increased in a cost-effective manner.

In the UK significant interconnections with other power systems are not available. However, wind power prediction could be used to limit the impact of increased use of wind power in the electricity system. For example, wind power prediction can be used to effectively schedule generating units to meet demand. Alternatively, it can be used to set the amount of excess capacity required to ensure reliable delivery of electricity. Furthermore, it can be used to allow wind generators to effectively compete in electricity markets and increase revenue. This, in turn, encourages investment in wind energy in the future. The contribution of wind power forecasting is not limited to the examples given above. A more detailed discussion of the applications of wind power is given in Section 2.3. However, the examples shown here indicate that wind power forecasting can play an important role in ensuring reliable and efficient operation of an electricity system with increased wind energy capacity. This will allow the targets outlined in Section 1.1 to be met and a low carbon electricity system to be realised. Because of this, new wind power forecasting techniques and methods which may increase forecasting accuracy continue to be important. Many forecasting techniques already exist, an overview of

which is given in Section 2.2. However, improvements can still be made and it is important to consider alternative techniques. Furthermore, the use of forecasts which provide a way of quantifying uncertainty is of growing interest. Quantification of uncertainty can help minimise or hedge against risk of financial loss or manage capacity requirements in a stochastic way in order to reduce system costs. Probabilistic forecasting is one way of providing an estimate of prediction uncertainty. Therefore methods for probabilistic wind power forecasting are an important area of research.

1.5 Aims and objectives

The overall aim of this thesis is to develop and test a method for short to medium term wind power forecasting. Here short to medium term refers to forecasts from 3 hours to 72 hours in advance. The method employed is a hybrid numerical weather prediction (NWP) and Gaussian process regression (GPR) model. This is first used to predict near surface wind speeds and then adapted to provide wind power predictions for both an individual turbine and a whole wind farm.

The main objectives of this thesis are:

(i) To investigate the use of GPR as a method for the prediction of wind speed and power output, focusing on how it can be used to improve upon NWP wind predictions and make wind power predictions.

(ii) To explore whether incorporating atmospheric stability into the GPR model leads to improvements when predicting wind speed and power output. This thesis will consider the atmospheric conditions under which NWP wind speed predictions could be improved and develop this to make wind power predictions.

(iii) To study whether a GPR model is appropriate as a probabilistic method for wind power forecasting and how best to formulate the model to make good probabilistic wind power predictions.

(iv) To evaluate the value of wind power forecasting from a user perspective, particularly focusing on the economic value of a forecast for a wind generator participating in the UK electricity market.

1.6 Outline of thesis

Chapter 2 begins by giving an outline of the concepts fundamental to wind and wind power generation. Next, a discussion of wind speed and power forecasting techniques is given. This covers both basic techniques and current state of the art methods. This includes an overview of current literature on wind speed and power forecasting techniques. In addition, applications of wind forecasting are discussed. A summary of electricity system management and electricity markets is given, followed by a discussion of current literature on the applications of wind forecasting.

In Chapter 3 a hybrid NWP and GPR model for wind speed prediction is developed. The model is used for the prediction of 10 m wind speeds at 15 sites across the UK and hub height wind speeds at one turbine. The model for the prediction of 10 m wind speed aims to correct NWP predictions, whilst for hub height wind speed prediction the model translates NWP predictions to hub height predictions. The impact of subdividing input data by atmospheric stability class on model performance is also investigated. Furthermore, the influence of improved wind speed predictions on power predictions is considered for one turbine.

In Chapter 4 the hybrid NWP and GPR model is developed further to give power predictions both for a single wind turbine and for 22 wind farms across the UK. In addition to this, the use of the Obukhov length stability parameter is introduced, and the impact of its inclusion as an input parameter in the GPR model on model performance is investigated.

Chapter 5 evaluates the hybrid NWP and GPR model for wind power prediction developed in Chapter 4 in a probabilistic framework. This allows the performance of the model for probabilistic predictions to be evaluated. In addition to this, the use of a beta likelihood function in the GPR model is discussed in comparison to a Gaussian likelihood function.

Chapter 6 evaluates the economic value of a wind power forecast from the perspective of a wind generator participating in the UK electricity market. The impact of forecast accuracy on price received for electricity generated is discussed.

Furthermore, the influence of ongoing changes to the imbalance pricing system in the UK on price received for electricity generated is considered. Finally, an example of how a deterministic or probabilistic forecast might be used to choose the volume of electricity to bid in the electricity market is given.

Chapter 7 gives the overall conclusions of this work and discusses potential further work.

Chapter 2. Background and Literature

This chapter introduces the underlying theoretical concepts upon which this thesis relies. In addition to this, a review of relevant literature is presented which gives the reader an overview of both the historical developments in the field and current state of the art techniques. First, the fundamentals of wind and wind power production are explored. This is followed by a review of forecasting methods and a review of state of the art wind speed and power forecasting techniques. A discussion of grid operations and power markets follows, giving context to the problems caused by wind energy as a power source. Once these issues have been considered, the applications of wind power forecasting are discussed.

2.1 Fundamental concepts

In this section three of the fundamental concepts related to wind forecasting are explored. These are: the atmospheric and physical processes which generate and affect wind, how wind speed increases with height above ground level and how wind is converted to power output. These concepts form an important background for considering how wind speed and power output are predicted. These concepts are included, either implicitly or explicitly, in many forecasting methods [24].

2.1.1 Wind meteorology

The troposphere is the lowest part of the atmosphere, which can be split into the planetary boundary layer (PBL) and the free atmosphere. The PBL is directly influenced by the Earth's surface, reacting to things such as frictional drag, evaporation, transpiration and heat transfer [25]. The PBL can be broken down further into the surface boundary layer, a well-mixed layer and a capping entrainment layer [26]. The PBL depth is variable in space and time. Wind occurs in the PBL due to the movement of air in the troposphere. This movement of air is driven by differences in atmospheric pressure, known as pressure gradients. Convective processes are driven by changes in atmospheric pressure due to uneven solar heating and cooling of the earth [27], a process which can happen on both a

regional and global scale. For example, on a small scale, pressure gradients created by variations in heating due to different surfaces, such as land and sea, lead to the movement of air. On a global scale, areas closer to the equator receive increased solar insolation comparative to the poles, leading to the circulation of air. Variations in heat flux in time occur due to seasonal, inter-annual and diurnal changes in solar heating. Variations in space are affected by height above ground and topography (both local and regional) [27]. Wind patterns are also affected by the Coriolis effect. This stops air simply flowing from areas of high pressure to low pressure causing circulation in global wind patterns. This influences prevailing wind conditions at a location [28].

Another key process which influences wind characteristics is mechanical turbulence. Mechanical turbulence is caused by the interaction between wind and the ground. Mechanical turbulence results from wind flowing over irregular terrain. This can include natural obstacles such as trees or hills or manufactured obstructions such as buildings. The magnitude, vertical and horizontal extent of mechanical turbulence depends on the wind speed, roughness of the terrain and atmospheric stability.

Atmospheric stability is an important component in modelling wind characteristics. Atmospheric stability is a measure of the atmosphere's tendency to encourage or deter vertical motion [29]. Under stable conditions, vertical motion is suppressed and under unstable conditions, vertical motion is encouraged. In the absence of either of these conditions, the atmosphere is said to be neutral. Atmospheric stability is affected by convective processes and wind speed. The convective processes affecting atmospheric stability follow a diurnal pattern. Typically, over land, at night, the ground becomes a heat sink and the lower atmosphere becomes stably stratified. In the morning, surface heating eventually becomes strong enough to cause convective mixing of air leading to unstable conditions before gradually becoming stable again as night-time heat patterns return. The rate at which this process occurs depends on the strength of surface heating and is therefore seasonal. As the atmosphere changes between stable and unstable conditions neutral conditions are seen where thermal processes have less effect on the vertical profile of flow. Neutral conditions are also seen when wind speeds are high. In these conditions, mechanical turbulence overrides convective processes resulting in neutral conditions. Atmospheric stability has a significant influence on atmospheric

circulation and momentum transfer [30]. This is particularly true for offshore wind climates as changes in surface roughness seen at the coastline affect wind profiles differently dependent on stability conditions [30, 31]. It is therefore important to consider atmospheric stability in wind prediction.

2.1.2 Vertical wind profiles

In addition to the processes outlined above, understanding how wind speeds change with height is an important part of making predictions relevant to the wind industry. Surface friction causes wind speeds to be lowest close to the ground and to increase with height. The power law, given by Equation 2-1 [32], is a simple way of estimating horizontal wind speed, u_2, at a height, z_2, above ground level given a reference wind speed, u_1, and height, z_1.

$$u_2 = u_1(z_2/z_1)^{\alpha} \qquad \text{Equation 2-1}$$

In this equation, α is the power law exponent which is determined empirically. The appropriate value depends on surface roughness and atmospheric stability. Von Karman showed that under some conditions $\alpha = 1/7$ is appropriate [33]. This value is used where no surface roughness or atmospheric stability data is available. The power law is frequently used due to its mathematical simplicity [34]. However, it does not account for the impact of surface roughness on vertical wind profiles.

An alternative way of estimating vertical wind profiles is using the logarithmic wind profile, developed from a combination of theoretical relationships and empirical research based on boundary layer fluid dynamics [27]. This is a relationship between three variables, the height above ground level, the roughness length and the frictional velocity given by Equation 2-2 [25].

$$u(z) = \frac{u_*}{k}\ln\left(\frac{z}{z_0}\right) \qquad \text{Equation 2-2}$$

where $u(z)$ is the wind speed at height z above the surface, u_* is the friction velocity, z_0 is the roughness length and k is the Von Karman constant. This is a simple version of the logarithmic wind profile, more complex versions include a displacement height. The roughness length z_0 is a parameterisation of the drag force exerted on the flow due to the roughness of the surface [35] and the frictional

velocity parameterises the frictional force between the flow and the ground [29]. The roughness length is usually estimated based on a general classification of the area in question. Roughness length can be estimated from land surface use as shown in Table 2-1. Classifications are frequently crude (as they are based on satellite imagery of large areas) which can introduce significant errors to calculations.

Table 2-1: Roughness lengths for various terrain types [36]

Surface	Description	z_0 (m)
Sea	Open sea, fetch at least 5 km	0.0002
Smooth	Mud flats, snow, little vegetation, no obstacles	0.005
Open	Flat terrain: grass few isolated obstacles	0.03
Roughly open	Low crops: occasional large obstacles	0.1
Rough	High Crops: scattered obstacles	0.25
Very rough	Orchards, bushes: numerous obstacle	0.5
Closed	Regular large obstacle coverage (suburban area, forest)	1.0
Chaotic	City centre with high and low rise building	>2

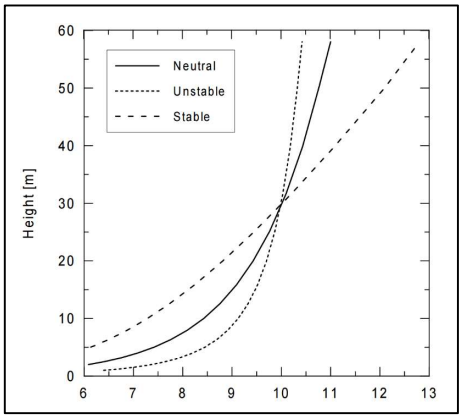
Vertical wind profiles are also affected by atmospheric stability. The wind speed gradient is reduced in unstable conditions and increased in stable conditions. This is because of the increased vertical mixing in unstable conditions and decreased vertical mixing in stable conditions. Neutral conditions are frequently assumed in order to simplify calculations. However, the effects introduced by non-stable conditions can be significant. It is suggested that a slightly more realistic estimation of vertical wind profile may be given by Equation 2-3, the logarithmic profile corrected for stability [29].

$$u(z) = \frac{u_*}{k}\left(\ln\left(\frac{z}{z_0}\right) - \psi\right) \qquad \text{Equation 2-3}$$

where ψ is a stability dependent function which is positive for unstable conditions and negative for stable conditions and all other parameters are as in Equation 2-2. As shown in Figure 2-1, further above ground level (e.g. above 30 m above ground level in this example) under unstable conditions wind speeds increase with height at a slower rate than under neutral conditions, and under unstable conditions the

increase in wind speed with height is faster than under neutral conditions. Closer to the ground (e.g. below 30m above ground level in this example), higher wind speeds are seen under unstable conditions for the same height above ground. Under stable conditions changes in wind direction with height are also observed [29]. The height above ground level at which changes like this will be seen is site dependent and must be observed empirically.

Figure 2-1: Wind profiles under different stability conditions according to Equation 2-2. Source: Peterson et al. [29]



The relationships discussed here can be used either implicitly or explicitly in wind power predictions to estimate hub height wind speed from near surface wind speed predictions or observations. This is useful as near surface wind speed observations and predictions are more frequently available than hub height wind speeds as these are valuable in many industries and not specific to the wind industry. However, there are many atmospheric variables which can affect the vertical wind speed profile and site specific vertical scaling relies on spatial and temporal averaging of wind conditions which can reduce the accuracy of predictions. Because of this, statistical forecasting techniques which aim to directly predict hub height wind speeds may produce more accurate predictions.

2.1.3 Converting wind speed to power

Thus far the meteorological mechanisms which generate wind have been introduced, followed by vertical wind profiles. The method by which this kinetic energy is converted into electrical energy is now discussed. Kinetic energy in an air flow may

be exploited by converting it into electricity using a rotor, which is connected to an electricity generator. This can be done via vertical or horizontal axis turbines, on a scale of a few kilowatt up to multi-megawatt units. The most frequently used turbines are 3 blade horizontal axis turbines. Power is a measure of kinetic energy per unit time. Using fluid mechanics, a derivation of the power '*P*' extracted from a volume of air passing the area of a flow at a given speed is given by Equation 2-4.

$$P = \frac{1}{2}\rho A u^3, \qquad \text{Equation 2-4}$$

where ρ is air density, A is the cross-sectional area of the flow and u is wind speed.

Equation 2-4 gives the maximum power which could be extracted from a flow. However, a wind turbine cannot extract all power from a flow of air. The power coefficient gives the fraction of power extracted by the turbine given the maximum power available. The German physicist Albert Betz concluded that no wind turbine can convert more than 59.3% of the kinetic energy from the wind into mechanical energy in the rotor. This is known as the Betz limit and is the theoretic maximum power coefficient [37]. In reality, this limit is not reached, and the actual limit is unique to a turbine and the operating conditions. Hence the actual extractable power from a wind turbine is given by Equation 2-5.

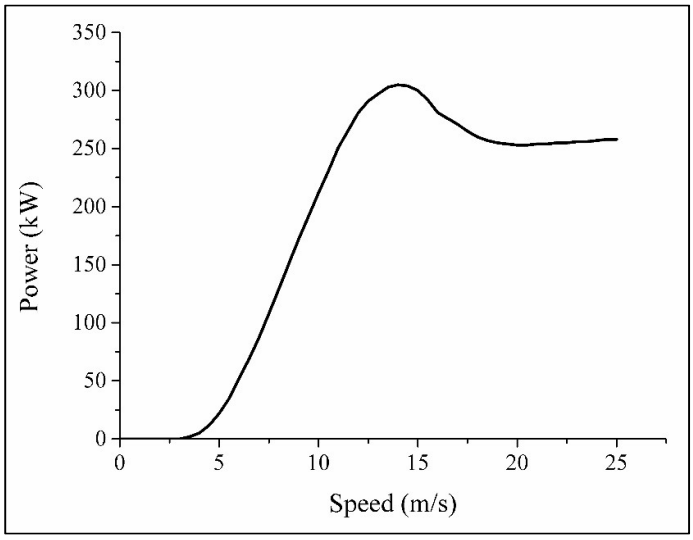
$$P = \frac{1}{2}\rho \pi r^2 u^3 C_{power}, \qquad \text{Equation 2-5}$$

where C_{power} is the power coefficient, r is the rotor radius and all other parameters are as defined in Equation 2-4. From this, it is obvious that the electrical output is related to both the wind speed and the rotor area of the turbine. The power coefficient, C_{power}, used in Equation 2-5 varies between different wind turbines and also varies with wind speed. For example it is affected by the aerodynamic efficiency of the turbine design, turbulence intensity of the wind and condition of turbine blades and other mechanical parts [38]. It is a measure of how efficiently a wind turbine extracts power from a flow.

The expected power output for a turbine at different wind speeds is given by a power curve. Power curves are usually found by taking concurrent measurements of wind speed and power output and plotting average values. They are normally

produced by a wind turbine manufacturer in order for potential customers to compare turbine performance. As an example, a power curve for a 300 kW turbine is given in Figure 2-2. Power curves can be estimated from measurements taken in the field or in a wind tunnel. For a power curve estimated from measurements taken in the field, concurrent wind speed and power output measurements are taken over a period of time and used to estimate the expected power output for a given wind speed. There is a degree of uncertainty in field measurements of wind speed which can cause problems in estimating wind power output at a given wind speed using this method. Anemometer measurements of wind speed can introduce some errors into the observed relationship between wind speed and power output. Additionally, the observed power output is measured as an average for a period of time, which may include some very short term fluctuations in wind speed. This can affect the observed relationship between wind speed and power.

Figure 2-2: Power curve for a Bonus 300 kW wind turbine. Data obtained from The Wind Power [39].



In some cases, power curves are made using measurements taken from a wind tunnel. Using this method, a constant wind speed is artificially generated and the power output measured as this is varied. Whilst this improves the accuracy of measured wind speed, they can be overly optimistic when compared to observations taken from operational wind turbines. This is due to an inability to model very short

term fluctuations in wind speed or other atmospheric conditions seen in operational conditions which can reduce power output.

Power curves published by manufacturers rarely include an estimate of uncertainty in the expected power output. In addition to this, both methods require some averaging of observed power output at a given wind speed. Whilst power curves can be very useful in some situations, the use of a power curve to predict power output from predicted wind speed can introduce significant errors and therefore power prediction methods which avoid this are valuable. For example, some statistical forecasting techniques use measured power output as an input variable in order to establish a relationship between predicted wind speed and power output rather than relying on a power curve. This approach is taken in the model introduced for wind power prediction in Chapter 3 and Chapter 4.

The relationship between wind speed and power output is an important consideration when predicting power output and power curves give some indication of how they are related for a single turbine. However, there are limitations to the use of power curves for predicting power output and uncertainty in the expected power output is important to consider.

The conversion of wind to power and the accuracy of power curves is also affected by atmospheric stability [40-42]. This is largely due to the difference in wind speed and turbulence across the rotor blade and leads to higher output under stable conditions and lower output under unstable conditions [40]. For example, under different stability conditions the rotor-averaged wind speed may differ from the hub height wind speed due to differences in wind shear. The use of power curves to estimate power output at a given wind speed assumes that the increase in wind speed with height will be linear over the area of the wind turbine rotor [42]. However, given the differences noted in Figure 2-1 this is not necessarily the case, particularly for modern turbines with large rotor areas. Wharton and Lundquist [40] found that for an American wind farm the rotor-averaged wind speed was higher than the hub height wind speed under stable conditions which led to an over performance in power output in comparison to what would be expected from a standard power curve. The opposite was observed under unstable conditions. Because of this, the inclusion of stability parameters in wind power predictions can lead to improved

accuracy [40]. This has motivated the discussion in this thesis into whether the inclusion of a stability parameter in the GPR model for wind power prediction can improve wind power predictions.

Overall, it can be seen from some of the issues noted here that the use of standard scaling methods to predict hub height wind speed and power output can introduce significant errors. Therefore the use of other prediction methods which take advantage of statistical techniques and numerical weather prediction are crucial.

2.2 Predicting wind speed and power output

Estimating wind speed or power output is not an easy task due to the unpredictable nature of global weather systems. However, it has many uses and applications. One of the key objectives of this thesis is to develop a model for wind speed and power prediction. Before doing this, currently available prediction techniques and the results that they achieve are discussed. Wind prediction falls into two areas: resource assessment and forecasting. Resource assessment is concerned with establishing the potential wind characteristics experienced by an area. This includes daily, seasonal and annual wind characterises as well as the uncertainty of the wind resource and turbulence in a given area. This is of use when considering wind farm siting and considering the strategic planning of wind energy generation. Secondly, wind prediction can refer to the prediction of wind speed and power at a specific point in time, from minutes to days ahead. This is used to aid electricity supply planning, grid balancing and turbine maintenance amongst other things discussed in Section 2.3. It is this definition of wind prediction that is the focus of this thesis. Wind prediction can be broken down into three areas, numerical weather prediction (NWP), statistical approaches and hybrid models. Models which employ a combination of NWP methods and statistical methods are known as hybrid models. Current techniques for both wind speed and power prediction are discussed in this section. Wind speed and power forecasts can be deterministic, where a single value is predicted, or probabilistic, where the probability of a possible value occurring is predicted. This section begins with a discussion of deterministic forecasting. This starts with an overview of the types of deterministic forecasts available, followed by forecast evaluation techniques and finally a summary of the results presented in the

literature. Subsequently, a discussion of probabilistic forecasting techniques is given. This covers the types of probabilistic forecasts available, how probabilistic forecasts can be used, how probabilistic forecasts are evaluated and a summary of results presented by literature.

2.2.1 Numerical weather prediction

Numerical weather prediction (NWP) is the process of predicting weather conditions through the numerical evaluation of differential equations using current weather observations. A basic summary of the physical approximations made by a numerical weather prediction is given by Peterson et al. [29] and is summarised here.

The state of the atmosphere at any time is defined by a number of parameters including pressure, temperature, density, moisture and three components of velocity. These parameters are governed by the first law of thermodynamics, Newton's second law and continuity equations. From these, fundamental physical laws of the conservation of mass, momentum and energy equations are derived. These form the basis for calculating the atmospheric state at a given point in time. These are outlined in Equations 2-6 to 2-9.

Momentum equations: $$\frac{\mathrm{D}\mathbf{v}}{\mathrm{D}t} + f\mathbf{k} \times \mathbf{v} = -\frac{1}{\rho}\nabla p - g\mathbf{k} + F$$ Equation 2-6

Continuity equation: $$\frac{\mathrm{D}\rho}{\mathrm{D}t} + \rho\nabla \cdot \mathbf{v} = 0$$ Equation 2-7

Thermodynamic equation: $$\frac{\mathrm{D}\theta}{\mathrm{D}t} = 0$$ Equation 2-8

Equation of state: $$p = \rho RT$$ Equation 2-9

where:

- $\mathbf{v}$ Velocity vector (u, v, w) for a flow
- $\mathbf{k}$ Vertical unit vector
- t Time
- f Coriolis parameter $f = 2\Omega\sin\varphi$ (Ω is the rate of the earths rotation, φ is latitude).

p Pressure of the gas in the flow (Pa)

ρ Density of air in the flow (kg m^{-3})

F Force of friction (N)

R Specific gas constant (J kg^{-1} K^{-1})

T Temperature (K)

θ Potential temperature (K), defined as $\theta = T\left(\frac{p}{p_0}\right)^{-R/c_p}$
(c_p is specific heat capacity (J kg^{-1} K^{-1}), p_o is reference pressure (Pa))

The Equations 2-6 – 2-9 allow the development of models which may be solved numerically to establish meteorological conditions at any point in time and space. The mathematical model can be solved through integration in time and space subject to initial boundary conditions. However, the continuous atmosphere must be approximated by discrete grid points, resulting in a loss of accuracy. Furthermore, the initial state of the atmosphere is not known for many locations across the earth due to a lack of observations and potential errors in observations. The spatial and temporal resolution of a model is constrained by the computer power available, as solving the numerical equations becomes more computationally intensive with more grid points. Generally, full global models are run on coarse resolution grids and then local area models are computed for areas of interest using the global model as input. Global models are usually run by large weather forecast providers such as The UK Meteorological Office (the Met Office). The Met Office unified model is an operational NWP model run in a number of configurations to provide short and long range weather forecasts on both a global and regional scale. The global model supports a higher resolution regional model which is able to capture more detail in atmospheric processes. The model has been updated to incorporate a variable resolution UK model (UKV). This resolves atmospheric conditions over two scales, a high-resolution inner domain (1.5 km) and a coarser resolution (4 km) near grid boundaries, which limits the effects of the boundaries on the forecasts [43]. The details of the global model and UKV including resolution, frequency which the model is run, forecast length and initial conditions are given in Table 2-2. Various other NWP models have been developed including Global Forecasting System, MM5, Prediktor and HIRLAM [44].

Weather observations are used to gain an accurate representation of the initial state of the atmosphere and inform the NWP. In a process known as data assimilation, recent observations are combined with previous weather forecasts to estimate the current atmospheric conditions [45]. The process of data assimilation allows for the model and observational uncertainty to be accounted for. Globally the vast majority of data is obtained from satellites [46]. However, radiosonde, surface land observations, marine observations, aircraft measurement and radar data are also used. The UK has an extensive network of synoptic observation stations, as shown in Figure 2-3, which give detailed observations which are used within the high-resolution model. At these stations, various observations are made such as wind speed and direction, rainfall and temperature. Measurements are taken at 10 m above ground level.

Table 2-2: Comparison of Met office global and UKV models [43]

Model	Grid length in mid-latitudes	Grid points	Vertical levels	Maximum height	Forecast length	Run times (UTC)
Global	17 km	1536 x 1152	70	80 km	6 days	00, 06, 12, 18
UKV	1.5 km inner 4 km outer	744 x 928	70	40 km	36 hours	03, 09, 15, 21

Previously, statistical interpolation and 3D variational (3DVAR) algorithms were the most frequently used data assimilation techniques for weather forecasting [47]. More recently 4D variational data assimilation has been incorporated into the Met office operational models, which allows the development of weather systems over time to be better modelled. This uses covariance matrices incorporating recent observations and predicted conditions allowing for relationships in time and space to be modelled [48].

Figure 2-3: UK synoptic observation network. Source: The Met Office [49]

Traditionally NWP models were deterministic weather predictions. However, innovation in computer capacity has allowed multiple models to be run at any one time, enabling forecasters to generate a representative sample of future meteorological scenarios. This process is known as ensemble forecasting and involves running multiple models with different initial conditions which allow the uncertainty to be assessed. Ensemble forecasting has many applications from ecology [50] to economics [51]. The use of ensemble forecasting for weather prediction has been subject to a number of studies. Eckel and Mass [52] reviewed the potential of using a short-range ensemble forecasting system to produce useful

information about the forecast probability of a mesoscale model for up to 2 days in advance, commenting that whilst ensemble forecasts were advanced for a range of up to 10 days, shorter timescales had received limited attention. Toth et al. [53] considered the use of an ensemble forecast to give a relative measure of uncertainty on forecasts and suggest that the economic advantage of using forecasts can be enhanced by incorporating a measure of uncertainty. Zhu et al. [54] compare the economic benefit of using an ensemble forecast over using a higher resolution mesoscale model, concluding that the probabilistic model is of higher economic benefit than a point forecast even if the point forecast is from a higher resolution model.

NWP models provide good forecasts for wind speed prediction particularly allowing medium to long term forecasts to be made. However, some of their key features limit the model accuracy. Firstly the grid length of an NWP model is limited by computer power. Higher resolution models require huge computational capacity. As computers improve and costs decrease the resolution upon which these models can be run has increased. For example, the Met Office UKV model with a 1.5 km grid lengths replaced the previous UK4 model with 4 km grid lengths in 2009 [55]. However, these higher resolution models can still only be run for shorter term forecasts (up to 36 hours for UKV) and may still not capture some terrain characteristics. For example, a hill, outcrop of rock or body of water contained in a 1.5 km area. Because of this, some adjustments might need to be made to NWP forecasts of wind speed to improve forecast accuracy. This can be beneficial for both shorter term higher resolution models and longer term lower resolution models. It is because of this that throughout this thesis a hybrid NWP and statistical model (GPR) is developed so that corrections can be made to NWP predictions where necessary.

2.2.2 Statistical approaches

NWP models are a good way of predicting weather conditions over a large area and can be effective for long forecast periods [24]. Despite this they can be expensive to run and, due to coarse resolution, may not be able to account for small-scale effects of terrain.

Statistical models make predictions based on analysis of past trends and patterns. Statistical models attempt to find relationships between explanatory variables and output variable employing a range of techniques. Explanatory variables can be NWP results, power output data, observed weather data, local area topographical data and more. Using a statistical approach, the wind speed and direction can be predicted and wind power estimated from this, or the wind power can be directly predicted.

Time series methods for wind forecasting largely focus on the autoregressive nature of wind speed time series. For very short-term wind forecasting, wind speeds are expected to follow relationships seen in the past few hours, and extrapolation is used to predict wind speed or power in the coming hours. The simplest type of time series model is the persistence model. The persistence model forecasts wind speed or power by assuming the forecasted value is equal to that of the prior time period. It relies upon the autocorrelation seen in wind speed and power time series and is commonly used as a benchmark for model performance to allow a comparison between forecasts when different datasets are used. Whilst simple, this method is only effective for time periods up to a few hours. A number of alternative methods have been explored, including ARMA (autoregressive moving average), ARIMA (autoregressive integrated moving average), fractional ARIMA (f-ARIMA), linear predictors or exponential smoothing. Torres et al. [56] outline a seasonally adjusted ARMA model to predict wind speeds up to 10 hours in advance. In this study, the time series was transformed to account for the non-Gaussian nature of wind speed evolution. The model was applied to data from 5 different locations in Spain over a period of 9 years and reports better results than a persistence model, but notes the model is only valid for a short time period. Kavasseri and Seetharaman [57] present an f-ARIMA model to forecast wind speeds up to 48 hours ahead. The results are combined with wind turbine power curves to produce wind power forecasts. This time the model is applied to 4 potential wind farm sites in North America and concludes the forecast is again better than a persistence model. However, the use of power curves in this model can propagate some errors due to the problems with power curves noted in Section 2.1.3. Time series approaches may perform well for short forecast horizons (minutes – a few hours ahead), but their performance is reduced for forecasts further in advance. Therefore, more complex methods have been investigated.

Artificial Neural Networks (ANN) are a family of statistical methods named as such due to their learning processes resembling that of a neural network in the brain. They are interconnected, independent computational units know as neurons. They are 'trained' with respect to certain data sets until they recognise patterns in the data presented to them. They can be used for the purpose of wind speed and power prediction and could be particularly useful as they can deal with noisy and incomplete data, which is often a feature of historical wind data [58]. A number of different ANN techniques exist, including feed forward, recurrent, multilayer perceptron (MLP), radial basis function (RBF) and adaptive linear element network (ADALINE). The use of artificial neural networks has been explored by a number of authors. These include More and Deo [59], Li and Shi [60], Li et al. [61], Ramirez-Rosado et al. [62] and Cadenas and Rivera [63].

More and Deo [59] compare the use of a feed-forward ANN and a recurrent ANN. The model inputs were daily average wind speeds for a coastal area in India over 12 years. The network was used to calculate monthly, weekly and daily average wind speeds. The study concluded that neither the feed forward nor the recurrent network produced consistently better results. In another study by Li and Shi [60] RBF, ADALINE and BP networks are compared. The results showed that the minimum errors depended on the individual site and the error metric used, with no formulation outperforming all others. This led to an approach introduced by Li et al. [61] which combines the use of a Bayesian combination algorithm and three neural network models. The neural network models used in this study are the ADALINE, BP and RBF networks. The method is used to predict the hourly wind speed for 120 hours at two sites in North Dakota. The author suggests some benefits of neural network forecasting including their ability to learn from past data, recognise hidden patterns and model linear and non-linear relationships. However, the study suggests that the results of different neural networks have been variable and no model reliably performs better than all others. The paper concludes that the individual neural networks are not consistent in predicting the one-hour-ahead wind speed under different evaluation criteria, but the Bayesian combination method provides an adaptive and comparatively reliable forecast. Ramirez-Rosado et al. [62] present a multi-layer perceptron neural network to predict power output for individual turbines in a wind farm, in order to predict total wind power output. This was done

for half hour periods up to 72 hours ahead. The models used wind power data from a SCADA system and an NWP model to train the model. Supervisory control and data acquisition (SCADA) systems are control systems which manage data collection from wind farms. The use of power data collected in real time from a SCADA system in a prediction model avoids the need to use a power curve to estimate power output given a wind speed and the issues associated with this. The use of power output data in a wind power prediction model is vital for estimating the overall wind farm output given the expected meteorological conditions. Cadenas and Rivera [63] present the use of ANN to predict an hourly time-series for a single month for a site in Mexico. A number of different network configurations are tested, which the study states show the technique is able to accurately predict short-term wind speed. However, the use of only one site and one month make the results difficult to evaluate in comparison to other techniques.

A number of other methods have been used to predict wind and power output, including wavelet transform, support vector machines and fuzzy logic methods. Wavelet transform methods have been used in a number of studies to process time series data prior to model input. In a study on wind power forecasting in Portugal by Catalão et al. [64], wavelet transform is used to decompose the wind power series into a set of filtered constitutive series. The processed time series is subsequently used to forecast the future power output through the use of a neural network. In this study wind power output for 3 hours ahead is predicted using the previous 12 hours of power output data with a time step of 15 minutes. Damousis et al. [65] present a method for predicting wind speed and power from 30 minutes to 2 hours ahead using a fuzzy model trained using wind speed and direction data from a neighbouring site. The fuzzy logic inference system provides estimates of the future wind speed based on variations at a neighbouring site. The study suggests that over flat terrain and where the reference and neighbouring site show sufficient correlation in wind speed and power output, the results for short-term forecasting were better than persistence.

The range of methods introduced here gives an overview into the many statistical methods which are available to predict speed and power output. In general these methods are only effective on short time scales, therefore to make longer term predictions must be combined with NWP models, as discussed in the next section.

Whilst numerous statistical methods exist for the prediction of wind speed and power there is still improvements to be made and new methods are continually sought to increases predictive accuracy.

2.2.3 Hybrid forecasting techniques

Different prediction techniques have different strengths and weaknesses. For example, statistical techniques generally perform better for short term predictions whilst NWP methods give better long term predictions. Because of this, different techniques are frequently combined to give hybrid forecasting techniques. For example, Liu et al. [66] proposed a hybrid approach using wavelet transform to analyse the original time series. In addition to this, an improved time series method (ITSM) was used to predict future value in the time series. The wavelet transform was used to decompose the time series and then a three-step approach was taken to implement the ITSM. Firstly, the Box-Jenkins methodology was used to establish the autoregressive integrated moving average (ARIMA) models which best fit the time series behaviour. Using the equation proposed by the ARIMA model the next value is predicted. The model parameters are then re-estimated using this predicted value. This new equation is used to calculate the step-ahead forecast. This method was applied to 200 hours of time series data, with errors considered for between 3 and 10 hours ahead. The study compares the method with a simple backpropagation neural network approach and concludes that the proposed method performs better according to a number of error measures such as mean absolute error (MAE), root mean squared error (RMSE) and mean absolute percentage error (MAPE).

Larson and Westrick [67] use hybrid NWP and statistical methods to predict hourly average wind speeds. The authors investigate the use of off-site observations combined with a numerical prediction model to increase forecast accuracy over short timescales. The study uses on and off-site observations and NWP forecasts as predictor variables and applies a number of statistical algorithms, including neural networks, support vector machines and conditional neural networks. This report only considered one site in the USA but reported that the use of on or off-site data improved the forecast accuracy of a 2 hour forecast.

Throughout this thesis, a hybrid NWP and GPR model is developed for wind speed and power forecasting. This allows the long term predictive ability of an NWP model to be combined with the skill of a statistical model for shorter term predictions.

2.2.4 Deterministic forecast evaluation techniques

In order to make a comparison of the numerous forecasting methods in use and to establish how well a forecast is performing, some numerical methods for measuring accuracy of a forecast need to be defined. This allows the objective comparison of different methods and consideration of the strengths and weaknesses a forecast exhibits. Many different ways of evaluating accuracy can be used. Jolliffe and Stephenson [68] give a full discussion of forecast verification. In general, a combination of 4 error metrics is used to compare forecasts. Throughout this thesis, a combination of these metrics will be used to evaluate and compare forecasts. Each metric provides different information on the performance of a forecast. The error metrics most commonly used to measure model performance are mean bias error (MBE), mean absolute error (MAE), mean absolute percentage error (MAPE) and root mean squared error (RMSE). These are calculated using Equation 2-10 to Equation 2-13. MBE is a measure of the overall over performance or under performance of a forecast. A large value for MBE can indicate a systematic tendency to deviate from the observed value which may need to be corrected. A small value for MBE may simply be considered as noise. Positive and negative error values cancel out in the bias calculation, therefore, other error metrics are also required. MAE is used to give an average magnitude of all deviations, giving a general value to the model performance. MAPE presents the average magnitude of errors as a percentage of the observed value, which allows some comparison between different datasets to be achieved. However, MAPE can be very sensitive to small deviations. RMSE is another metric which shows the average magnitude of errors. In comparison to MAE, RMSE gives a higher weighting to large error values. MAE gives equal value to all errors whilst larger error values affect the RMSE more heavily. This means that the RMSE can be more sensitive to outliers. A number of authors have compared the relative merit of using MAE and RMSE. Wilmott and Matsuura [69] conclude that MAE is favourable over RMSE as RMSE is a function

of both the magnitude of errors and the distributions of errors. They also state that RMSE is ambiguous and hence should not be used to compare model performance. On the other hand, Chai and Draxler [70] refute the claims made by Willmott and Matsuura [69], concluding that there are some circumstances where using RMSE can be beneficial and in general using a variety of error metrics (not limited to RMSE and MAE) should be used to evaluate and compare models. In Equations 2-10 to 2-13 y_t = observed wind speed or power and $\widehat{y_t}$ = predicted wind speed or power.

$$\mathrm{MBE} = \frac{1}{n}\sum_{t=1}^{n}(\widehat{y_t} - y_t) \qquad \text{Equation 2-10}$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{t=1}^{n}|y_t - \widehat{y_t}| \qquad \text{Equation 2-11}$$

$$\mathrm{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\frac{|y_t - \widehat{y_t}|}{y_t} \times 100 \qquad \text{Equation 2-12}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \widehat{y_t})^2} \qquad \text{Equation 2-13}$$

In the case of wind power forecasts frequently it is desirable to compare model performance at different sites. In this case, it is necessary to normalise error metrics to allow the comparison of results at different sized wind farms. Error metrics can be expressed as a percentage of the total wind farm capacity. These are referred to as normalised error metrics. Calculation of normalised MBE, MAE and RMSE is achieved using Equations 2-14 to 2-16. In these equations C refers to the installed capacity of a wind power producer under evaluation. This could be an individual turbine, a full wind farm or part of a wind farm. In Equations 2-14 to 2-16 y_t = observed wind power and $\widehat{y_t}$ = predicted wind power.

$$\mathrm{NMBE} = \frac{1}{n}\sum_{t=1}^{n}(y_t - \widehat{y_t}) \times \frac{100}{C} \qquad \text{Equation 2-14}$$

$$\mathrm{NMAE} = \frac{1}{n}\sum_{t=1}^{n}|y_t - \hat{y}_t| \times \frac{100}{C}$$ Equation 2-15

$$\mathrm{NRMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2} \times \frac{100}{C}$$ Equation 2-16

In order to define the quality of a forecast, it is common practice to compare performance measures against a baseline forecast. This baseline could be a number of things. A commonly used baseline is 'persistence'. This is a forecast which assumes the future value of the forecast variable will be the same as the currently observed value. In short term wind speed and power forecasts this can be quite successful and a skilful forecast must be able to perform better than this. An alternative baseline could be another forecast which is known to be relatively successful. For example, an ARIMA model may be compared against an ANN for the same site to establish which gives the better performance.

In the evaluation of model performance presented later in this thesis, a combination of error metrics is used to provide an overview of model performance. Furthermore, results are compared to either a persistence model or other forecasting techniques depending on the most appropriate baseline.

2.2.5 Summary and comparison of deterministic forecasting methods

Sections 2.2.1, 2.2.2 and 2.2.3 cover some of the currently available deterministic forecasting methods. In this section, the results demonstrated by different authors for different prediction techniques are presented. The error metrics used to present results vary across the literature, but an attempt is made to consolidate the results in a comparable way. In Table 2-3 an overview of some of the currently available deterministic forecasting techniques is given and the key points form this are discussed here. For wind speed forecasting methods RMSE, MAE and MAPE are shown where possible. For wind power forecasting RMSE and MAE are normalised by capacity to allow comparison between installations of different sizes, giving NRMSE, NMAE and MAPE as comparison statistics for power forecasts. MBE and

NMBE is discussed too infrequently in literature to be included in comparisons. Where sufficient detail is given in the literature, the results are shown in comparison to the persistence model.

The wind speed prediction methods reviewed in Table 2-3 comprise 4 statistical methods and 2 hybrid methods and cover a variety of timescales, from 1 hour ahead to 120 hours ahead. Firstly comparing short term predictions, Chen et al. [71], Li and Shi [60], Li et al. [61] and Torres et al. [56] present results for prediction up to 4 hours in advance. The method presented by Chen et al. [11] is a hybrid method, whilst the other are statistical methods. For these methods, MAE and RMSE do not differ significantly. MAE varies between 0.7 and 1.14 ms^{-1} for a forecast 1 hour ahead and RMSE varies between 0.96 and 1.5 ms^{-1} for a forecast 1 hour ahead. MAPE has the largest variation for forecasts 1 hour ahead with the largest (23%) reported by Li and Shi [60] and the smallest (11%) reported by Chen et al. [71]. The method presented by Li et al. [61] has the highest MAE and RMSE at 1 hour ahead, with an MAE of 1.14 ms^{-1} and an RMSE of 1.5 ms^{-1}. The hybrid model presented by Chen et al. [71] shows very similar results to the statistical models presented by other authors at this timescale. The use of statistical and hybrid models for very short term models is popular due to their ability to exploit the fact that wind conditions tend to persist for periods of time. In addition to this, they are frequently less computationally intensive than NWP models to generate and so are less expensive. However, the methods presented here do not incorporate an NWP model in any way therefore may fail to predict more extreme or changeable weather conditions.

Kavasseri and Seetharaman [57], Torres et al. [56], Chen et al. [71] and Louka et al. [72] show results for forecasting methods further in advance (5-120 hours ahead). The methods discussed by Kavasseri and Seetharaman [57] and Torres et al. [56] are statistical, whilst Chen et al. [71] and Louka et al. [72] present hybrid methods. The results presented by Kavasseri and Seetharaman [57] are difficult to compare to the other methods due to how the results are aggregated. Of the other three methods, Torres et al. [7] only show results for forecasts up to 10 hours ahead whilst Louka et al. [72] presents results up to 120 hours ahead and Chen et al. [71] up to 72 hours ahead. The hybrid methods presented by Louka et al. [72] and Chen et al. [71] perform much better than the statistical method presented by Torres et al. The hybrid

method by Louka et al. has an MAE of 2.04 ms^{-1} and an RMSE of 2.88 ms^{-1} as far ahead as 120 hours in advance, whilst the statistical method has an MAE of 2.5 ms^{-1} and an RMSE of 3 ms^{-1} at just 10 hours ahead. MAPE is only given by Chen et al. [73] hence this is not comparable. This indicates, as other literature suggests, that hybrid methods can perform well in the short term and frequently outperform statistical methods further in advance. This is due to the reliance of statistical methods on the persistence of past wind conditions to predict future condition whilst hybrid methods usually incorporate NWP models which are able to better model weather systems further in advance. There is still room for improvements in predictive accuracy in current state of the art hybrid models. Because of this, the investigation of new hybrid models offers an interesting area of research in wind speed and power prediction models.

Five wind power prediction methods are compared in Table 2-3, including 3 hybrid methods and 2 statistical methods. The first statistical method, presented by Catalão et al. [64], is difficult to compare to the other techniques due to the errors shown. The only other statistical method is that presented by Ramirez-Rosado et al. [62] which reports results for forecasts up to 72 hours in advance. However, RMSE averaged over 3 time periods is shown (12-24 hrs ahead, 24-48 hrs ahead and 48-72 hrs ahead), again making it difficult to compare to other methods. For the remaining three hybrid methods only NMAE and NRMSE are compared as MAPE is only given in one case. The hybrid methods presented by Chen et al. [73] and Shu et al. [74] report very similar results. The model presented by Chen et al. [73] reports results for forecasts from 1 – 24 hours ahead, with an NMAE of between 7.5 and 11.1% and an NRMSE of between 11 and 16%. The model given by Shu et al. [74] gives results for forecasts from 1 – 48 hours ahead, with an NMAE of between 7 and 15% and an NRMSE of between 11 and 21%. The model presented by Louka et al. [72] shows results for forecasts between 24 and 120 hours ahead. This model seems to outperform others with an NMAE of between 11 and 15.5% and an NRMSE of between 15 and 21%.

The methods presented in Table 2-3 give an idea of the results which are possible with currently available forecasting techniques. It can be seen from these methods that for good wind speed forecasts, MAE should be around 1 ms^{-1} for forecasts up to 3 hours ahead and 2 ms^{-1} for forecasts further in advance. In addition to this, for

good wind power forecasts, NMAE should be below 11% for forecasts up to 3 hours ahead and below 16% for forecasts up to 24 hours ahead. The forecast results are highly dependent on the data sets used, some sites may have complex terrain which is more difficult to predict whilst others are simpler. However, this gives some point of reference for a new forecast developed. The results shown in this table suggest that statistical methods are appropriate for forecasting up to a few hours in advance and hybrid models which incorporate an NWP model enable accurate predictions further in advance. This indicates that a flexible model which is able to rely more heavily on prior observations for short term predictions and NWP outputs for predictions further in advance will enable good predictions from 1 hour up to days in advance, This thesis focuses on the development of a forecast for wind speed and power output from a few hours up to 3 days in advance. Because of this a hybrid NWP and statistical model is used.

Table 2-3: Overview of currently available deterministic forecasting techniques

Authors	Year	Speed or power	Method type	Method summary	Data	Forecast period	Results
Catalão, et al. [64]	2011	Power	Statistical	ANN + wavelet transform	All wind farms in Portugal that connect with the national electric grid.	3 hours ahead.	MAE not given RMSE: 392.3 MW (Persistence not given) MAPE: 7% (Persistence 19%). Total capacity forecasted not given so cannot compare NRMSE.
Chen, et al. [71]	2013	Speed	Hybrid	Wavelet and Gaussian process	Wind farm in southern China. 15 turbines, installed capacity 2000kW	1-4 hours ahead, and 1-3 days ahead	MAE: 0.72-1.6 ms^{-1} (Persistence: 0.74 – 1.83ms^{-1}) RMSE: 0.96 -2.04 ms^{-1} (Persistence: 1.0 – 2.23 ms^{-1}) MAPE: 11.24 -44% (Persistence: 11.1 – 42%)
Chen, et al. [73]	2014	Power	Hybrid	Gaussian process and NWP	3 wind farms in China. 3 years for 2 wind farms and 2.5 months for one.	1-24 hours. Results not shown separately	Results given for 4 wind farms: NMAE: 7.5-11.1% (Persistence 9.8 – 18.6%) NRMSE: 11.69 – 15.96% (Persistence 15.7 – 26.3%) MAPE: 7.6 – 11.12% (Persistence 10.1 – 18.4%) Best results for the largest wind farm.
Kavasseri and Seetharaman [57]	2009	Speed	Statistical	f-ARIMA	Wind speed from 4 potential wind farm sites in North Dakota	24 and 48 hours ahead	MAE not given RMSE: 5.35% (Persistence: 8.43%) MAPE (24hrs): 47% (Persistence 117%).
Li and Shi [60]	2010	Speed	Statistical	Comparison of 3 ANN methods	1 year of hourly mean wind speed from two observation sites in North Dakota	1 hour ahead	MAE: 0.9-1.05 ms^{-1} RMSE:1.2-1.4 ms^{-1} MAPE:19%-23% No single ANN outperforms the other on all error metrics (MAE, RMSE, and MAPE). No comparison to persistence given.
Li, et al. [61]	2011	Speed	Statistical	Bayesian adaptive combination from ANN models	1 year of hourly wind speed at 10m above ground	1 hour ahead	MAE: 1.137 ms^{-1} RMSE: 1.508 ms^{-1} MAPE: 18% No comparison to persistence given

Louka, et al. [72]	2008	Speed and Power	Hybrid	Kalman filtering to post process NWP	1 year wind speed and power data at Rokas wind farm	24, 48, 72, 96 and 120 hours ahead	Speed: MAE: 1.75 - 2.04 ms^{-1} RMSE: 2.38 – 2.88 ms^{-1} Power: NMAE: 11 - 15.5%; NRMSE: 15 - 20.5% MAPE not given, comparison to persistence not given.
Ramirez-Rosado, et al. [62]	2009	Power	Statistical	Two ANN methods (FORECAS and SGP)	Wind farm with rated power of 21,600kW, 12 turbines of 1.8MW.	0.5 - 72 hours, time step 0.5 hours.	Average RMSE given for 3 time periods: 12-24h, 24-48h and 48-72h. FORECAS: 14-19.7%, SGP: 14-18.8%, Persistence: 31.2- 37.5%. MAE and MAPE not given
Shu, et al. [74]	2009	Power	Hybrid	Two-stage hybrid network with Bayesian clustering and SVR.	74 MW wind farm in Oklahoma, US.	1-48 hours ahead	Errors are given for 1, 24 and48hrs. NMAE: 7-15% (Persistence 8-25%), NRMSE: 11-21% (Persistence 11-34%) MAPE not given
Torres, et al. [56]	2005	Speed	Statistical	ARMA	5 locations, 9 years. Wind measured every 10 mins at 10m and averaged over 1 hour.	up to 10 hours ahead	MAE: 0.9 - 2.5 ms^{-1} (Persistence 0.9 - 2.9 ms^{-1}) RMSE: 1.2 - 3 ms^{-1} (Persistence 1.25 - 3.7 ms^{-1}). MAPE not given

2.2.6 Probabilistic forecasting

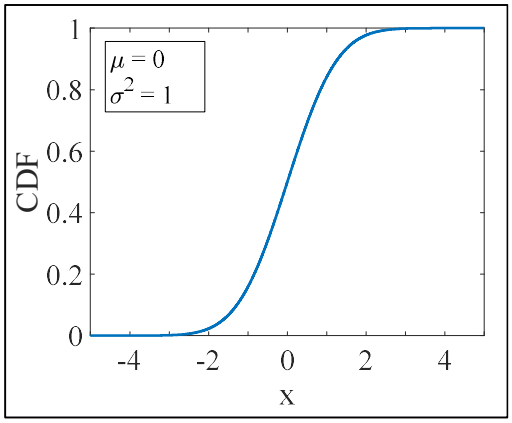
The forecasting methods presented so far provide point predictions of expected wind speed or power at a given point in time. In contrast, probabilistic forecasts consider the probability of a range of possible values occurring by treating the forecast as a random variable. This gives more information regarding the confidence in a forecasted value and a range of probable values. This information can be incorporated into the scheduling of power generating units, managing financial risk whilst bidding in electricity markets or managing reserve levels to minimise risk. The uses and applications of probabilistic wind power forecasting are discussed in more detail in Section 2.3. All forecasts contain a certain level of uncertainty, and in probabilistic forecasting this uncertainty is conveyed by expressing the probability of an event occurring. Probabilistic forecasts provide additional information which can be useful to a forecast user and add value in the wind industry. However, they can be difficult to evaluate and implement in an operational setting. In this section, the main types of probabilistic forecast available will be introduced and a forecast evaluation framework will be discussed. Finally, a summary and comparison of state of the art wind power forecasting methods will be given.

An introduction to the types of probabilistic forecast is given by Pinson et al. [75], a summary of which is given here. If a forecast is made at time t for k hours ahead, then for a probabilistic forecast it is assumed that the actual output at time $t+k$ is an unknown random variable, Y_{t+k}. The observed power output at time $t+k$ is y_{t+k}which is a realisation of the random variable Y_{t+k}. For Y_{t+k}, a probability density function (pdf) is given by $f_{t+k}$and a cumulative density function (CDF) is given by F_{t+k} . The CDF, given by Equation 2-17, is a function which defines the probability of different values of Y_{t+k} occurring.

$$F_{t+k}(x) = P(Y_{t+k} \leq x) \qquad \text{Equation 2-17}$$

where the values x can take depend on the distribution of Y_{t+k}. For example for a random variable $X \in [-\infty, \infty]$ which is normally distributed with parameters $\mu = 0$ and $\sigma = 1$, the CDF, $F_{t+k}(x)$, evaluated at $x \in [-4,4]$ is shown in Figure 2-4.

Figure 2-4: Cumulative distribution function for a random variable with a standard normal distribution



The aim of probabilistic forecasting is to predict the distribution of Y_{t+k}. Probabilistic forecasts can loosely be categorised as quantile forecasts, interval forecasts or density forecasts. The different types of forecast can give subtly different information, however are often linked. Quantiles and intervals can be obtained from a density forecast, or they can be forecast independently. Alternatively, a density forecast can be formulated from quantile or interval forecasts, or it can be predicted separately.

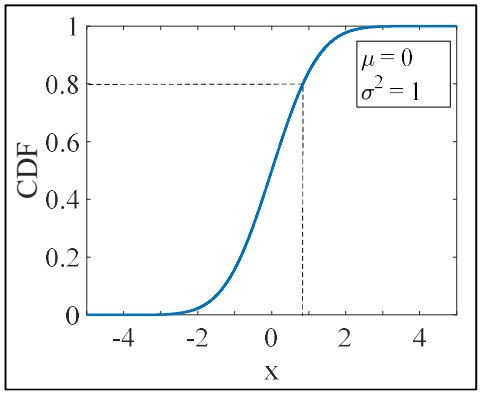
A quantile forecast is a point in the predictive distribution which corresponds with a certain probability of an event occurring. Assume that the predictive CDF of Y_{t+k} is given by $\hat{F}_{t+k}(x)$. Provided the CDF, $\hat{F}_{t+k}$, of a random variable Y_{t+k} is a strictly increasing function, a quantile with proportion $\alpha \in [0,1]$ is defined by Equation 2-18. A quantile forecast, $\hat{q}_{t+k}^{(\alpha)}$ produced at time t for time $t+k$ aims to estimate a quantile based on the predictive information available.

$$\hat{q}_{t+k}^{(\alpha)} = \hat{F}_{t+k}^{-1}(\alpha) = inf\{x \in \mathbb{R} : \hat{F}_{t+k}(x) \geq \alpha\} \qquad \text{Equation 2-18}$$

where *inf* is the infimum of the set, the largest value which satisfies the conditions. Considering the standard normal distribution as shown in Figure 2-5, a prediction may be desired where there is a high probability that the observed value will be less than the predicted value. In this case, $\alpha = 0.8$ may be chosen. Hence a prediction of

$\hat{q}_t^{(0.8)} = \hat{F}_t^{-1}(0.8) = 0.84$ would be made. The quantiles corresponding to different values of α can either be obtained by using a model which predicts the full CDF, $\hat{F}_{t+k}(x)$ or by using a model which predicts the values of $\hat{q}_{t+k}^{(\alpha)}$ for values of α which are of interest.

Figure 2-5: Quantile forecast with $\alpha = 0.8$ for a standard normal distribution.



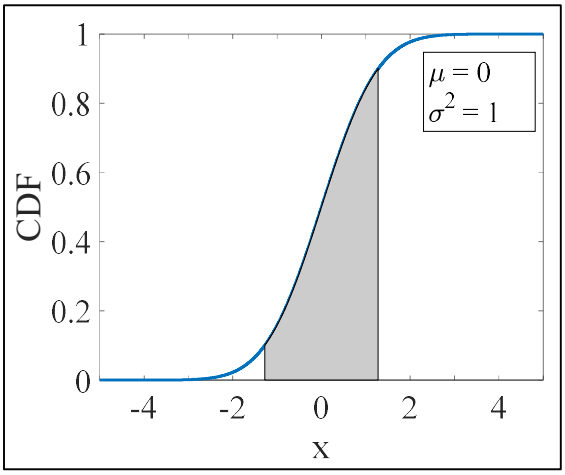
Alternatively, an interval forecast gives a range of values in which the expected value is expected to lie with a certain probability. The interval lower and upper bounds can be defined in terms of quantiles. The probability of the actual value falling with the range of values predicted is called the coverage rate and is given by $(1-\beta), \beta \in [0,1]$. An interval forecast for β is given by Equation 2-19.

$$\hat{I}_{t+k|t}^{\beta} = [\hat{q}_{t+k|t}^{(\alpha_l)}, \hat{q}_{t+k|t}^{(\alpha_u)}] \qquad \text{Equation 2-19}$$

where $\alpha_u - \alpha_l = 1 - \beta$

A prediction interval is not uniquely defined by a coverage rate. The way the interval is centred on the CDF also needs to be defined. Most commonly the interval is centred on the median. Returning to the standard normal distribution, if $\beta = 0.2$ and the prediction interval is centred on the median then $\alpha_u = 0.9,\ \alpha_l = 0.1$ and $\hat{I}_{t+k|t}^{0.2} = \left[\hat{q}_{t+k|t}^{(0.1)}, \hat{q}_{t+k|t}^{(0.9)}\right] = [-1.28, 1.28]$. This is demonstrated visually in Figure 2-6.

Figure 2-6: Interval forecast with $\beta = 0.2$ for a standard normal distribution.



As with quantile forecasts, interval forecasts can be obtained from a model where the full predictive CDF is provided, as in the example above, or intervals can be directly predicted from a model using parameterisations.

The final type of probabilistic forecast considered here is a density forecast. In this type of forecast, the entire predictive distribution is forecast. The predictive distribution can be modelled through the use of parametric and non-parametric models by assuming the predictive distribution type and learning suitable parameters. Alternatively, a discrete CDF can be obtained by predicting multiple quantiles. Using a model to predict a set of m quantiles, the predictive CDF is given by Equation 2-20.

$$\hat{F}_{t+k|t} = \left\{ \hat{q}_{t+k|t}^{(\alpha_i)}, i = 1, \dots, m \,\middle|\, 0 \leq \alpha_1 < \cdots < \alpha_m \leq 1 \right\} \qquad \text{Equation 2-20}$$

The definitions and example given above show how a density forecast can be obtained from quantile or interval forecasts, or vice versa depending on the model. In some applications forecasting specific quantiles or intervals is all that is required, in which case this can be simpler than producing a full predictive distribution. In other applications, the full predictive distribution is of interest. However, looking at specific quantiles can still offer useful insights.

2.2.7 Forecast evaluation for probabilistic forecasts

In Section 2.5.1.4 forecast evaluation techniques were discussed for deterministic forecasts. In order to quantify the quality of a deterministic forecast the deviation between the predicted value and observed value is considered. However, as noted by Pinson et al. [75] and Gneiting et al. [76], it is not quite so simple to evaluate the quality of a probabilistic forecast. For example, just because the observed value fell outside the range of values predicted with a certain probability does not mean that the predicted interval was wrong. Therefore an alternative framework for evaluating the success of a probabilistic model is required. Frameworks for evaluating the quality of a probabilistic forecast are discussed by Pinson et al. [75] and Gneiting et al. [76]. Gneiting et al. [76] propose evaluating forecasts according to 3 main features: calibration, sharpness and resolution. Furthermore, Gneiting et al. [77] suggest the use of skill scores to evaluate the relative contribution of calibration, sharpness and resolution in different models and to compare between models.

Calibration is defined as the statistical consistency between the predictive distribution and the observed values [76]. This is also referred to as reliability. Whilst reliability is important, it does not alone define a good probabilistic forecast. Take, for example, if the long term frequency distribution of a meteorological variable is used to make a probabilistic forecast. In the case of a wind speed forecast, this could be the long term observed frequency of wind speeds at a site. This would be a reliable forecast, yet is not the most valuable forecast [68]. This is because over a significant number of forecasts the predictive distribution would be consistent with the observed distribution, however, it would offer the user limited insight into short term predicted conditions. In addition to being reliable, the predictive distributions or quantiles should be focused and able to differentiate between different predictive conditions in order to be useful. The first property is known as sharpness and the second as resolution. Sharpness quantifies a forecasts ability to minimise uncertainty surrounding a prediction. Resolution quantifies a forecast ability to deviate from a simple climatological predictive distribution [75]. Gneiting et al. [76] consider a good probabilistic forecast to be one which produces

narrow prediction intervals (maximised sharpness) subject to calibration. It is worth noting at this point that a calibrated probabilistic forecast with zero spread is equivalent to a perfect deterministic forecast.

Assume a probabilistic prediction of power output made at time *t* for time *t+k* is given by a CDF $\hat{F}_{t+k}(x)$. The actual output at time *t+k* is assumed to be an unknown random variable Y_{t+k}. The observed power output at time *t+k*, y_{t+k} is a realisation of the random variable Y_{t+k}. The CDF for Y_{t+k} is given by $F_{t+k}(x) = P(Y_{t+k} \leq x)$. F_{t+k} is unobserved as only one realisation of Y_{t+k} is ever seen.

First, consider reliability. Reliability is assessed by considering the observed frequency of an event and the forecast probability of an event. Reliability is concerned with whether the forecast has the correct statistical properties.

For example, consider forecasts are given for 100 hours. The hour the forecast is made is denoted by *t* and the forecast horizon is denoted by *k*. At each of the time periods, the forecast is given by a probability distribution which defines the probability of a value or range of values occurring. The predicted CDF at time *t+k* is $\hat{F}_{t+k}(x) = P(Y_{t+k} \leq x)$. The CDF for a normally distributed random variable is given by Equation 2-21.

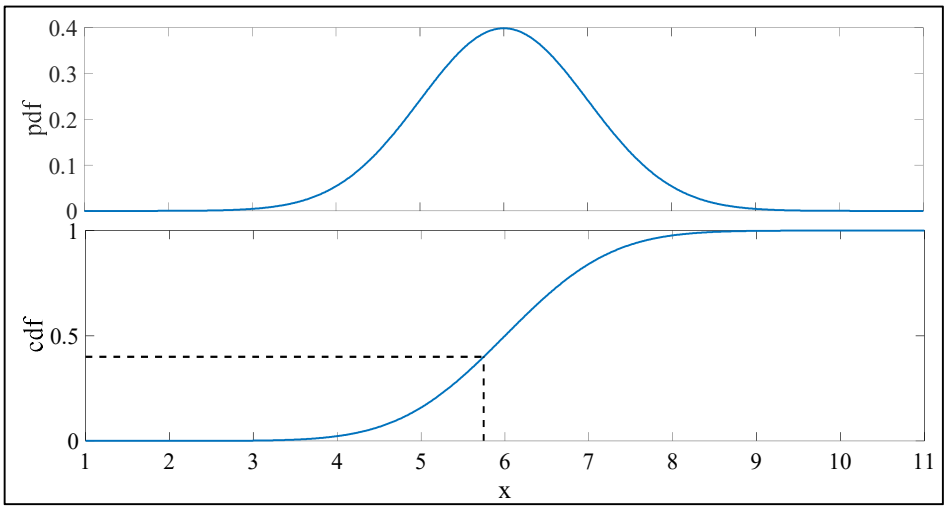
$$\hat{F}_{t+k}(x) = P(Y_{t+k} \leq x) = \frac{1}{2}\left[1 + erf\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right] \qquad \text{Equation 2-21}$$

where *erf* is the error function given by $\text{erf}(x) = \frac{1}{\sqrt{\pi}}\int_{-x}^{x} e^{-t^t} dx$. For example, assume that a forecast for time *t+k* = 4 is given by a normal distribution with parameters $\mu = 6, \sigma = 1$. The CDF and the pdf of the forecast are shown in Figure 2-7. From the CDF we can calculate the probability that the observed value will be less than a certain value. For example, the probability that the observed value will be less than 4 is $\hat{F}_4(4) = 0.0228$. Alternatively, if we wish to find the value, *x* for which the probability that the observed value will be less than or equal to with probability *p* the inverse CDF is used, given by Equation 2-22. In Equation 2-22 *inf* refers to the infimum of the set, the largest value which satisfies the conditions.

$$F_{X_t}{}^{-1}(p) = inf\{x \in \mathbb{R} : F_{X_t}(x) \leq p\}$$ Equation 2-22

For example, using the normal distribution with parameters $\mu = 6, \sigma = 1$ the value x for which $P(Y_{t+k} \leq x) = 0.4$ is $x = 5.74$. This is also shown on Figure 2-7.

Figure 2-7: pdf, CDF and $F_{t+k}^{-1}(0.4)$ for normal distribution with μ=6, σ=1



Once the observed value at the forecasted time is known, then for any value of x the event $Y_{t+k} \leq x$ either occurs or it does not. For example, using the forecasted distribution above the probability that the observed value is less than 5.74 is 0.4. This is equivalent to saying the probability of the event $Y_{t+k} \leq 5.74$ occurring is 0.4. However if the actual value is $Y_{t+k} = 3$ then the event occurs, whilst if $Y_{t+k} = 6$ the event does not occur. If the forecast is reliable then over a sufficient number of forecasts the observed frequency of an events occurrence will equal the forecasted probability of its occurrence. This is how reliability is measured.

The method for calculating reliability for a forecast is given by Pinson et al. [75]. To compare the forecasted probability of an event occurring with the frequency with which an event occurred, a probability $\alpha \in [0,1]$ is chosen. In the work of Pinson et al. [75] and Gneiting et al. [76], the value α is referred to as the "nominal probability". For each forecast and for each value of α calculate $\hat{F}_{t+k}^{-1}(\alpha)$ and consider the proportion of times over all forecasts that the observed value is less than this value. For a perfect forecast this proportion should be equal to α.

Mathematically, for a nominal probability, α, reliability is calculated over the entire forecast using a two-step process. First, for each prediction in the forecast an indicator variable is calculated using Equation 2-23. This indicator variable takes a value of 1 if the observed value at time $t+k$ is less than $\hat{F}_{t+k}^{-1}(\alpha)$, and takes a value of 0 otherwise.

$$\varepsilon_{t+k}^{\alpha} = \begin{cases} 1 \text{ if } y_{t+k} < \hat{F}_{t+k}^{-1}(\alpha) \\ 0 \text{ otherwise} \end{cases} \quad \text{Equation 2-23}$$

An estimate of the observed frequency of the event $y_{t+k} < \hat{F}_{t+k}^{-1}(\alpha)$ for a given look-ahead time k is given by Equation 2-24 for a test set of size N.

$$\hat{a}_k^{\alpha} = \frac{1}{N}\sum_{t=1}^{N} \varepsilon_{t+k}^{\alpha} \quad \text{Equation 2-24}$$

For example, assume hourly power output is predicted for a wind farm on the 1st January from 10 am – 3 pm. The forecasts are made 3 hours ahead, hence $k = 3$ and the first forecast is made at 7 am on 1st Jan ($t = 1$). All the forecasts are normally distributed with mean μ and standard deviation σ, hence $\hat{Y}_{t+k} \sim N(\mu, \sigma^2)$. Examples of the predictions and observed values are given in Table 2-4.

Table 2-4: Example probabilistic predictions and observed values

Forecast period (t+k)	Predictive distribution of $\hat{Y}_{t+k}$	Observed output (MW) (y_{t+k})
4	N(6,1)	10
5	N(7,2)	3
6	N(4,1)	8
7	N(8,2)	6
8	N(9,3)	2

Next, take a nominal probability value for example, $\alpha = 0.4$. For each forecast $F_{t+k}^{-1}(0.4)$ and $\varepsilon_{t+k}^{0.4}$ can be calculated. The values for these are given in Table 2-5.

Table 2-5: Inverse cumulative distribution and indicator variable for example forecasts

Forecast period (t+k)	Predictive distribution of $\hat{Y}_{t+k}$	Observed output (MW) (y_{t+k})	$\hat{F}_{t+k}^{-1}(0.4)$ (MW)	$\varepsilon_{t+k}^{0.4}$
4	N(6,1)	10	5.75	0
5	N(7,2)	3	6.49	1
6	N(4,1)	8	3.75	0
7	N(8,2)	6	7.49	1
8	N(9,3)	2	8.24	1

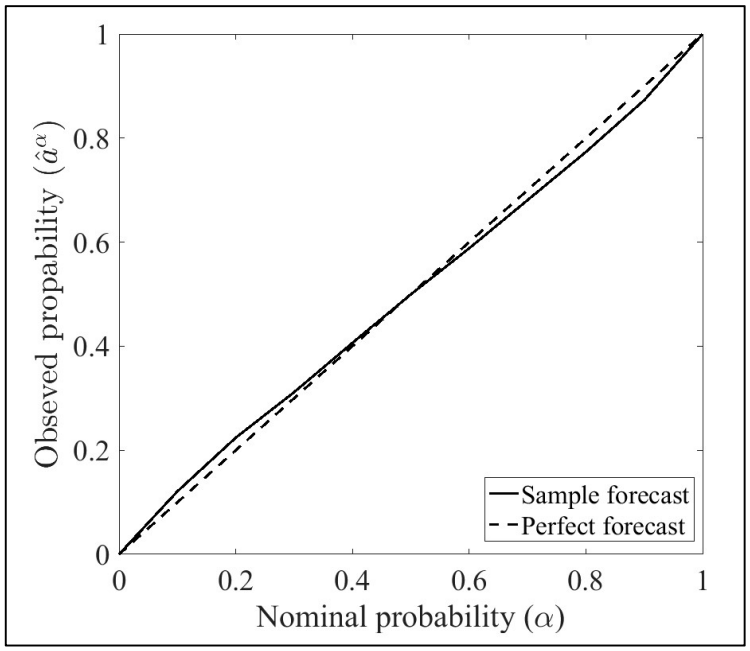
From Table 2-5, $\hat{a}_k^{\alpha}$ can be calculated for $k = 3$ and $\alpha = 0.4$ using Equation 2-24. This gives $\hat{a}_3^{0.4} = \frac{1}{5} \times 3 = 0.6$. In a perfect forecast $\alpha = \hat{a}_k^{\alpha}$ for all values of $\alpha \in [0,1]$.

Reliability is frequently considered in terms of the deviation from perfect reliability, b_k^{α}, given by Equation 2-25. In the example given above, this would give a value of $b_3^{0.4} = 0.4 - 0.6 = -0.2$.

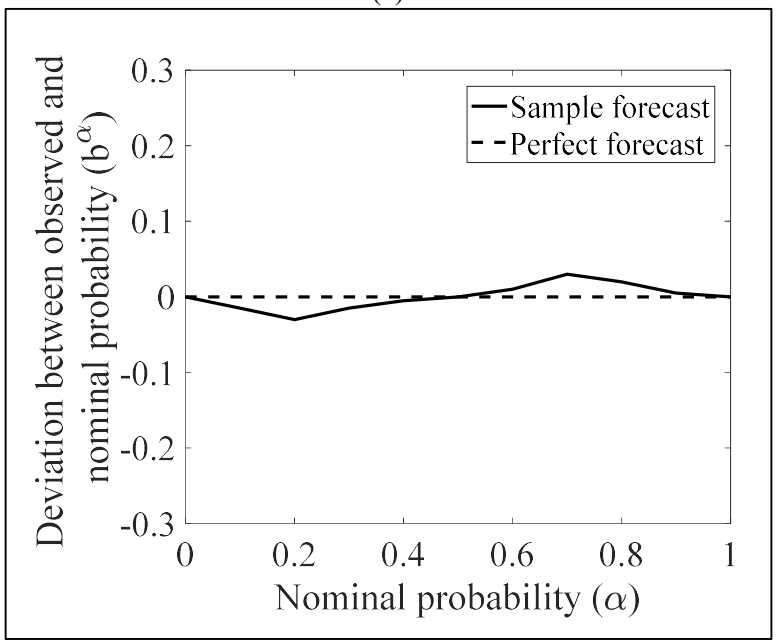
$$b_k^{\alpha} = \alpha - \hat{a}_k^{\alpha} \qquad \text{Equation 2-25}$$

Reliability can be displayed graphically by using a reliability diagram. In this either $\hat{a}_k^{\alpha}$ or b_k^{α} is plotted against α, as shown in Figure 2-8. In this a sample forecast is shown in comparison to a perfect forecast. When comparing results from different studies, the maximum deviation between observed and nominal probability for any value of α is compared. This is calculated by evaluating $\max\{|b^{\alpha}|\}\ \forall \alpha$.

Figure 2-8: Example reliability diagrams showing (a) observed probability against nominal probability and (b) deviation between observed probability and nominal probability against nominal probability.



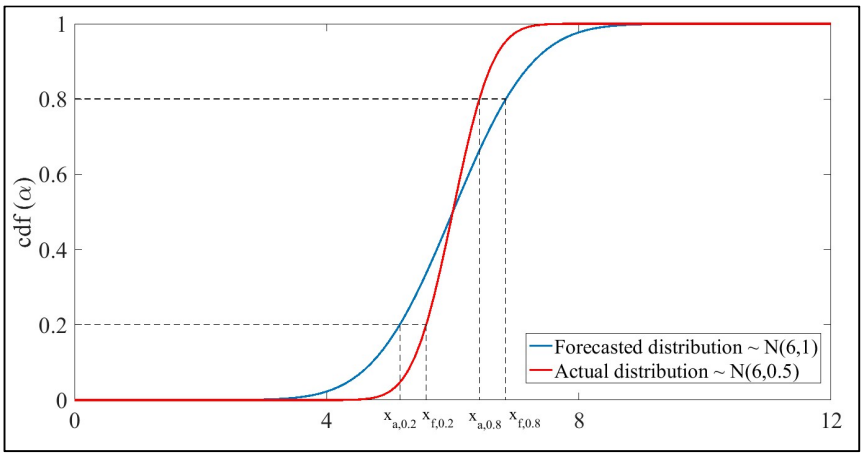
(a)



(b)

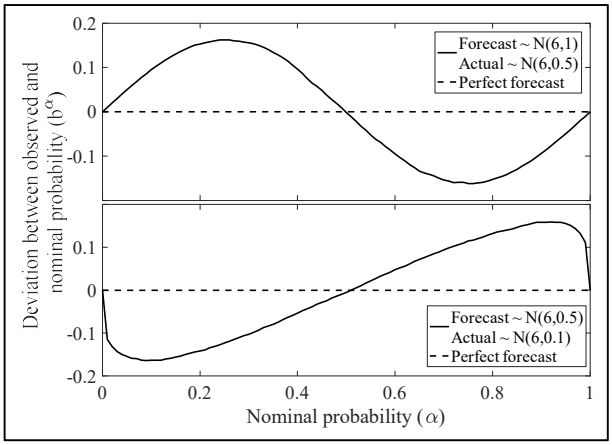
Visually, a few features of the forecasted distributions can be seen from a reliability diagram. Consider a simplified forecast made *k* hours ahead where for all time periods *t,* the forecast is a normal distribution with $\mu = 6, \sigma = 1$. If this forecast was correct for all values of *t* then $b_k^{\alpha} = 0$ for all values of α. If the observed values at all values of *t* were actually from a normal distribution with $\mu = 6, \sigma = 0.5$ then the forecasted standard deviation is larger than the actual standard deviation. In this case, for low values of α, e.g $\alpha = 0.2$ the observed frequency of the event $\left(y_{t+k} \leq \hat{F}_{t+k}^{-1}(0.2)\right)$ will be less than 0.2. When calculating reliability for the forecast, for $\alpha = 0.2$, the value x_f would be calculated such that $P\left(\hat{Y}_{t+k} \leq x_f\right) = 0.2$ holds for the forecasted distribution. In this example, the forecasted distribution for all *t* is normal with $\mu = 6, \sigma = 1$ and so $x_f = 5.16$. To evaluate reliability, calculate the proportion of observations that are below 5.16. The actual distribution is normal with $\mu = 6, \sigma = 0.5$. For the actual distribution, $F_{t+k}^{-1}(0.2) = x_a = 5.58$. Because of this, $\hat{a}_k^{0.2}$ would be less than 0.2, as in the actual distribution a proportion of 0.2 of the observations should fall below 5.58. The observations are a finite set, so the actual proportion of observations which fall below 5.58 may not exactly equal 0.2, but it should be close for a set of sufficient size. Conversely, for higher values of α, e.g $\alpha = 0.8$ the observed frequency of the event $P\left(y_{t+k} \leq \hat{F}_{t+k}^{-1}(0.8)\right)$ will be greater than 0.8. To add clarity the forecasted and actual CDFs are shown in Figure 2-9 with the values for x_a and x_f for $\alpha = 0.2$ and $\alpha = 0.8$.

Figure 2-9: CDF of an example forecast distribution in comparison the actual distribution.



Because of this property, if the standard deviation of the forecasted distribution is larger than that of the actual distribution a plot of $b_k{}^{\alpha}$ against nominal probability (α) will follow a pattern similar to that seen in the top graph in Figure 2-10. On the other hand if the standard deviation of the forecasted distribution is smaller than that of the actual distribution a plot of b_k^{α} against nominal probability (α) will follow a pattern similar to that seen in the bottom graph in Figure 2-10.

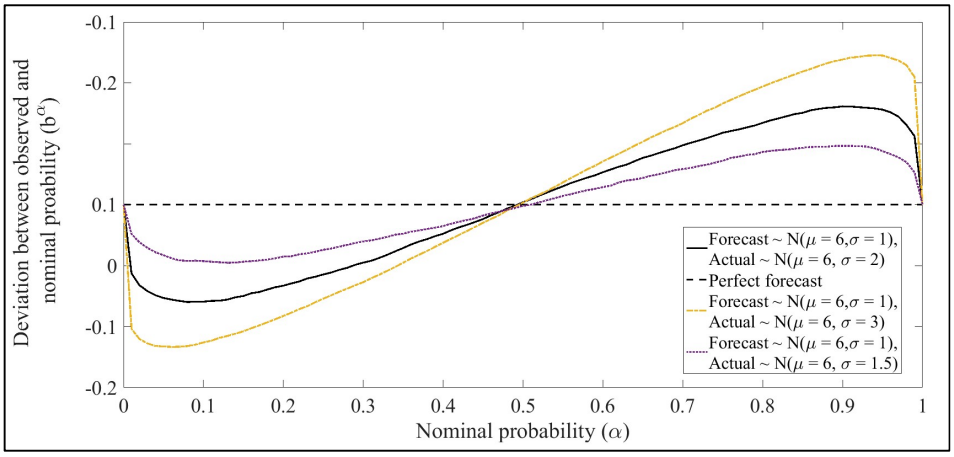
Figure 2-10: Deviation between observed and nominal probability for a forecast with a larger standard deviation than the observed values and a forecast with a smaller standard deviation.



Regardless of whether the forecasted standard deviation is larger or smaller than the actual distribution, the closer the forecasted distribution is to the actual distribution the smaller $\max\{|b_k^\alpha|\}\ \forall\alpha$ will be. Because of this, in a plot of b_k^α against nominal probability, α, the peak will be smaller for a forecasted distribution that is closer to the actual distribution. This is demonstrated in Figure 2-11.

For an actual forecast, the interpretation of a reliability diagram is more complex. For a real probabilistic forecast, the forecast distribution is not the same for all values of t, and reliability is calculated for a finite set. Hence $b^\alpha = 0\ \forall\alpha$ will not occur due to sampling errors and the interpretation of the reliability diagram will not be as clear as those in Figures 2-10 and 2-11. However, the closer $b^\alpha = 0\ \forall\alpha$ and the smallest value of $\max\{|b^\alpha|\}\ \forall\alpha$ indicates the most reliable forecast.

Figure 2-11: Deviation between observed and nominal probabilities for 3 different example forecasts.



Next, consider a method for measuring sharpness. Such a method is given by Pinson et al. [75]. Sharpness is a measure of mean interval size for different prediction intervals from a forecast. Sharpness is measured by considering the mean central prediction interval size for a chosen nominal coverage rate β for the whole forecast made at look ahead time k. This is calculated using Equation 2-26. Sharpness of

different forecasts can be compared by calculating $\overline{\delta_k^{\beta}}$ for a full range of nominal coverage rates or, most commonly, for specific rates such as $\beta = 0.5, 0.75, 0.9$.

$$\overline{\delta_k^{\beta}} = \frac{1}{N}\sum_{t=1}^{N}\left(\hat{F}_{t+k}^{-1}\left(\frac{1+\beta}{2}\right) - \hat{F}_{t+k}^{-1}\left(\frac{1-\beta}{2}\right)\right) \quad \text{Equation 2-26}$$

For example, returning to the example forecasts given in Table 2-5. In Table 2-6 the interval lengths are calculated for $\beta = 0.75$ and $\beta = 0.9$ for intervals centered on the median.

For the examples in Table 2-6, $\overline{\delta_k^{0.75}} = 3.22$ and $\overline{\delta_k^{0.9}} = 5.92$. In general forecasts with a smaller mean interval size contain less uncertainty and are therefore more useful. However, it is important to consider sharpness and reliability together. This is because as the mean interval size decreases, reliability can decrease if the standard deviation of the forecasted distribution is less than the actual distribution.

Table 2-6: Interval lengths for $\beta = 0.75$ and $\beta = 0.9$ for example forecasts

Forecast period (t+k)	Predictive distribution of $\hat{Y}_{t+k}$	$\delta_k^{0.75}$	$\delta_k^{0.9}$
4	N(6,1)	2.30	3.29
5	N(7,2)	4.60	6.58
6	N(4,1)	2.30	3.29
7	N(8,2)	4.60	6.58
8	N(9,3)	6.90	9.87

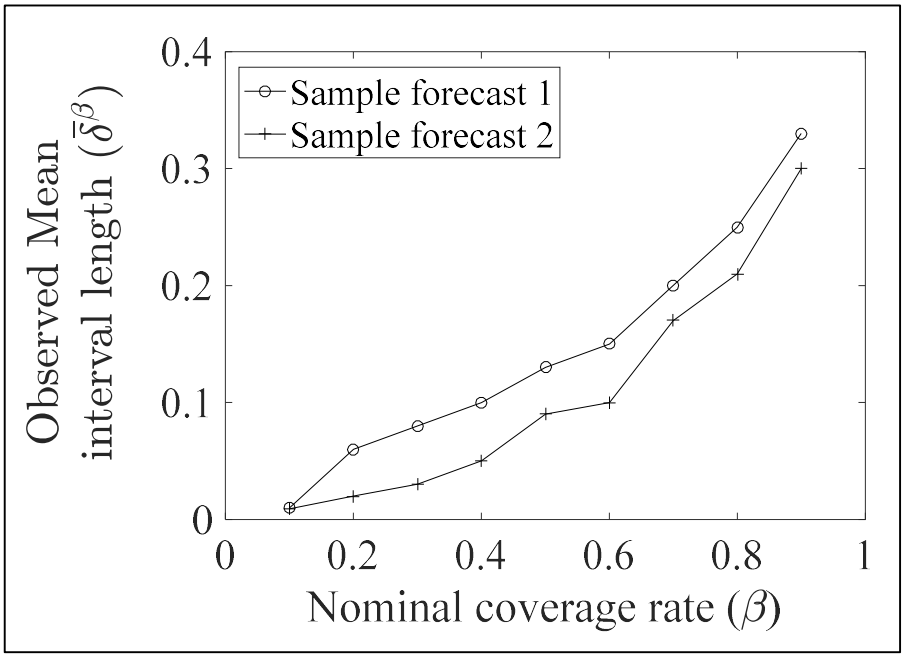
Finally, consider resolution. The most frequently used measure of resolution is the standard deviation of central prediction interval size. This gives an estimate of how different levels of uncertainty are represented by the model. Resolution σ_k^{β}, can be calculated for different values of β using Equation 2-27.

$$\sigma_k^{\beta} = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(F_{t+k}^{-1}\left(1-{}^{\beta}/_{2}\right)-F_{t+k}^{-1}\left({}^{\beta}/_{2}\right)-\overline{\delta_k^{\beta}}\right)^2} \qquad \text{Equation 2-27}$$

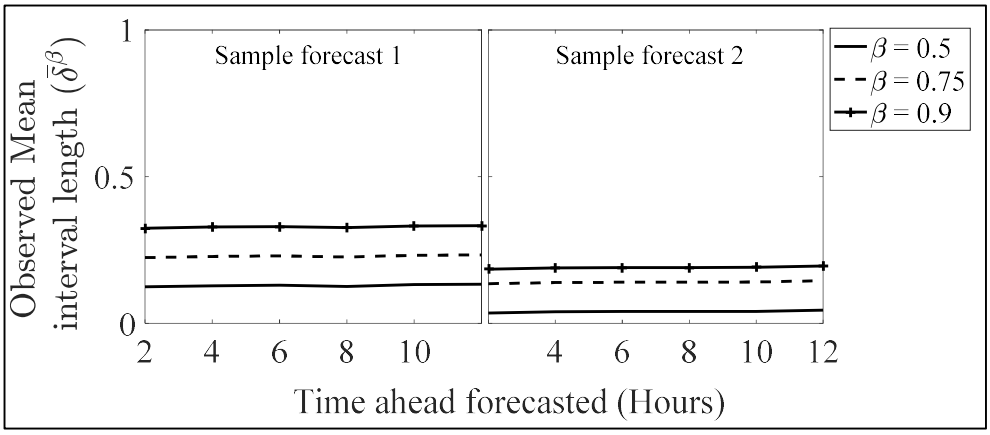
A larger value for σ_k^{β} indicates a more useful forecast, as this indicates a forecast is better able to represent uncertainty in different conditions. For example, if for all values of *t* the forecast distribution is normal with $\mu = 6, \sigma = 1$, then for all values of β $\sigma_k^{\beta} = 0$. Whilst this might be a reliable forecast, it does not give the user any information about the uncertainty at time *t*.

As with reliability, sharpness and resolution can be displayed graphically. Sharpness can be considered in terms of forecast horizon or nominal probability, as shown in Figure 2-12. A forecast which gives a lower mean interval length is generally more useful, as there is less uncertainty surrounding predictions. For example, for the sample forecasts given in Figure 2-12 “Forecast 2” gives a more useful forecast than “Forecast 1”. In contrast, a higher value for resolution indicates a more valuable forecast. This is because a larger standard deviation of interval size indicates a more expressive forecast and ability to differentiate between predictive conditions.

Figure 2-12: Example sharpness diagram is shown (a) for increasing nominal probability and (b) for a selection of nominal probabilities for increasing look ahead time.



(a)



(b)

Whilst reliability, sharpness and resolution give an overview of model performance, it can be difficult to compare competing models on the basis of these properties. For this reason, scoring rules are used to give a numerical value to forecast performance. A variety of probabilistic scores are available. However, the most commonly

reported amongst wind power prediction literature is the continuous ranked probability (CRPS) score given by Equation 2-28 [68].

$$CRPS\left(\hat{F}_{t+k}, y_{t+k}\right) = \int_{-\infty}^{\infty} \left(\hat{F}_{t+k}(x) - H(x - y_{t+k})\right)^2 dx \quad \text{Equation 2-28}$$

where $\hat{F}_{t+k}(x)$ is the forecasted CDF, y_{t+k} is the observed value and H is the Heaviside step function given by:

$$H(z) = \begin{cases} 0, & z < 0 \\ 1, & z \geq 0 \end{cases}$$

CRPS is a strictly proper scoring rule which is equivalent to MAE for deterministic forecasts. It provides a summary statistic for evaluating probabilistic forecasts [77]. Whilst CRPS is a useful method for evaluating probabilistic forecasts, finding a closed form expression for Equation 2-28 causes difficulties in practical implementation. Lerch et al. [78] present equations for calculating CRPS for different predictive distributions. For a normal distribution with parameters μ and σ a closed form expression for CRPS is given by Equation 2-29 [78] where $\hat{F}_{t+k}$, the forecasted distribution is normally distributed with parameters μ and σ, and y_{t+k} is the observed value at time $t+k$. In Equation 2-29 φ and Φ refer to the pdf and CDF of the standard normal distribution respectively.

$$CRPS\left(\hat{F}_{t+k}, y_{t+k}\right) = \sigma\left[\frac{y_{t+k}-\mu}{\sigma}\left(2\Phi\left(\frac{y_{t+k}-\mu}{\sigma}\right) - 1\right) - 2\varphi\left(\frac{y_{t+k}-\mu}{\sigma}\right) - \frac{1}{\sqrt{\pi}}\right] \quad \text{Equation 2-29}$$

CRPS over the full forecast is given as an average of CRPS calculated for all t. The average value over the full forecast is what is usually referred to as CRPS in literature, hence this notation is used here. Average CRPS is calculated using Equation 2-30.

$$CRPS(\hat{F}_k) = \frac{1}{N}\sum_{t=1}^{N} CRPS(\hat{F}_{t+k}, y_{t+k})$$ Equation 2-30

As CRPS is equivalent to MAE for deterministic forecasting, a forecast with a lower CRPS is better than one with a higher CRPS. Other closed form equations for CRPS relevant to the discussion in this thesis are given in Chapter 5.

2.2.8 Summary and comparison of probabilistic forecasting methods

In this section, current state of the art forecasting techniques are discussed. This gives an idea of the methods employed to generate probabilistic forecasts for wind forecasting and the results reported for these methods. Table 2-7 gives an overview of current probabilistic forecasting techniques described in the literature.

Beginning with quantile forecasts, Bremnes [79] used local quantile regression to provide forecasts from 24 – 47 hours ahead. Haque et al. [80] also use quantile regression, this time providing forecasts from 1-24 hours in advance. Sideratos and Hatziargyriou [81] used a neural network to provide quantile forecasts from 1-60 hours ahead. Whilst the results of this method seem promising, it is time consuming and complex to train for numerous relevant quantiles. Messner et al. [82] used an inverse power curve to try and estimate uncertainty, giving forecasts 24 and 48 hours in advance. The main issue with this method is that there is uncertainty introduced by different meteorological conditions when making power predictions from power curves, as discussed in Section 2.1.3. For example, different stability conditions have been shown to affect the appropriate power curve for a site [40].These conditions are not accounted for when an estimate of uncertainty is given. Finally, Nielson et al. [83] used ensemble wind speed forecasts to give quantile forecasts of wind power production. This was done for longer term forecasts than any of the other methods, giving forecasts up to 7 days ahead. The use of NWP ensemble forecasts allows for good predictions and estimates of uncertainty to be made quite far in advance. However, this method could be impractical for

many users as it requires an NWP model to be run with different starting conditions which is expensive. The results for these methods are given in terms of length of average prediction intervals, or sharpness as described by Equation 2-26. Most authors at least gave results for 50% and 90% prediction intervals. However, some only gave average width of 40% or 80% intervals, making it difficult to compare. Because of the different forecast horizons used, and in some cases different prediction intervals it can be difficult to compare the results. From the results shown in Table 2-7 it can be seen that for forecasts at 24 hours ahead, the average length of a 50% or 90% confidence interval is around 20% or 50% of capacity respectively. These results are very similar for the methods presented by Bremnes [79], Haque et al. [80] and Nielson et al. [84]. There is no single method which outperforms the others and the results are consistent despite the capacity of the test site varying from 2 MW to 168 MW. Quantile forecasts offer a good way of highlighting the quantities of interest in a probabilistic forecast and allow for easier visualisation and comparison. However, difficulties can be encountered when many scenarios need to be considered when applying a forecast, in which case a full predictive distribution can be more valuable. One issue noted with quantile regression is that that model needs to be re-optimized to predict more than one different quantile, which can make the model slow and complex if many quantiles or forecast periods are required.

Only one author reviewed described their method as an interval forecast. Carpinone et al. [85] used a Markov chain model to estimate an interval forecast from a point forecast. In this case the results were given in terms of NRMSE so cannot be compared to other probabilistic forecasts.

Finally, density forecasts are considered. The most common method for density forecasting is a kernel density estimator. Bessa et al. [86] presented a method which used a kernel density estimator to give a predictive pdf for forecasts from 6 to 48 hours ahead. The advantage of this model is its ability to be updated to reflect new data obtained. Reliability and sharpness were given for a single wind farm (unknown size) and 15 hypothetical sites. At the single wind farm the maximum

deviation from nominal probability ($\max(|b_k^\alpha|)\ \forall\alpha$) was 8% and for the 15 hypothetical sites it was 4%. The average length of a 50% prediction interval was 12% over the 15 sites and 20% at the single wind farm. Juban et al. [87] also used a kernel density estimator, this time presenting results averaged across three wind farms. The reliability and sharpness were similar to the results presented by Bessa et al. [86] for the 15 hypothetical wind farms. Zhang et al. [88] also used a kernel density estimator to predict power from 1 – 24 hours in advance. Maximum deviation from nominal probability was given as 3%, and average length of a 50% interval was 12% of capacity. Again, this is very similar to the results given by Bessa et al. [86] and Juban et al. [87]. Kou et al. [89] used a Gaussian process model with warping function to predict power output from 15 minutes to 36 hours ahead. The only comparable results given is reliability, which shows a maximum deviation from nominal probability of 3%. This is similar to the reliability shown by Juban et al. [87]. Tatsu et al. [90] generated predictive distributions based on a point forecast and average errors in the area. Forecasts were given from 15 minutes – 8 hours ahead. The only comparable result presented is CRPS which is between 4 and 11% of capacity. The work outlined by Kou et al. [91] indicates that the GPR model is successful in modelling wind power and uncertainty in these forecasts. However the few results are shown in a framework which allows comparisons to other methods to be drawn. Density forecasts are a popular method for estimating uncertainty in wind power predictions. This is because, despite their complexity, they typically offer more information than quantile or interval forecasts and can be adapted well to suit the user. This allows wind power uncertainty to be incorporated into optimization problems, for example establishing an appropriate trade-off between risk and profit when generators are competing in electricity markets.

It can be seen from the methods presented above that there is a large range of suitable techniques which have been used for generating probabilistic wind forecasts. Overall no single method outperforms other methods reliably. However, there is still room from decreasing prediction errors. Furthermore, the results show the value of providing a number of evaluation metrics to enable comparison of results with other methods.

The practicality of density forecasts in the wind industry, and the promising results shown using a hybrid NWP and GPR model make this an interesting area for future research. GPR has been shown to be flexible and able to model the complex relationships between wind speed and power output. However, further work is required to establish whether the inclusion of further meteorological variables can improve model performance. In addition to this, appropriate ways to implement the model given the non-Gaussian nature of wind power predictions require further research. It is for these reasons that this thesis focuses on the use of GRP for wind speed and power predictions.

Table 2-7: Overview of currently available probabilistic forecasting techniques

Author	Year	Title	Forecast type	Summary	Test data	Forecast period	Results
Bessa et al. [82]	2012	Time-adaptive quantile-copula for wind power probabilistic forecasting	Density forecast	Uses kernel density estimator to give a predictive pdf. Results are compared to splines quantiles regression. Predicts wind power at one wind farm in the mid-west US	15 hypothetical sites in Illinois and one wind farm in mid-west US	6 to 48 hours ahead	Reliability: up to 4% (15 sites), up to 8% (single wind farm) deviation from nominal proportions. Average length of 50% interval: ~12% (15 sites), ~20% (single wind farm). Average length of 90% interval: ~30% (15 sites), 60% (single wind farm).
Bremnes [76]	2004	Probabilistic wind power forecasts using local quantile regression	Quantile forecast	Predictions of different quantiles are made for predictions based on regression.	Vikna wind farm in Norway (2.2 MW)	24, 30, 36, 42 and 47 hours ahead.	Average length of 50% prediction interval: 0.25 - 0.45 MWh. Average length of 90% prediction interval: 1.2 - 1.5 MWh.
Carpinone et al. [81]	2010	Very short-term probabilistic wind power forecasting based on Markov chain models	Interval forecast	Uses a discrete time Markov chain model to estimate an interval forecast from a point forecast.	One American wind farm	10 mins intervals up to 200 mins ahead	NRMSE (% of nominal power): 4 - 36%
Haque et al. [77]	2014	A hybrid intelligent model for deterministic and quantile regression approach for probabilistic wind power forecasting	Quantile forecast	Uses quantile regression to predict wind power output at a wind farm in Colorado	Cedar Creek wind farm in Colorado	1 - 24 hours ahead	Average length of 50% prediction interval: ~20% % of capacity. Average length of 90% prediction interval: 51 % of capacity. Reliability: deviation from nominal proportion 1-4%.

Juban et al. [83]	2007	Probabilistic short-term wind power forecasting based on kernel density estimators	Density forecast	Produces pdf based on kernel density estimation.	3 wind farms in France.	up to 60 hours ahead	Reliability: +/-2% deviation from nominal probabilities. Average length of 50% interval: ~13%. Average length of 90% interval: ~32%. CRPS: between 0 and 12%.
Kou et al. [85]	2014	Probabilistic wind power forecasting with online model selection and warped Gaussian process	Density forecast	Extends a standard Gaussian process by including a warping function. Uses an online ensemble model.	A 300 MW wind farm	15 minute intervals up to 36 hours ahead	Reliability: +/-3% deviation from nominal probabilities.
Messner et al. [79]	2014	Probabilistic wind power forecasts with an inverse power curve transformation and censored regression	Quantile forecast	Generates quantiles using an inverse power curve.	A 2 MW Turbine in Austria/	24 and 48 hours ahead	Average length of 40% prediction interval at 24 hours ahead: 15-16%. Average length of 80% prediction interval at 24 hours ahead: 41-42%. Also shows a market skill score.
Nielsen et al. [80]	2006	From wind ensembles to probabilistic information about future wind power production - results from an actual application	Quantile forecast	Uses ensemble wind speed prediction to generate quantile forecasts	166 MW Offshore wind farm, Denmark (Nysted), and the region of western Denmark, 2200 MW.	Up to 7 days	Average length of 50% prediction interval: 20 - 60 % of capacity (for single farm), 10 - 40% of capacity (for region. Average length of 80% prediction interval: 50 - 90 % of capacity (for single farm), 20 - 60% of capacity (for region).

Sideratos and Hatziargyriou [78]	2012	Probabilistic wind power forecasting using radial basis function neural networks	Quantile forecast	Uses point predictions of an existing wind power forecasting model and forecasts the prediction uncertainty due to the inaccuracies of the NWP. Predicts set of quantiles with predefined nominal probabilities	2 wind farms: Lasithi, Crete (18 MW) and Klim, Denmark (21 MW).	up to 60 hours ahead	CRPS: 10% (Lasithi) 6-10% (Klim). Reliability: +/-2% (Lasithi), +/- 1.5% (Klim) deviation from the nominal probabilities. Average 50% interval length: 20% (Lasithi), 15% (Klim). Average 90% interval length: ~50% (Lasithi), 40% (Klim)
Tastu et al. [86]	2014	Probabilistic forecasts of wind power generation accounting for geographically dispersed information	Density forecast	Generates predictive distribution based on average point forecast errors in the local area.	offshore wind farm in Denmark (Nysted, 165 MW)	15 mins to 8 hours	CRPS between 4% and 11%
Yan et al. [88]	2016	Hybrid probabilistic wind power forecasting using temporally local Gaussian process	Density forecast	Uses a temporally local GP employing nearby local datasets and a moving window technique	A 300 MW wind farm in the USA and a 60 MW wind farm in Ireland.	1 - 12 hours ahead	The results are not evaluated in a probabilistic framework.
Yang et al. [89]	2012	Probabilistic short-term wind power forecast using componential sparse Bayesian learning	Density forecast	Generates a predictive distribution by estimating the probabilistic density of the weights of Gaussian kernel functions. Tested on a 74 MW wind farm in Oklahoma. Time horizon 48 or 72 hours ahead	74 MW wind farm in Oklahoma	48 and 72 hours ahead	The results are not evaluated in a probabilistic framework.

Zhang et al. [84]	2015	Probabilistic wind power forecasting based on logarithmic transformation and boundary kernel	Density forecast	Uses a kernel density estimator to give non-Gaussian predictive distributions. Evaluated for a 542 MW wind farm in Iowa. Forecast horizon 1 - 24 hours	542 MW wind farm in Iowa	1 - 24 hours ahead	Reliability: +/-3% deviation from nominal probabilities. Average 50% interval length: ~12%. Average 90% interval length: ~30%. CRPS: 0.02 - 0.08 (1-24 hours).

Probabilistic forecasting is an area of research which presents an opportunity for methods which would aid the advancement of wind energy integration and potentially improve planning and balancing mechanisms. The development of probabilistic prediction techniques allows those responsible for grid management to have not only an estimation of the potential power output but also an idea of uncertainty and range for potential generation. This enables backup power and balancing mechanisms to be planned more efficiently which can reduce costs. Whilst some methods have been developed as mentioned above, this area has received relatively little attention and comparisons of methods are sparse. The hybrid NWP and GPR model presented in this thesis allows for probabilistic forecasting and provides an alternative way of developing probabilistic wind power forecasting. The method and its contribution to probabilistic literature are discussed further in Chapters 3, 4 and 5. In the next section, the applications of both deterministic and probabilistic forecasts are discussed.

2.3 Applications of wind power forecasting

So far, the available techniques for wind speed and power prediction have been discussed alongside methods for evaluating their performance. In this section, the applications of wind power forecasting are examined. This allows us to consider how wind speed and power forecasts might be used in a commercial setting and why improving these forecasts is important. In Section 1.4 three ways in which wind power forecasting is used were introduced. These were: using wind power forecasting to schedule generating units to meet demand, to set reserve requirements to maintain system reliability and to enable wind power producers to compete in electricity markets. These applications are discussed further here. In addition to these applications, wind power forecasting can help with the design and operation of electricity storage facilities. Knowledge of the normal discrepancy between predicted power output and actual power output helps users to assess the storage capacity which would be beneficial and forecasting future power output allows operators to assess when to charge and discharge from storage facilities. For example, Pinson et al. [92] develop a method for optimising required storage capacity on a dynamic basis based on the uncertainty in wind power forecasts,

concluding that various non-dispatchable electricity generators could use forecasting and combined storage capacity to reduce costs. Alternatively, Garcia-Gonzales et al. [93] developed a method which takes into account the uncertainty in both day ahead prices and wind generation to optimise the use of energy storage facilities. Currently electricity storage devices are infrequently used in the UK, particularly alongside renewable generators to manage intermittency hence storage is not considered in detail in this thesis. However, as the cost of storage capacity decreases and if market mechanisms encourage investment storage may become more relevant to the UK electricity system. In order to consider how wind forecasting may be helpful in these scenarios, an explanation of electricity system management and electricity markets is given. This allows a discussion of how wind forecasting can be used to balance supply and demand and to improve revenue for a wind farm. The applications discussed here highlight why wind power forecasting is important and how improvements in wind power forecasting might impact various aspects of the electricity system. The concepts introduced here inform the discussion of how the value of a forecast might be evaluated, one of the objectives of this thesis highlighted in Section 1.5. The concept of the value of forecasting is discussed mostly in Chapter 6. However, the concepts introduced here are fundamental to considering why improved wind forecasting is valuable.

2.3.1 Electricity system management and grid balancing

Electricity systems are a complex network linking the electricity generators to end users by a collection of transmission and distribution systems. Electricity is produced by generators who deliver electricity to suppliers via the transmission network. Suppliers then deliver electricity to consumers via the distribution network. In the UK, the transmission system is owned and operated by the National Grid. The UK electricity system requires continuous supply and demand balancing to manage variations in supply and demand for electricity and to maintain a secure supply at all times. The National Grid is responsible for this system balancing, ensuring that supply meets demand on a second by second basis. The process of planning the delivery of electricity to meet demand is known as dispatch planning or unit

commitment planning. This is a complex optimisation problem referred to as the unit commitment problem. The planned delivery of electricity is usually optimised to reduce costs or minimise the chance of there being insufficient generation to meet demand (loss of load probability). This has to be done taking into account constraints such as individual unit capacity, time taken to start up and minimum run time of units [94]. Methods employed to solve this include dynamic programming, integer programming, Lagrangian relaxation and Tabu search [95]. The unit commitment problem is the subject of ongoing research, with attempts being made to solve the problem in more efficient ways which better reflect real conditions.

In most countries with a well-established electricity system, the supply network first began operating with the majority of power supplied by coal-fired power stations, with gas power stations adding to capacity subsequently. With the addition of variable renewable energy to the electricity system the unit commitment problem becomes even more complicated. Not only do the start-up times, minimum run times and the unit capacity need to be taken into account but also the uncertainty in variable renewable generation.

In addition to planning generation of electricity to meet expected demand, some capacity must be available to meet demand if units are unable to generate as scheduled. For example in the event of a technical failure. This capacity is known as reserve. Reserve can be either capacity available to either increase generation or decrease demand. Reserve capacity comprises of spinning (synchronised) and non-spinning (non-synchronised) reserve. Spinning reserve is additional capacity available from units already providing power at that moment and non-spinning reserve is reserve available from units which can be brought online in addition to units already providing power. Spinning reserve can usually respond more quickly than non-spinning [96]. Reserve is sometimes categorised as frequency response reserve and capacity replacement reserve. Most electricity system users require frequency to be stable. In the UK the National Grid is obligated to maintain the frequency within operational limits of 49.8 Hz – 50.2 Hz. Frequency response reserve is fast response reserve used to return the system frequency to an acceptable

level. Capacity replacement reserve is capacity available over a longer time period to cover the loss of generation. For example, due to a technical failure at a power plant. The amount of reserve required depends on the volume and type of electricity required at any one time. The allocation of reserve must balance the cost of providing reserve with the risk of loss of load.

Planning the dispatch of units to maintain system reliability is a complicated task. Both demand and supply for electricity are variable over time, and predictions of each contain some uncertainty. In order to effectively match supply and demand energy systems must contain some flexibility. Denholm and Hand [97] define system flexibility as “the general characteristics of the ability of the aggregated set of generators to respond to net load”. Traditional thermal electricity generators such as coal and gas provide flexibility in the supply due to vast stores of fuel. They are only constrained by ramping capacity and total capacity and can typically respond to fluctuations in demand needed to maintain a stable grid. Maintaining a portfolio of different types of power plant allows the provision of a flexible reaction to demand [98]. Kondziella and Bruckner [99] state that the magnitude of costs of integration of large amounts of renewable energy is directly linked to system flexibility. Costs for system balancing, grid-related costs and curtailment (hence the subsequent loss of revenue) all increase if a system is unable to respond to fluctuations in net load and deviations from predicted supply and demand. Increasing non-dispatchable renewable energy generation in a system requires a more flexible generation mix, relying on capacity which can be called upon on different time scales. Different technologies can contribute to this flexibility. For example, new gas turbines can generate power at full capacity in up to 1 or 2 hours (depending on the type) while coal power stations take much longer (between 12 and 72 hours) [100]. Because of this coal and nuclear power plants are used to provide baseload power and gas plants can be used to react to demand. In addition to this, pumped hydropower can provide power at very short notice. In the UK flexibility in the electricity system is achieved in a number of ways. Pumped hydro facilities located in North Wales can provide an extra 2 GW of power the UK network in up to 16 seconds [101]. Interconnections also exist between the UK and France, Ireland, the Isle of Man and the Netherlands,

allowing 4 GW capacity to be transferred in total [102]. Further capacity is available from fast response plants such as open-cycle gas turbines which can provide some power in a few minutes [103]. Some power stations are paid to run partly loaded so that power may be increased where necessary, known as spinning reserve. Alternatively, some large consumers of electricity can be called upon to reduce their demand to balance available supply of electricity. All these methods allow the power supply to be balanced on a very short term basis to account for any mismatch between supply and demand. Other ways that an electricity system can be flexible include demand side response and the use of energy storage.

Given the importance of a flexible system for the successful integration of variable renewable energy, we must consider how a flexible electricity system can be maintained with reduced use of fossil fuels and continued security of supply. Wind power is non-dispatchable and cannot provide the same ability to meet demand. Hence in order to effectively integrate a large amount of wind into the electricity system other strategies must be considered. One way of doing this is through the introduction of market mechanisms aimed at supporting the growth of renewables and ensuring sufficient capacity is available. Some mechanisms introduced in the UK to support the growth of renewables are discussed in Section 2.3.3. Forecasting output from variable renewable energy sources is another way of reducing the total system requirement for flexibility. If wind power can be accurately forecasted there is less requirement for a large amount of plant with very fast start-up times in order to cover a potential lack of generation. This is because electricity can be procured further in advance to meet demand allowing more efficient use of power plants. The use of wind power forecasting for system management is discussed further in Section 2.3.4.

2.3.2 Electricity markets

As mentioned previously, in the UK electricity is produced by generators which deliver electricity to suppliers via the transmission network. Suppliers then deliver electricity to consumers via the distribution network. Suppliers estimate

consumption in any given settlement period and contract to buy electricity from generators. Electricity is traded in half hour settlement periods. Trading of contracts occurs from many years in advance up to one hour before the settlement period, known as gate closure [102].

The National Grid electricity transmission (NGET) is responsible for managing supply and demand on a second by second basis, as suppliers and generators may not always match supply and demand perfectly before gate closure. This is done using the balancing mechanism which allows NGET to manage contracts for increases or decreases in both supply and consumption at very short notice. The adjustments made at this time are known as balancing actions. After the trading period, the actual volume of electricity supplied or used is calculated for each participant in the electricity market. If suppliers or generators demand or supply more or less than they contracted for, they are liable to an imbalance price for the difference, based on the NGET cost of balancing. Prior to November 2015, this was achieved through a dual pricing system. Under this system, energy generated that exceeded the final contracted volume of electricity is sold at the system sell price (SSP), and generators must pay for energy not generated at the system buy price (SBP) [23]. If a generator does not generate sufficient electricity to meet their contracts they are short on their contracts, whereas if they produce an excess of electricity they are long on their contracts. If at gate closure the total volume of contracts to supply electricity is less than the total volume of contracts for demand then the system is said to be short. If the opposite is true, the system is said to be long. Under this system, SBP was always greater than SSP, and the market price would fall between the two. The idea of this was that the payment for being long should be less than the market price and the charge for being short should exceed the market price [104].

Since November 2015, through the implementation of balancing and settlement code (BSC) modification P305, this has moved to a single imbalance price (SIP) which is paid to the generator for surplus energy and the generator must pay to the National Grid for energy deficits [105]. The imbalance price is based on the average

cost of balancing actions. This is calculated by averaging over a specified volume of the most expensive balancing actions in a half hour delivery period. This volume is known as the price average reference (PAR) volume. Prior to the implementation of P305, a PAR volume of 500 MWh was used (PAR500). Under the system prior to November 2015 the most expensive 500 MWh of balancing actions was used to set a value for the main imbalance price (SSP if the system was long, SBP if the system was short) [104]. The alternative imbalance price was set as the market index price (MIP). This is the average value of contracts traded in a settlement period, and reflects the wholesale price of electricity at that time [106]. Since the implementation of P305, the average of the most expensive balancing actions is used to calculate the SIP. In addition to the change from a dual pricing system to a single pricing system, the PAR volume was reduced from 500 MWh to 50 MWh in November 2015. This will change further in November 2018 when the SIP will be based on the average price of the most expensive 1 MWh of balancing actions. To illustrate this an example of prices for a hypothetical settlement period is shown. In this period there were 600 MWh of balancing actions required, made at prices ranging from £70/MWh to £96/MWh as shown in Table 2-8. These are shown ordered by descending price in Figure 2-13.

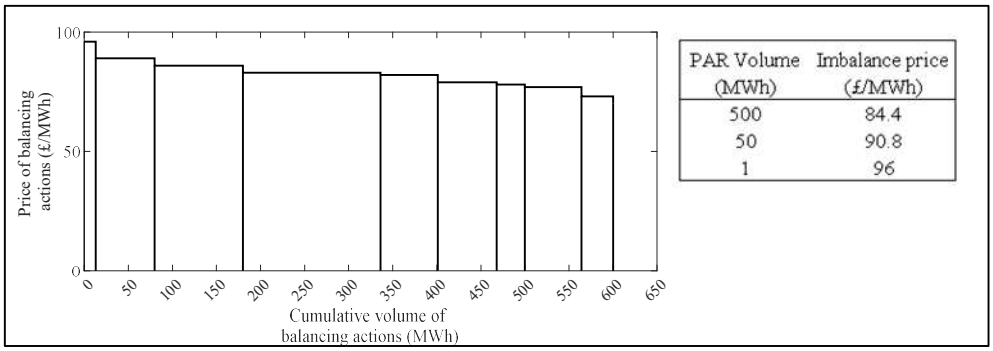
When the PAR volume is 50, the imbalance price is the average of 50 cumulative MWh of the most expensive balancing actions. From the example given in Table 2-8 and Figure 2-13, the imbalance price when PAR volume is 50 MWh is given by $SIP_{PAR50} = \frac{(13\times96)+(37\times86)}{50} = £90.8/\text{MWh}$. With a PAR volume of 500 MWh the imbalance price would be £84.40/MWh, increasing to £96/MWh when the PAR volume is decreased to 1 MWh. The changes to the imbalance pricing system and PAR volumes have led to an increase in volatility in system prices [107]. This is discussed further in Chapter 6.

Electricity is bought and sold through contracts made through either power exchanges or bilateral agreements made between generators, electricity suppliers, large consumers and non-physical traders. This trading creates a liberal market where prices are established according to supply and demand.

Table 2-8: Cost and volume of balancing actions for example trading period

	Volume of balancing action (MWh)	Price of balancing action (£/MWh)
	156	82
	100	83
	67	78
	67	86
	65	79
	64	73
	36	70
	32	77
	13	96
Total:	600	724

Figure 2-13: Cost of balancing actions in example balancing period shown in Table 2-8 and resultant imbalance prices.



Whilst trading is mostly governed by profitability, certain behaviour needs to be encouraged through financial incentives. For example investment in low carbon technologies to reduce the emissions from electricity generation and investment in capacity to ensure the security of supply in the future must be incentivised

financially. In the next section the market reforms which have been introduced in the UK are discussed. These are designed to encourage investment in renewable energy and to encourage the system flexibility required to support increased renewable energy.

2.3.3 Electricity market reform

Various market mechanisms have been introduced to encourage investment in low carbon technologies and maintain system reliability. The electricity market reform (EMR) was proposed by the UK government in 2011, and embedded in legislation through the energy act in 2013 [108]. The EMR was introduced to encourage investments in low carbon electricity to support future growth in renewables whilst improving security of the UK electricity supply and maintaining affordability. Since electricity generation units in the UK are owned by private companies the government cannot simply build capacity required to meet future demand. The aim of the EMR is to encourage low carbon investment in the UK electricity supply to ensure sufficient capacity is available to meet demand in the future. The mechanisms introduced aim to encourage investment in low carbon technologies and develop an electricity system with sufficient flexibility to enable the integration of variable renewable energy.

The EMR proposes the introduction of two mechanisms; feed-in tariffs with contracts for difference and capacity agreements within a capacity market supported by a carbon price floor and emissions performance standard [109]. The details of these are outlined in the following sections.

2.3.3.1 Feed-in Tariffs with Contracts for Difference

Contracts for difference (CFD) are private contracts between low-carbon electricity generators and the Low Carbon Contracts Company which is owned by the UK government [110]. This is a mechanism for payment for low carbon generation in which a pre-agreed electricity price is guaranteed to the generator. This means that when the price of electricity falls, an additional payment will be made to the

generator and when it rises the generator will pay some money back. CFDs replace renewable obligation certificates (ROC). Under the ROC system, electricity generators receive ROCs from Ofgem for renewable energy generated [111]. Electricity suppliers were obligated to obtain a (annually increasing) proportion of their annual sales from renewable sources [112] giving evidence of this by buying ROCs. If a supplier did not buy enough ROCs they were required to pay a buy-out price which was collected by Ofgem and redistributed to suppliers with ROCs [111]. This gave another source of revenue for renewable energy generators [113]. The ROC scheme was closed to new onshore wind generators in May 2016 and other new generators in March 2017 [114]. However, support will continue for those already registered for 20 years from registration [113].

The new system of CFDs aims to encourage investment in low-carbon generation by providing a stable and predictable income. However significant issues surround how to price technologies such that private companies will invest, whilst keeping electricity prices from rising drastically. The first CFD allocation was completed in 2014, with the results published in 2015 [115]. The total allocation of contracts and the average strike price are listed by technology type in Table 2-9. From this it can be seen that the strike price for offshore wind was high in comparison to other technologies.

Table 2-9: Summary of CFD allocation round 1 [115]

Technology	Total Contracts (MW)	Average strike price (£/MWh)
Advanced Conversion Technology (ACT)*	62	118.05
Energy from waste with combined heat and power	94.75	80
Offshore Wind	1162	117.14
Onshore Wind	748.55	81.94
Solar PV**	71.55	67.53

*ACT refers to a range of technologies which convert waste to energy.
**Solar PV refers to larger solar photovoltaic farm projects, with an average capacity of 14 MW.

The second auction was held in April 2017 [116]. This included caps on the maximum strike price for some technologies, outlined in Table 2-10. These maximum prices are designed to encourage the industry to reduce the price of technologies. The contract amounts and values by technology are shown in Table 2-11 [116].

Table 2-10: Maximum strike prices for CFD allocation 2 [117]

	£/MWh, 2012 prices		
Technology	2018/19	2021/22	2022/23
Offshore wind	140	105	100
Advanced conversion technologies	140	125	115
Anaerobic digestion	140	140	135
Biomass with CHP	125	115	115
Wave	305	310	300
Tidal stream	305	300	295
Geothermal	140	Under consultation	

Table 2-11: Summary of CFD allocation round 2 [116]

Technology	Total Contracts (MW)	Average strike price (£/MWh)
Advanced Conversion Technology (ACT)*	64.31	40.43
Biomass with combined heat and power (CHP)	85.64	74.75
Offshore Wind	3196	62.14

*ACT refers to a range of technologies which convert waste to energy.

From Table 2-11 it can be seen that there was a significant reduction in the strike prices of all technologies in the second CFD allocation round. For offshore wind there was a 48% decrease in the strike price, indicating there has been significant accomplishments in reducing the cost of delivering offshore wind projects in the last few years. This is likely to lead to increasing numbers offshore wind projects in the

future, which will need to be managed carefully to avoid increases in the cost of delivering electricity due to managing the integration of variable renewable electricity. Onshore wind was not included in the second round of CFD allocation as this is considered a mature technology. Hence onshore wind energy developers are subject to market prices without the support or security of guaranteed income from CFD. This suggests that for new onshore wind farms the ability to predict power output and participate in electricity markets effectively will become increasingly important as alternative revenue streams have been removed.

2.3.3.2 Capacity Markets

Traditionally in the UK investment in energy relied on high peak energy prices to encourage investment. However, in order to maintain a secure supply and support highly polluting power stations coming offline, a capacity market has been introduced in the UK which sets a price for available capacity, not just generating capacity. The main aim of the capacity market is to ensure the security of supply in the coming years. Generators and consumers with demand management processes can bid for capacity provision 4 years in advance, receiving payments for capacity provided and incurring charges if capacity is unavailable [109]. This will reward availability of supply when the system is heavily loaded and penalise capacity that is not available when needed. Generators, consumers who can manage their demand and power storage assets can all participate. The first capacity auctions were held at the end of 2014 for capacity delivery in 2018/2019 [118] and subsequent auctions have been held annually the results of which are summarised in Table 2-12.

Table 2-12: Details of T4 capacity auctions held from 2014 - 2016

Auction held	December 2014 [119]	December 2015 [118]	December 2016 [120]
Year to be delivered	2018/2019	2019/2020	2020/2021
Total capacity agreements	49.2 GW	46.3 GW	52.4 GW
Clearing price	£19.40 / kW / year	£18.00 / kW/ year	£22.50 / kW/ year

Whilst the UK government is optimistic about the ability of this mechanism to deliver a secure electricity supply, others have indicated that the uncertainty surrounding this new concept may stagnate investment as investors seek stability and assurance before committing to projects. However, it will hopefully encourage flexibility in the electricity grid which is beneficial for the integration of renewable generation [98]. A significant problem with this mechanism is that it encourages and continues to pay for fossil fuel based generators, which does not aid the aim of decarbonising the electricity supply. For example for electricity capacity procured in 2020/2021 over 50% will come from gas turbines of different types. This encourages a reliance on fossil fuel generators for many years to come. The capacity market has the potential to offer a revenue stream to electricity storage operators, which may support the integration of variable renewable electricity. However, so far very few storage projects have been successful in the capacity market, with less than 6% of the capacity procured coming from storage projects in the auctions held so far [118-120].

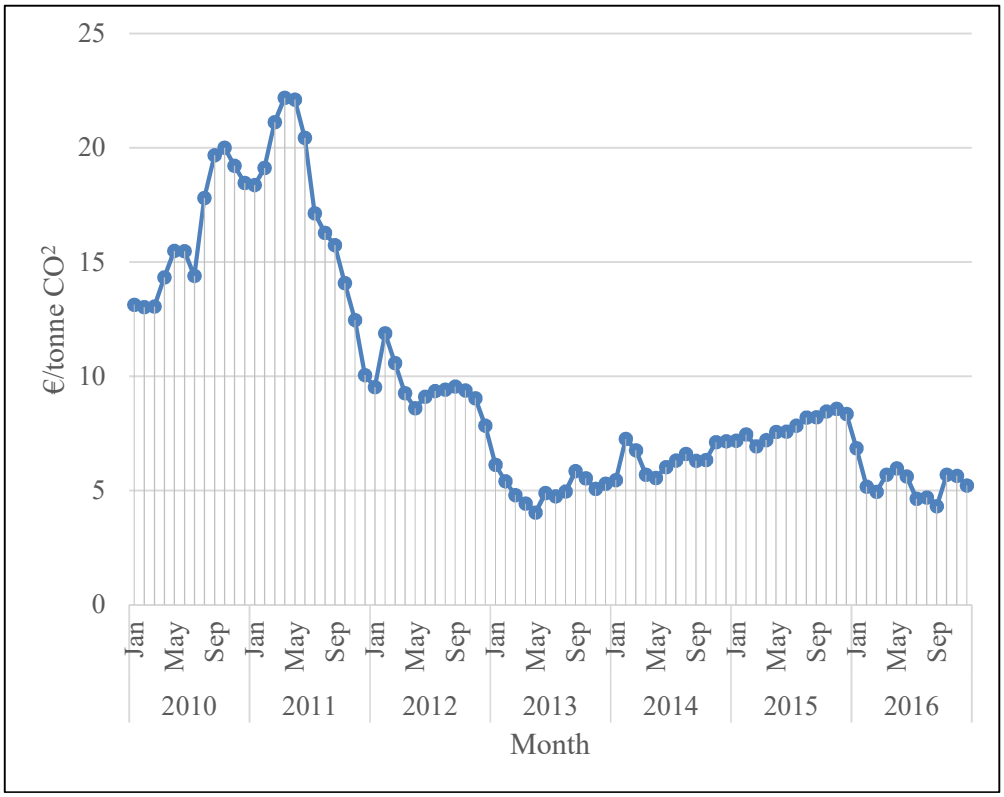
2.3.3.3 Carbon Price Floor

Electricity generation companies in the EU are required to buy permits to emit greenhouse gases under the EU emissions trading scheme (ETS). The idea of this was to give a financial incentive to reduce the use of cheap polluting fuels and encourage investment in low carbon technologies. This market has been volatile and in April 2013 the daily price fell to €3.13/tonne of CO_2 [121]. As can be seen in Figure 2-14, the monthly average price around this time remained low, at around €4/tonne of CO_2. Despite an increase in prices between 2013 and 2015, prices fell again in 2016 with the monthly average price remaining at around €5/tonne CO_2. The low prices have largely been attributed to an excess of permits being available and has led to an inadequate incentive to reduce emissions [122].

The carbon floor price was introduced in the UK in April 2013 [122]. The carbon floor price meant that there would be a minimum payment for emissions, and if the

EU ETS allowance price fell below this companies in the UK would be required to make top-up payment to the UK government, known as the carbon price support. This was introduced in 2013 and was proposed to steadily rise from £15.70/ tonne CO_2 in 2013 to £30/ tonne CO_2 in 2020. The carbon floor price was levied via the existing climate change levy. The problem with this is that some argued the suppliers will simply pass on increased costs to consumers, maintaining profits with no need to invest in low-carbon technologies or drive out high emission businesses to other countries.

Figure 2-14: Average Monthly Price of EU ETS from 2010 to 2016 [121]



The actual price for the carbon price support was set at £4.94/tonne CO_2 for 2013, £9.55/tonne CO_2 for 2014 and £18.08/tonne CO_2 for 2015 [122]. However, in the 2014 budget the carbon price support was frozen at £18.08/tonne CO_2 until 2020 to avoid excessive cost to business and rises in electricity prices for customers [123].

This price freeze was extended to 2021 in the 2016 budget [124]. Whilst the price has not increased as originally planned this mechanism has contributed to a reduction in the amount of electricity generated from coal.

2.3.3.4 Emissions Performance Standard

The emission performance standard (EPS) introduced in February 2014 sets an upper limit on emissions from power plants. It is applied to new power plants over 50 MW and the replacement of any main boiler after February 2014. This prevents coal-fired power stations without carbon capture and storage (CCS) technology being built. It is set at 450g CO_2/kWh and fixed until 2044 [109]. In a year a power plant must not exceed the total volume of CO_2 emissions if the plant were operated at 85% of full capacity for the whole year. For example, a plant with a capacity of 60 MW operating at 85% capacity over the year would generate $60 \times 24 \times 365 \times 0.85 = 446{,}760$ MWh of electricity. Hence the limit on its CO_2 emissions would be 201,042 tonnes/year. The aim of this is to maintain the use of some fossil fuel plant to provide flexible base load supply whilst encouraging investment in CCS and biomass plants. Any plant with CCS installed will be exempt from the EPS until 31 December 2027 [109]. The implementation of the EPS could have an impact on wind energy. For example, if this is successful as a policy, some of the electricity previously supplied by coal is likely to be replaced by other sources of electricity. Whilst some of this may come from renewable sources, some may be replaced by gas power plants. Gas power plants, particularly of modern designs, are more flexible than coal power plants resulting in an electricity system which may be able to accommodate wind power more easily [97]. This is the trend which has been observed in the UK between 2012 and 2017, with an increase in electricity generated from gas and a decrease in electricity generated from coal [16, 125].

2.3.4 Uses of wind power forecasting

There are various uses for wind forecasting for example turbine control, integration in the grid or turbine maintenance planning. The optimal time scale of a forecast

depends on its intended use. Forecasts from milliseconds to seconds ahead are used for turbine controls. For integration in the electricity grid, longer predictions are important. Forecasts from a few minutes up to an hour are used to support balancing mechanisms, whilst forecasts from 1 hour to a few days ahead are used to optimise dispatch of conventional power and allocate reserve power [24]. Forecasts from a few days to a few months ahead are used to schedule maintenance activities, which is of particular importance with the increased developments in offshore power as the conditions for maintenance may be hazardous. An overview of applications of wind power forecasting is given by Soman et al. [44] and summarised in Table 2-13.

Table 2-13: Applications of different forecasting horizons [44]

Range	Use
Few seconds to 30 minutes ahead	-Electricity market clearing -Regulation actions
30 minutes to 6 hours ahead	- Economic load dispatch planning - Load increment/decrement decisions
6 hours to 1 day ahead	- Use of generators/ backup power - Participation in day-ahead electricity market
1 day to 1 week or more	- Unit commitment decisions - Reserve requirement decisions - Maintenance scheduling to reduce operating costs

The forecasts developed in this thesis are for between 3 and 72 hours ahead. Therefore the forecasts are of most relevance for unit commitment decisions, setting reserve requirements and participation in electricity markets. Alstrom and Zavadil [126] explored the role of wind power forecasting in making unit commitment decisions. They tested the impact of wind on an area in the USA with peak demand of 9000 MW and wind capacity of 300 MW estimating there was a cost of $4.37/MWh to provide electricity contracted but not produced. The authors concluded that this could be reduced through the use of wind forecasting, however

they did not assess the cost benefit which could be achieved through better forecasting. Doherty and O'Malley [127] used wind power forecasting to quantify the level of reserve required in the Irish electricity system, concluding that increasing wind penetration would require higher levels of reserve but decreased wind forecast errors decreased the reserve requirements. In this paper the authors consider the correlation between standard deviation of wind power forecast error and reserve requirement. This is not an error metric used in many studies, making it difficult to compare the costs estimated in this paper with the forecast accuracy achieved by state of the art forecasts. Lowery and O'Malley [128] also considered the impact of wind forecasting error on unit commitment decisions and how forecasting can be used to reduce system costs.

Forecasts can enable wind energy to become cost-competitive with other, more established, generation methods [67]. This is particularly important given that onshore wind energy is considered an established technology and so does not qualify to bid for CFDs under the electricity market reforms. Barthelmie et al. [23] looked at the benefits of using short-term forecasting to participate in electricity markets from an economic point of view. The paper created a simulation of 2003 electricity prices for different forecasting strategies. The authors concluded that optimal price per kWh of electricity is achieved when the smallest possible discrepancy between the forecast and produced power and the optimal trading window in the market are obtained. The authors also noted that ideally wind energy would be accurately predicted and sold at the highest trade price. Failing that, it is preferable to under predict output, rather than over predict and risk being charged the current system buy price. The difference between the system buy price and the system sell price is used in this study to quantify the benefit of an accurate forecast but this difference is variable. Using the method outlined by Barthlemie et al. [23] if a perfect forecast is achieved, the average benefit of the forecast for 2003 – 2007 was estimated to be £4.50/MWh. This estimate just considers the price of electricity and does not take into account the traded price of a ROC, which also makes up some of the income for wind energy. The work presented by Barthelmie et al [23] looks at several scenarios: no forecast being used, a simple mean annual wind speed, a perfect forecast, and a

simulation based on time series data. The results show that for reasonable wind speed forecast errors, wind speed forecasts are economically valuable. It suggests that once the systematic error in forecasted winds speeds exceeds +/- 10% the benefit of a forecast is close to zero. It aims to understand the size of the wind farm for which short-term forecasting becomes economically viable. The work done in this paper does not directly evaluate the impact of random error (MAE) on price received for electricity generating and only concludes that systematic error (forecast bias) has most impact on prices. The prices used for this study are from 2003 under a dual pricing system and the same conclusions may no longer hold in the current UK electricity markets. As stated in Section 2.2.4 bias is rarely discussed in literature therefore it is difficult to assess whether current state of the art forecasting techniques achieve the forecast accuracy required to be economically successful.

As seen in the work performed by Barthelmie et al. [23] accurate forecasting allows operators to trade electricity for the maximum possible income. Further to this, forecasting reduces the need for balancing power and reserve load, which in turn leads to lower integration costs for wind, lower emissions from power plants used for balancing and subsequently a higher value for wind energy [20]. If no wind forecasting were performed, then wind energy would require significant balancing.

In addition to the use of deterministic wind power forecasts, incorporating uncertainty into the operational management of wind farms through the use of probabilistic forecasting can be beneficial over the simple use of deterministic forecasts. Probabilistic forecasts have been used to define reserve levels, determine electricity storage system size and determine an optimum bidding strategy. Some examples of the uses of probabilistic wind power forecasts are discussed here.

Matos and Bessa [129] used probabilistic forecasting to develop a reserve management tool which enables system operators to define reserve needs for day-ahead and intraday markets. The tool aimed to balance risk and cost. In this the benefit is measured in terms of the reduction in the expected energy not supplied. The trade-off is then between reducing the expected energy not supplied whilst maintaining acceptable costs. The acceptable costs in this case can be defined by the

user. Bludszuweit et al. [130] proposed a method for energy storage system operation using probabilistic forecasting. It aimed to reduce uncertainty in wind power delivery up to 48 hours ahead. The method optimised the size of an energy storage system given a defined acceptable level of uncertainty in wind power delivery. The method allowed the optimum size of an energy storage system to be based on the cost-benefit analysis of the cost of storage compared to the benefit of reducing unserved energy. Botterud et al. [131] looked at how probabilistic wind power forecasting can be used to determine reserve requirements and how these change with uncertainty. They also discussed how demand dispatch and wind power forecasting should be factored into unit commitment and economic dispatch decisions. A cases study of the electricity market in Illinois, USA was used to show how more efficient scheduling of resources can be achieved by setting reserves based on uncertainty rather than fixing levels. This, in turn, reduced operating costs. Zhou et al. [132] also considered the use of a probabilistic wind power forecast to schedule operating reserve in a market with day-ahead and real-time trading. This again showed that dynamic operating reserve setting is more efficient than using fixed reserve setting. However neither Botterud et al. [131] nor Zhou et al. [132] discuss how the accuracy of a power forecasts affects operating decisions.

In addition to the use of probabilistic forecasts for defining reserve requirements, a number of authors have looked at operational strategies for trading wind energy in a liberalised energy market using probabilistic forecasts. Zugno et al. [133], Botterud et al. [134] and Pinson et al. [135] presented optimal bidding strategies for wind power producers participating in electricity markets making use of wind power forecasts. Zugno et al. [133] used probabilistic forecasts of both wind power and market prices to develop a bidding strategy which optimises market revenue. The methods developed were tested on the Eastern Danish price area of the Nordic Power Exchange for 10 months of 2008. They concluded that the use of a bidding strategy alongside forecasts lead to increased revenue through a decrease in imbalance costs of 2.3%. Botterud et al. [134] proposed a methodology for optimal day-ahead bidding again concluding that the use of wind power forecasts led to increased profit and decreased financial risk. This was due to the avoidance of

penalties for deviating from amount of electricity contracted. Pinson et al. [135] used a probabilistic wind power forecast to develop an optimal bidding strategy. When tested on a Dutch wind farm this lead to an average reduction in cost of regulation of €2.20 - €5.90 per MWh. Robu et al. [136] introduced the idea of a payment mechanism based on the use of probabilistic forecasts of power production from small renewable energy sources. The aim was to present a method for small renewable energy cooperatives to estimate a confidence interval on power output estimates to aid grid planning, hence aiding integration of renewable energy in the grid.

As discussed in Chapter 1, moving towards a low carbon electricity system is crucial to avoiding increasing emissions. Many different scenarios have been outlined in literature for how a low carbon electricity system might be achieved. In Table 1-1 one such proposal of how sufficient electricity might be generated from renewable sources is outlined. In this, as in most low carbon electricity scenarios, a high proportion of variable or intermittent renewable energy is anticipated. In addition to this, some level of demand side management of electricity is likely to be required. As we move towards a system with a greater level of renewable electricity, all the applications of wind power forecasting mentioned so far are likely to become increasingly important. Firstly, with more renewable electricity the need to predict power output in advance will increase to allow efficient planning of power production in order to meet demand. In addition to this, in future low carbon energy systems predicting periods of high power production in advance will allow some electricity consumption to be shifted to utilise power production fully. For example, batteries or electric vehicles could be charged during periods of high production. Similarly, consumption could potentially be reduced during periods of low power. However, in order to do this and avoid disruption to consumers predictions will need to be accurate. These techniques allowing effective use of resources will be essential in ensuring that sufficient power can be generated from renewable resources and forecasting can enable this.

The applications and literature shown here indicate that there is value in forecasting power for many different users in a variety of applications. The value can be evaluated in many ways, however multiple studies conclude that more accurate forecasting is beneficial to the user. Few authors discuss the impact of predictive accuracy on the value added by a forecast. However, there is some indication that more accurate forecasting increases the value added by a forecast. Because of this, continuing research into new methods for forecasting wind power output is both interesting and important.

The literature available which discusses applications of wind power forecasting indicates that there is value in improving wind power forecasts. It also indicates that probabilistic wind power forecasts can be effectively used for both grid management and participation in electricity markets. The potential applications considered above indicate that wind power forecasting can provide value for a number of users. The value of wind power forecasting and the appropriate way to quantify this is discussed further in Chapter 6. In Chapter 6 the impact of forecast accuracy on value is also considered. Within this, the impact of the changes to the UK electricity markets discussed in Section 2.3.2 on the value of a wind forecast are considered. Thus far, there is limited literature which assesses the impact of these changes to the electricity market on renewable energy generators. This is a significant gap in literature which this thesis aims to address.

Chapter 3. Near surface wind speed prediction using a hybrid NWP and Gaussian process regression model.

3.1 Overview

As outlined in Section 2.3 wind power forecasts can be valuable to a number of users. Literature outlined in Section 2.2 suggests that for short to medium range forecasts hybrid statistical and NWP methods are likely to provide the best results. This is because the combination of a statistical and NWP approach enables the NWP predictions to be corrected to provide better predictions for specific locations. Furthermore, the use of the NWP model allows predictions to be made further in advance than would be possible with a purely statistical model. Because of this, throughout this thesis, a hybrid method for predicting wind power output is developed and evaluated. To make an initial assessment of the suitability of the method in the context of wind energy this chapter focuses on wind speed predictions. This chapter presents a hybrid numerical weather prediction model (NWP) and a Gaussian process regression (GPR) model for near-surface wind speed prediction up to 72 hours ahead. It also considers whether subdividing data by atmospheric stability class can aid model performance. In Section 2.1 atmospheric stability was introduced and its influence on wind speed profiles and the conversion of wind to power is discussed. Because of the influence of atmospheric stability on the key components of wind speed prediction, one objective of this chapter is to assess whether sub-dividing data by atmospheric stability class can improve model performance. In this chapter stability conditions are estimated based on observations at the time of prediction, using the Pasquill-Gifford-Turner classification scheme given in Section 3.2.5. This shows the potential of the method using commonly

available data. However, the method is not technically a forecast as it relies on observations of stability conditions rather than predicted stability conditions. The results shown in this chapter indicate motivation for developing this method further, based on data obtained from publicly available resources. In Chapter 4 this is developed further into a predictive model using additional data obtained from the Met Office.

In this chapter the NWP performance is discussed, giving motivation for the consideration of atmospheric stability in wind speed prediction. Next, the model is introduced and the results are shown for a selection of 15 weather observation sites across the UK. The model is also tested for the prediction of hub height wind speeds for one turbine in the UK. At this site, predicted wind speed is compared to measured wind speed. Finally, the impact of improved wind speed forecast on power forecasting is considered.

A three hourly wind speed forecast from an NWP provided by the Met Office was corrected using a GPR model, where the data was subdivided using the atmospheric stability class calculated using the Pasquill-Gifford-Turner method based on observations at the time of prediction. The method was validated using data from 15 UK MIDAS (Met Office Integrated Data Archive System) sites with a 9-month training and 3 month test period. Results are also shown for hub height wind speed prediction at one turbine. Additionally, power output is predicted for this turbine by translating hub height wind speed to power using a turbine power curve. While various forecasting methods exist, limited methods consider the impact of atmospheric stability on prediction accuracy. Outputs show the GPR model improves forecast accuracy over the original NWP data and subdividing data by atmospheric stability class further reduces prediction errors. Comparing predicted power output to measured output reveals power predictions are also enhanced, demonstrating the potential of this novel wind speed prediction technique.

The main objectives of this chapter are:

(i) To develop a hybrid numerical weather prediction model (NWP) and Gaussian process regression (GPR) model for near surface wind speed prediction up to 72 hours ahead.
(ii) To investigate the impact of using data partitioned on atmospheric stability class on model performance.
(iii) To evaluate the model performance for wind speed prediction at 15 MIDAS sites across the UK.
(iv) To evaluate the model performance for prediction of hub height wind speed at one wind turbine site in the UK.

This chapter will be structured as follows: Section 3.2.1 gives details of the data used in this section to train and evaluate the model. Section 3.2.2 gives the definition of atmospheric stability used in the current work, its potential role in wind forecasting, and methods for calculation. Sections 3.2.4 to 3.2.6 introduce the GPR model, giving an overview of the mathematical concepts. Section 3.2.7 gives details of the model formulation. Section 3.3 presents the results for prediction of 10 m wind speed at 15 MIDAS sites, hub height wind speed at one site and the significance of these results in power forecasting for a single turbine. Section 3.4 gives conclusions and outlines how this work will be continued in the remainder of this thesis.

3.2 Methodology

3.2.1 Introduction to Gaussian process regression

Regression analysis is the process of finding a relationship between a variable and a number of independent variables or predictors. This is achieved by estimating the relationship between observed values of the dependent and independent variables. This may then be used to estimate future values of the dependent variable, given values for the predictor variables. Given n observations of a dependent variable $\mathbf{y} = [y_1, \dots, y_n]$ a relationship is estimated by considering $f(\mathbf{X})$, the underlying function

of $\mathbf{X} = [\mathbf{x_1}, \ldots, \mathbf{x_n}]$ where $\mathbf{x_i}$ is a vector of predictor variables for observation $i \in [0, n]$. To predict the values of a dependent variable based on new input values a training dataset is used develop a general relationship between input and output variables [137].

This type of analysis is known as supervised learning. Traditionally this is done using a parametric model, where the relationship is defined and parameters learnt from data. Once parameters have been fit this can be used to make predictions [138]. One example is linear regression, where the function is assumed to be linear and the parameters defining the relationship can generally be calculated analytically. Functions could also be assumed to be polynomial, exponential or any other type. This type of regression can suffer due to the process of choosing the class of functions to consider. If the underlying function, *f*, is not modelled well by the class chosen, predictions made will be inferior. Increasing the number of variables included to attempt to give a good prediction can lead to overfitting, where the chosen function fits the training data well but is a poor generalisation of the function. This leads to substandard predictions.

An alternative solution is provided by Gaussian process regression. GPR is a supervised learning method where an input-output mapping is learnt from empirical data [137]. It is a non-parametric Bayesian regression technique which does not initially restrict the relationship between the target and input variables to a specific form. This allows a flexible model. Prior knowledge is combined with observed data to determine posterior predictive distributions for future inputs.

Rather than defining a relationship and learning parameters, Gaussian process regression defines a probability of every function possible and evaluates the likelihood of these by updating prior knowledge of the probabilities given some observed data. This is done by defining a prior probability distribution over all functions possible [137]. This is a non-parametric technique as it involves attempting to infer how measured data is correlated rather than fitting parameters in a model [138]. A Gaussian process is a collection of random variables, any finite

number of which have a joint Gaussian distribution [137]. Gaussian processes can be considered as a generalisation of the Gaussian probability distribution to infinitely many variables [138]. The advantage of GPR is its ability to provide uncertainty estimates and to learn parameters from data [138]. Rasmussen and Williams [137] provide an extensive mathematical background of GPR, of which a summary is provided here. The Gaussian process $f(\mathbf{x})$ is completely specified by its mean and covariance function,

$$f(\mathbf{x}) \sim GP(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}'))$$

This Gaussian process is used as a prior for Bayesian inference which is updated to reflect the training data available. Different mean and covariance functions can be used to specify some properties of the functions, these are discussed further in Section 3.2.2.

GPR aims to identify a relationship between input variables and target variables, based on the observational data available. Assume a set of measured values, $\mathbf{y} = [y_1, \dots, y_n]$ are observed at points $\mathbf{X} = [\mathbf{x_1}, \dots, \mathbf{x_n}]$. These measurements represent noisy observations of a function, f, which is affected by noise, $\varepsilon \sim N(0, \sigma_n)$. The aim is to discover the underlying function that satisfies $y_i = f(\mathbf{x}_i) + \varepsilon_i$ where $y_i \in \mathbf{y}$ is the target variable, $\mathbf{x}_i \in \mathbf{X}$ are the input variables and ε is normally distributed additive noise. This is used to define a distribution over functions which can be updated using training data. The prior distribution is the initial specification of the distribution which gives information on the mean and covariance functions used. Being a linear combination of Gaussian variables, $\mathbf{y}$ is also Gaussian, with distribution $\mathbf{y} \sim N(m(\mathbf{x}), K(\mathbf{X}, \mathbf{X}) + \sigma^2\mathbf{I}$ where $K_{\mathrm{i,j}} = k(\mathbf{x_i}, \mathbf{x_j})$, is the covariance matrix calculated from the covariance function. The covariance is a crucial part of the model specification, as it includes assumptions about the functional relationship. Despite this, establishing the correct covariance function for a regression problem is a significant issue in the inference process. The covariance function is discussed in more detail in Section 3.2.2.

Given the training set $\mathbf{D} = (\mathbf{X}, \mathbf{y})$, and a new set of inputs $\mathbf{X}_*$ the aim is to estimate output values $\mathbf{f}_*$ where $\mathbf{f}_* = f(\mathbf{X}_*)$. $\mathbf{y}_*$**,** the target variables are the noisy observed values of the underlying function, $\mathbf{f}_*$. The set $\mathbf{D}_* = (\mathbf{X}_*, \mathbf{y}_*)$ of new input and target variables is known as the test set. In the Bayesian framework, this is done by calculating the conditional distribution of $\mathbf{f}_*$ given $\mathbf{y}$, $\mathbf{X}$ and $\mathbf{X}_*$. Alternatively, this can be thought of as conditioning the prior distribution on the observations. This conditional distribution is referred to as the posterior distribution. Bayes rule can be used to calculate this conditional distribution, as shown in Equation 3-1.

$$p(\mathbf{f}_*|\boldsymbol{y}, \mathbf{X}, \mathbf{X}_*) = \frac{p(\mathbf{y}|\mathbf{f}, \mathbf{X})p(\mathbf{f}|\mathbf{X})}{p(\mathbf{y}|\mathbf{X})} \quad \text{Equation 3-1}$$

$$\text{posterior} = \frac{\text{likelihood} \times \text{prior}}{\text{marginal likelihood}}$$

The prior distribution incorporates assumptions about the model before evidence from the data is taken into account. The likelihood allows assumptions about the underlying function. In this chapter, the likelihood is assumed to be Gaussian as this allows exact inference. However, the choice of likelihood function is explored further in Chapter 5. The joint distribution of the training data and the predicted output is Gaussian, given by Equation 3-2 [137].

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{f}_* \end{bmatrix} \sim N\left(\begin{matrix} m(\mathbf{X}) \\ m(\mathbf{X}_*)' \end{matrix} \begin{bmatrix} K(\mathbf{X}, \mathbf{X}) + \sigma^2 \mathbf{I} & K(\mathbf{X}, \mathbf{X}_*) \\ K(\mathbf{X}_*, \mathbf{X}) & K(\mathbf{X}_*, \mathbf{X}_*) \end{bmatrix} \right) \quad \text{Equation 3-2}$$

The principle of joint Gaussian distributions allows the prediction results for the target to be inferred from the posterior distribution given by Equation 3-1, leading to equations for the mean and covariance of $\mathbf{f}_*$ [137] given by Equation 3-3 and Equation 3-4.

$$\bar{\mathbf{f}}_* = E[p(\mathbf{f}_*|\mathbf{X}, \mathbf{y}, \mathbf{X}_*)] = K(\mathbf{X}_*, \mathbf{X})[K(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbf{I}]^{-1}\mathbf{y} \quad \text{Equation 3-3}$$

$$cov(\mathbf{f}_*) = K(\mathbf{X}_*, \mathbf{X}_*) - K(\mathbf{X}_*, \mathbf{X})[K(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbf{I}]^{-1} K(\mathbf{X}, \mathbf{X}_*) \quad \text{Equation 3-4}$$

GPR has been used for prediction in a number of applications. For example, spectroscopic calibration [139], robot control [140] and image processing [141]. In

the context of wind energy applications, a method which uses GPR to assess resource availability for small-scale wind energy projects was developed by Weekes [142]. In this context GPR was used in a measure-correlate-predict method to estimate wind resource at potential locations for wind turbines. Through these applications, GPR has shown an ability to predict well in situations where complex nonlinear relationships exist between variables. Because of this, it is a good method for wind speed prediction, given the typically complex patterns and relationships between wind and other weather variables. Chen et al. [73] describe a method in which an NWP model is combined with a GPR model to predict wind speeds up to 1 day ahead. The corrected wind speeds are used to predict wind power using another GPR model. In this example, three data sets from different wind farms in China are used to validate the method, reporting reductions in mean absolute error compared to an Artificial Neural Network (ANN) model. In a different study Chen et al. [71] present the potential for a composite wavelet analysis and GPR forecasting technique. Small improvements over a simple GPR model were noted, demonstrating that the concept merits further investigation. Zhang et al. [143] combine an autoregressive model with GPR for probabilistic wind speed forecasting. The model was used to predict mean hourly wind speed one hour ahead for wind speeds at 3 wind farms in China. Furthermore, Hu et al. [144] combine empirical wavelet transform, partial auto correlation function and GPR to predict wind speeds at one wind farm in China. The results are shown for both half hourly wind speed prediction (up to 2 hours ahead) and hourly wind speed prediction (up to 4 hours ahead). These studies show some potential for GPR models to predict wind speed and power output well. However, they focus on short term predictions, mostly up to a few hours ahead. The model presented in this chapter focuses on wind speed forecasts further in advanced, presenting a hybrid NWP and GPR model for wind speed predictions up to 72 hours ahead. In addition, the impact of subdividing data by atmospheric stability on model performance is discussed.

3.2.2 Choosing mean and covariance functions

The covariance function incorporates important information about the model. It is described informally by Rasmussen and Williams [137] as the similarity between data points. GPR assumes that points which are 'close' in the input space are likely to have similar target values. This means that for the prediction of a point in the test set, training points with input points which are close to the test point will inform the prediction. How points which are close are related and how this relationship changes with increased distance is defined mathematically by the covariance function. The covariance between points x_1, x_2 is given by a function $k(x_1, x_2)$. The covariance matrix is given by **K** where the element in position (i, j) is given by $k_{i,j} = k(x_i, x_j)$. The function k can be any function which gives a positive semi-definite symmetric covariance matrix **K**. Many structures can be incorporated into a covariance function for example periodicity, interactions between variables and symmetry [137]. The squared exponential covariance function, defined by Equation 3-5 is a smooth, infinitely differentiable function which is commonly used in GPR models. It is a universal function, thus can be used to approximate any continuous function given enough training data [145, 146]. It is a very flexible and adaptive covariance function. Because of this, the squared exponential covariance function is used throughout this thesis.

$$k_{SE}(x_1, x_2) = \sigma^2 exp\left[-\frac{1}{2}\left(\frac{x_1 - x_2}{l}\right)^2\right] \quad \text{Equation 3-5}$$

where σ, l are the noise and lengthscale parameters respectively. Another property of covariance functions is that they can be combined using addition or multiplication. Multiplying or adding positive definite covariance functions will always result in a positive definite function. This property is used by Duvenaud [145] to define covariance functions which can be used for models in multiple dimensions. This is used to give a covariance function with a different lengthscale parameter for every input variable in the model. For example, for a model with D input variables the covariance function is given by Equation 3-6.

$$k_{SE,D}(\mathbf{x},\mathbf{x}') = \prod_{d=1}^{D}\sigma_d^2 exp\left[-\frac{1}{2}\left(\frac{x_d - x_d{}'}{l_d}\right)^2\right] = \sigma_f^2 exp\left[-\frac{1}{2}\sum_{d=1}^{D}\frac{(x_d - x_d')^2}{l_d^2}\right] \quad \text{Equation 3-6}$$

The set of length scale parameters $\mathbf{l} = \{l_1, \dots, l_D\}$ allows the relevance of each variable to be determined separately. This is known as automatic relevance determination (ARD). If the lengthscale parameter of a particular variable is small, then the parameter has a large impact on the predicted output and is therefore highly relevant. If the parameter is large it has little impact. This allows the most relevant variables to be given a higher weighting in the model.

The mean function incorporates less information about the model and generally receives less attention in the literature. Usually, a simple mean function is employed, either assuming the mean function is equal to zero everywhere or using a linear mean function. The covariance function allows sufficient expressivity within the model for the mean function to be less important [145]. However, it can incorporate assumptions about model prediction at points which are far away from training data. Whilst making predictions based on extrapolation far from the training set should generally be avoided, the mean function can add information from the modeller's expertise about areas of the input space which have limited training data. In this chapter, a linear mean function with a constant term is used, given by Equation 3-7.

$$m(\mathbf{x}) = b + \sum_{d=1}^{D} a_d x_d \quad \text{Equation 3-7}$$

where the hyperparameters $b, a_1, \dots, a_D$ are set in the training phases.

3.2.3 Learning the hyperparameters

To enable flexible modelling, there are a number of free parameters in the mean, covariance and likelihood functions. These are known as hyperparameters. Usually, the value of these is not known in advance. Because of this, the set of

hyperparameters, $\boldsymbol{\theta}$, is considered as a random variable [147] and through the model training phase the probability of the hyperparameters given the data is considered. This is known as the a posteriori hyperparameter likelihood. In order to find the optimum value for the hyperparameters, a maximum of Equation 3-8 is sought. Using Bayes rule Equation 3-9 is obtained.

$$p(\boldsymbol{\theta}|\mathbf{y},\mathbf{X}) \qquad \text{Equation 3-8}$$

$$p(\boldsymbol{\theta}|\mathbf{y},\mathbf{X}) = \frac{p(\mathbf{y}|\boldsymbol{\theta},\mathbf{X})p(\boldsymbol{\theta}|\mathbf{X})}{p(\mathbf{y}|\mathbf{X})} \qquad \text{Equation 3-9}$$

If no prior knowledge of the hyperparameters is incorporated into the model, $p(\boldsymbol{\theta}|\mathbf{X})$is constant. $p(\mathbf{y}|\mathbf{X})$, is also constant with respect to the hyperparameters therefore the a posteriori hyperparameter likelihood is proportional to $p(\mathbf{y}|\boldsymbol{\theta},\mathbf{X})$. The optimal values for $\boldsymbol{\theta}$ can then be obtained by considering the likelihood of $\mathbf{y}$ given the input data $\mathbf{X}$. Hence finding the optimum values for the hyperparameters is equivalent to maximising the marginal likelihood of the target variables given the hyperparameters. For ease of mathematical manipulation this is done using an equivalent optimisation problem of minimising the negative log marginal likelihood. When the Gaussian likelihood function is used, the log marginal likelihood function can be inferred exactly. Given these conditions the log marginal likelihood can be calculated using Equation 3-10.

$$\log p(\mathbf{y}|\boldsymbol{\theta},\mathbf{X}) = -\frac{1}{2}\mathbf{y}^T(\mathbf{K}+\sigma_n^2\mathbf{I})^{-1}\mathbf{y} - \frac{1}{2}log|\mathbf{K}+\sigma_n^2\mathbf{I}| - \frac{n}{2}\log(2\pi) \qquad \text{Equation 3-10}$$

The terms in the log marginal likelihood enable model fitting whilst avoiding overfitting. Overfitting occurs when a model is parameterised to ensure agreement with measured values in such a way which impedes generalisation to an unseen data set. The second term in the log marginal likelihood function adds a penalty as the model complexity increases. This is similar to the role of a regularisation term in parametric regression techniques. This is balanced by the first term which reduces with improved model fit.

This minimisation cannot be solved analytically therefore numerical methods must be used [137]. Throughout this thesis the GPML toolbox [146] is used. In this, a minimisation function is included which is used to optimise the hyperparameters. The function uses a Polack-Ribiere conjugate gradient method which is a non-linear optimisation technique. It implements a line search using quadratic and cubic approximations. Furthermore, Wolfe-Powell stopping criteria and the slope ratio method are used to calculate the initial step size. An in depth discussion of these techniques is given in [148]. Whilst the optimisation of marginal likelihood functions allows the most appropriate values for hyperparameters to be chosen, there are some difficulties with implementation which need to be considered. Optimisation requires initial values for the hyperparameters to be chosen. A common problem is that the marginal likelihood may not have a single global optima with respect to the hyperparameters [149]. Chen et al. [149] investigate the sensitivity of the hyperparameter estimation on the choice of initial values and the influence this has on GPR predictability. The authors conclude that the prior choice of hyperparameters has little influence on the performance of the GPR model, particularly when a squared exponential covariance function is used, as it is here.

3.2.4 NWP prediction and meteorological observations

Wind speeds predicted using an NWP model and observed wind speeds are used as inputs in the GPR model. A detailed discussion of the model inputs is given in Section 3.2.7. Before this, the data sets which will be used are discussed.

The predicted wind speeds which are used to inform the hybrid model are taken from an NWP model developed by the Met Office. This NWP model provides three hourly forecasts up to 5 days in advance, employing a global forecast model to predict longer range weather forecasts (48+ hours ahead) combined with a mesoscale model to generate a more accurate short-range forecast. The forecast data used in this work are a weighted combination of the Met Office UKV and Euro4 models. UKV is a variable resolution deterministic model, with a resolution of 1.5 km over the UK and decreased resolution at the model boundaries to aid integration

in a nested model [150]. Euro4 is a 4 km resolution deterministic model covering Europe. UKV runs up to 36 hours in advance and Euro4 up to 120 hours. The forecast data is available from the UK governmental public data website [151] for over 6000 sites. The meteorological observations which have been used for reference have been taken from the Met Office Integrated Data Archive System (MIDAS), available from the British Atmospheric Data Centre (BADC) [152]. The archive consists of UK land surface observations, global marine observations, and a selection of radiosonde observations both in the UK and at international stations operated by the Met Office. This data provides hourly observations of a selection of meteorological variables including wind speed and direction, cloud cover, temperature, air pressure and humidity amongst others. The data provided including units is shown in Table 3-1. The MIDAS observation stations are set up so that the observation data can be the best quality possible with details given by the BADC [152]. Cup anemometers are used to measure wind speed, at a height of 10 m above ground level. The site must be free from obstructions to avoid measurements in the wake of obstructions and quality control is performed to avoid inclusion of spurious data where possible. For example, automatic algorithms are applied to ensure consistency of wind measurements with other local stations. From the available data, nearly 300 sites were identified where both observational and forecast data was available. Not all weather variables are available at every MIDAS site and data coverage is variable, dependent on factors such as equipment failure. Subsequently, from the available sites, sites were considered where 80% data coverage was available. Furthermore, certain meteorological variables are not available at all sites. Due to the variables required to estimate stability conditions only sites where information on cloud coverage was available were considered.

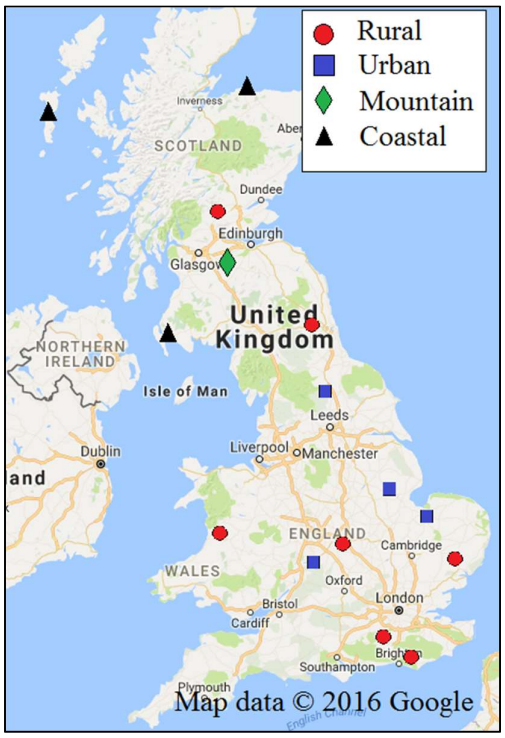
Table 3-1: Met Office forecast variables

Variable	Description or unit
Wind speed	Miles per hour (mph)
Wind Direction	16 point compass
Temperature	Degrees Celsius (^{O}C)
Significant weather	Description of weather i.e. sunny, mist, fog, light rain etc.
Relative humidity	Percentage (%)
Visibility range	Given within a range of km
Visibility description	Poor to excellent
Wind gust	Miles per hour (mph)
Feels like temperature	Degrees Celsius (^{O}C)
UV Index	Index of 1-11. In the UK this does not exceed 8.
Precipitation probability	Percentage (%)

From this reduced set of potential sites, a sample of 15 sites across the UK was investigated. A map of these is shown in Figure 3-1. The MIDAS datasets are taken from various locations across the UK, with different weather conditions and site characteristics across the selection. The sites were categorised into 4 types; rural, urban, mountain and coastal. Categories were chosen for the sites based on visual inspection of the site itself and the local area, considering the proximity to the coastline, building density, elevation and terrain complexity. The model performance was considered within the different categories as well as overall. The locations and classification of these sites are shown in Figure 3-1.

To demonstrate the potential for wind power prediction, the model was also tested for one location in the UK where hub height wind speed and power were available. The data comprised of measured wind speed data at approximately 65 m above ground level and power output from a 1.5 MW turbine in a suburban location.

Figure 3-1: Map of MIDAS sites used, including site classification



3.2.5 Treatment of atmospheric stability

In addition to the performance of the hybrid NWP and GPR model, this chapter considers whether an improvement is seen in near surface wind speed predictions when the input data is subdivided by atmospheric stability class. In order to do this, atmospheric stability is now discussed alongside methods for classifying atmospheric stability conditions.

Atmospheric stability is a measure of the atmosphere's tendency to encourage or deter vertical motion [29]. Vertical motion in the PBL can be driven be convective processes (i.e. temperature gradient) or mechanical processes (i.e. wind shear). When dominated by wind shear the PBL is said to be neutral. When convective processes dominate, two situations can occur. When heat flows from the surface of the earth to the atmosphere a convective current is generated. This encourages

vertical motion and resulting in unstable conditions. When a convective current is not generated vertical stratification occurs, resulting in stable conditions.

Neutral conditions occur during high winds and when cloud cover prevents strong heating or cooling of the earth's surface. Unstable conditions occur when strong surface heating and low wind speed conditions occur, encouraging vertical motion of air. Stable conditions usually occur as a result of a cool surface, either the earth at night or over cool oceans. The flow of air is affected by atmospheric stability and consequently a number of different aspects of wind power forecasting can be affected. NWP models must parameterise numerous atmospheric processes in order to provide predictions of many variables at a point in time. This is a complex process which is difficult to balance as changes to the model may increase model performance in the prediction of some variables at the expense of others. The representation of turbulent mixing in stable atmospheric conditions is one such problem in most NWP models as discussed by Sandu et al. [153] and Holtslag et al. [154]. Artificially increasing turbulent mixing in stable conditions can improve prediction of near surface temperature, however, reducing turbulent mixing in stable conditions is required to improve near surface wind speed predictions [153]. The need to balance NWP model performance for many atmospheric variables means the predictions may not be optimised for every process. Because of these issues, the correction of the Met office NWP predictions of 10 m wind speed under different stability conditions is of interest.

In addition to the issues with predicting 10 m wind speeds under different stability conditions, stability impacts other aspects of wind speed and power forecasting. For example, Peterson et al. [29] document the difference in vertical wind profiles under different stability conditions. The difference in the power law under different stability conditions is also explored by Irwin [32]. This was further investigated empirically by Focken and Heinemann [155], using data from a meteorological observation mast at Cabouw in the Netherlands. In addition to affecting meteorological conditions, it has been observed that the efficiency of extracting energy from the wind is affected by stability, particularly at offshore sites [156]. For

example, Jenson [157] reported a difference of 6% between stable and unstable conditions in annual mean array efficiency at an offshore wind farm. Because of this, consideration of atmospheric stability in hub height wind speed prediction and in wind power prediction is potentially important.

Numerous methods exist for classifying stability, each requiring a range of meteorological parameters for calculation. Most methods involve some way to estimate contributions of convective and mechanically driven turbulence. Some examples include the Obukhov length, Richardson number, temperature gradient, wind speed ratio and Pasquill-Gifford stability class. The main issue surrounding calculation of some stability parameters is that they require estimates of variables such as friction velocity and heat flux which are not commonly available from either forecasts or meteorological observations. The Pasquill-Gifford method was developed to categorise the stability class based upon variables that are commonly measured at meteorological stations. The method uses solar insolation as an indication of convective turbulence and wind speed as an indication of mechanical turbulence [158]. This method for calculating stability was developed predominantly for the purpose of pollutant dispersion models but has become a commonly used classification scheme. It requires wind speed at one height, daytime solar insolation or night time cloud cover. This was further modified by Turner by using net radiation index (NRI) to estimate solar insolation based on cloud cover and cloud ceiling height, resulting in the Pasquill-Gifford-Turner (PGT) method for stability condition classification. The PGT method classifies 7 different stability conditions as given in Table 3-2.

Table 3-2: Stability categories for PG and PGT stability methods

PGT class	Stability condition
1	Highly unstable or convective
2	Moderately unstable
3	Slightly unstable
4	Neutral
5	Slightly stable
6	Stable
7	Extremely stable

The first step in obtaining the stability classification is to calculate the insolation class number. This is obtained based on solar altitude as outlined in Table 3-3. NRI is calculated using the algorithm given in Figure 3-2, where cloud cover is given in tenths, with 1/10 indicating low cloud cover and 10/10 indicating opaque cloud. Finally, using NRI and wind speed, the stability classification is obtained from Table 3-4.

Table 3-3: Insolation class number

Solar Altitude (φ)	Insolation	Insolation class number
$60 < \varphi$	Strong	4
$35 < \varphi \leq 60$	Moderate	3
$15 < \varphi \leq 35$	Slight	2
$\varphi \leq 15$	Weak	1

The PGT method allowed stability conditions to be estimated based on MIDAS observations. However, the forecasted variables available from the Met Office forecast data do not include sufficient details of cloud conditions to allow the use of the PGT method. Because of this, stability conditions used in this chapter have been based on MIDAS observations at the time of the prediction. Hence this chapter presents preliminary results exploring the benefit of using atmospheric stability in a GPR model for predicting wind speed. Chapter 4 explores how this work could be extended to explore the impacts of using predicted stability conditions.

Figure 3-2: Algorithm for calculating net radiation index [159].

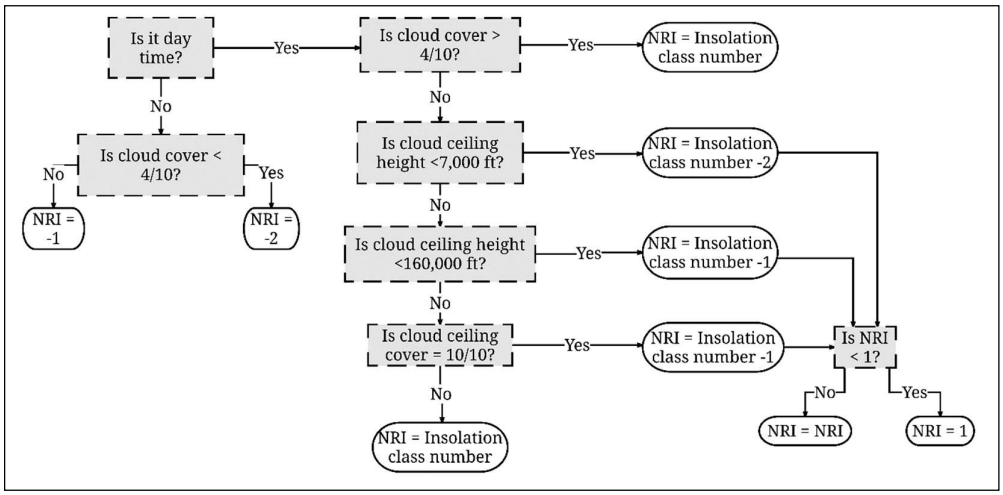


Table 3-4: PGT stability classes

Wind speed (ms^{-1})	Net radiation index (NRI)						
	4	3	2	1	0	-1	-2
0-0.7	1	1	2	3	4	6	7
0.8-1.8	1	2	2	3	4	6	7
1.9-2.8	1	2	3	4	4	5	6
2.9-3.3	2	2	3	4	4	5	6
3.4-3.8	2	2	3	4	4	4	5
3.9-4.8	2	3	3	4	4	4	5
4.9-5.4	3	3	4	4	4	4	5
5.5-5.9	3	3	4	4	4	4	4
≥6	3	4	4	4	4	4	4

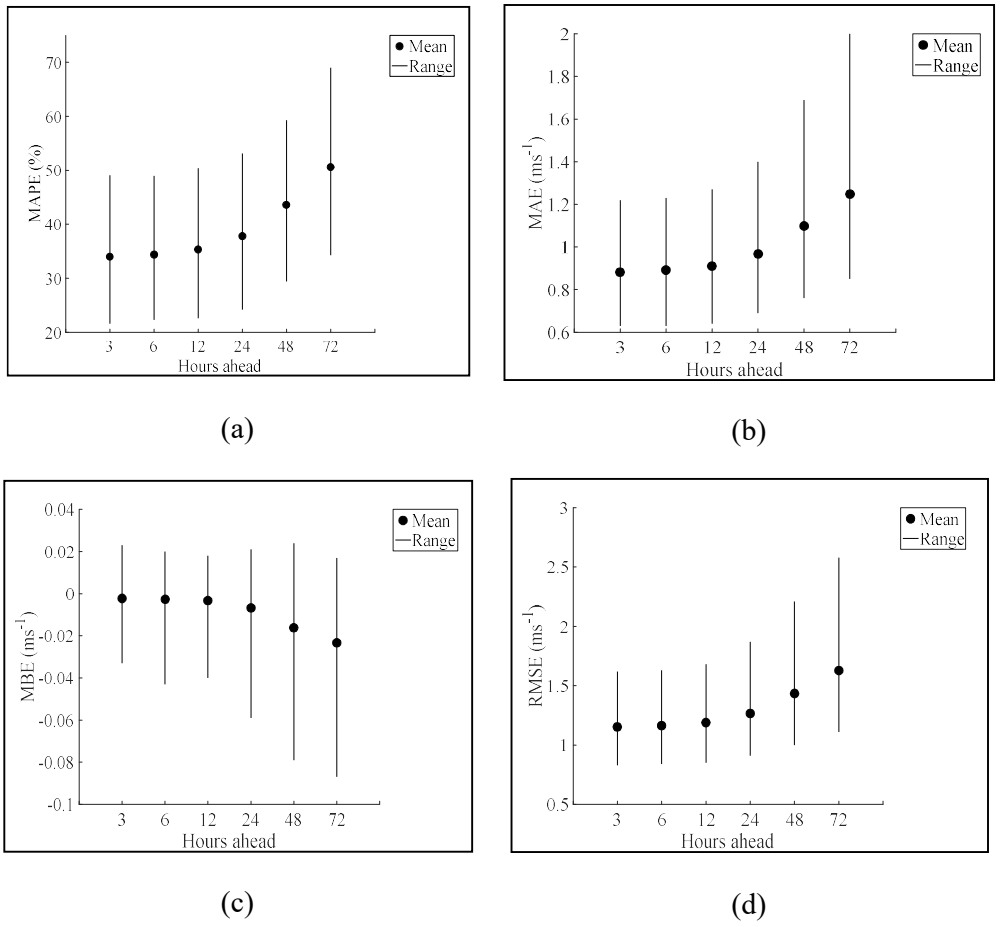
3.2.6 NWP performance

In this section, the performance of the Met Office NWP predictions for near-surface wind speed are evaluated. The NWP predictions of near surface wind speed contain some errors, which the hybrid NWP and GPR model for near surface wind prediction aims to improve upon. In order to understand the conditions under which the NWP predictions of 10 m wind speed may require improvement, a brief analysis

of the NWP performance is performed. The model performance for different prediction time frames and how model performance is affected by atmospheric stability is considered. This provides some motivation for the work performed. The Met Office NWP is evaluated at 15 sites over a range of forecast horizons to consider how these predictions can be improved through the use of GPR model.

In Figure 3-3 the mean error statistics over the 15 sites are shown alongside the highest and lowest errors and any individual site (the range). In Figure 3-3 RMSE ranges between approximately 0.8 and 1.6 ms^{-1} for 3 hours ahead and between 1 ms^{-1} and 2.5 ms^{-1} at 72 hours ahead. The lowest RMSE varies between site 6 and site 2, whilst the highest is at site 11. This is similar to what is seen for MAE, where the lowest values are seen at site 6 (with site 2 a close second), and the highest values seen at site 11. For MBE, positive values imply an over prediction whilst negative values imply an under prediction. This indicates that at most sites there is a slight under-prediction. Overall, the magnitude of bias is small but increases as the forecast horizon increases. This indicates a tendency in the NWP wind speed predictions which could potentially be reduced using the GPR model. There is a significant difference in the performance of the NWP model at each site. This indicates that some sites potentially have more complex site characteristics that are not always correctly captured by the NWP model. For example, where there are trees, bodies of water, hills or outcrops of rock which are smaller than the grid resolution of the NWP model the impact of these on wind profiles are unlikely to be well modelled. Furthermore, it highlights that any results and potential improvements of a model will be highly site specific which should be considered carefully when drawing comparisons with other predictive models.

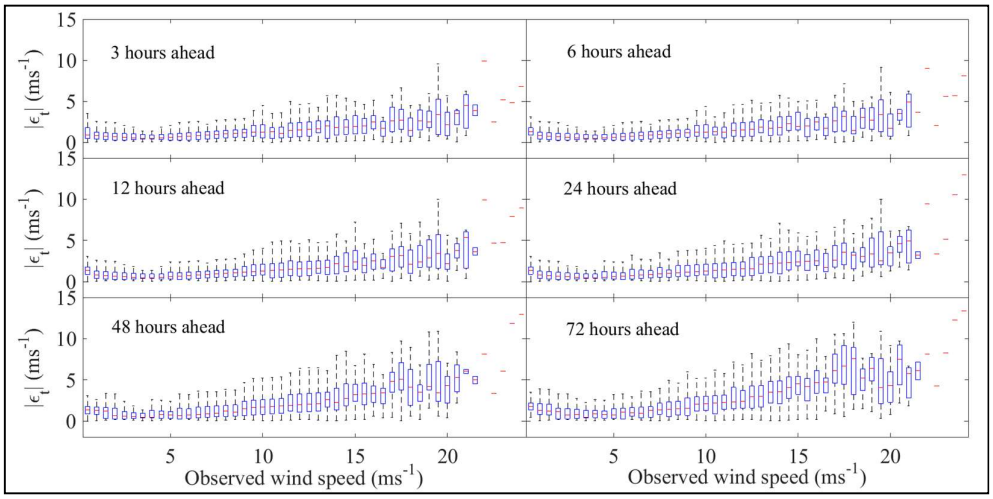
Figure 3-3: Summary of NWP error statistics at all 15 MIDAS sites. Mean error over all 15 sites and the highest and lowest errors at any one site are shown.



Further to considering the error statistics at different time horizons, the variation in error with different wind speed is considered. As some higher wind speeds are only observed at some sites, the absolute error over all observations at 15 sites are considered. These are shown in Figure 3-4. It can be seen that there is a larger variation in absolute error at low wind speeds (up to about 3 ms^{-1}) with lower absolute errors seen between 3 and 5 ms^{-1}. Then as the wind speed increases both the MAE and the range of absolute errors increases with wind speeds above 5 ms^{-1}. When considering models for wind power production low wind speeds are an issue when considering turbine cut in, and hence whether any power will be produced, and for higher wind speeds the difference in power output with wind speed will be

more pronounced. This can be seen from the power curve shown in Section 2.1.3. For the turbine shown, the curve is steepest for wind speeds between 5 and 15 ms^{-1}. Hence, differences between observed and predicted speeds around this range will result in larger changes in power output. Furthermore, because the datasets considered here are for 10 m wind speeds the errors may be compounded when scaled to hub height as wind speeds generally increase with height above ground level as discussed in Section 2.1.3. Hence it is worth considering errors over the full range of wind speeds in order to look at how to reduce errors.

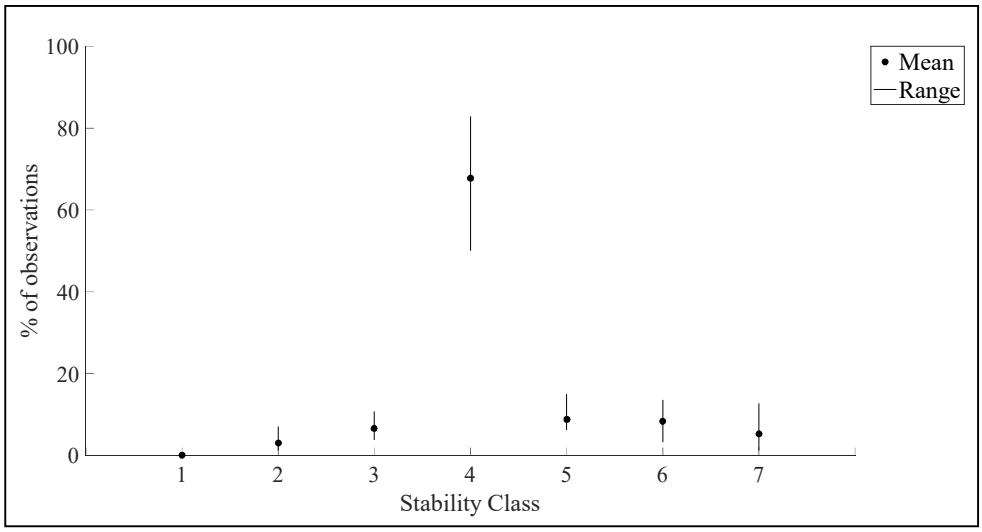
Figure 3-4: Graphs showing the change in absolute error with wind speed for all observations at all 15 sites. This shows a boxplot of the spread of absolute error, $|\epsilon_t| = |\hat{y} - y_t|$, at different observed wind speeds. The blue box represents the interquartile range, the horizontal red line represents the median and the vertical dashed line represents the range (maximum and minimum values).



Following an analysis of the change in errors with wind speed, the difference in error under different stability classes was considered. For the sites chosen, the PGT stability class was calculated for every hour of observations using the method outlined in Section 3.2.5. Figure 3-5 shows the percentage of observations in each stability class, shown as the range and mean overall 15 sites. Most observations fall into the neutral stability class, with the maximum at any site being 89% and an average of 66%. The lowest percentage of observations falling into the neutral

category at any site was 49%. The stability conditions vary from site to site and are influenced by the local site characteristics. Whilst neutral conditions are most frequent, and most frequently attributed to higher wind speeds, considering stability conditions can help to understand the NWP performance.

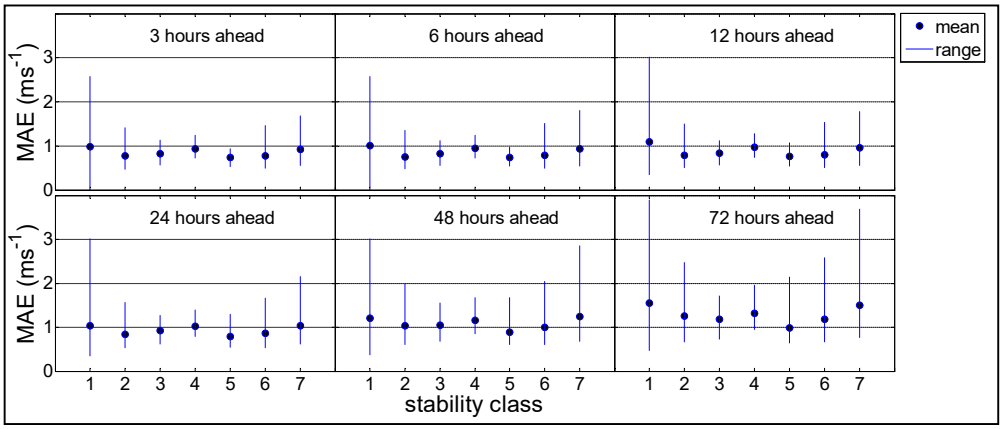
Figure 3-5: Percentage of observations in each stability class.



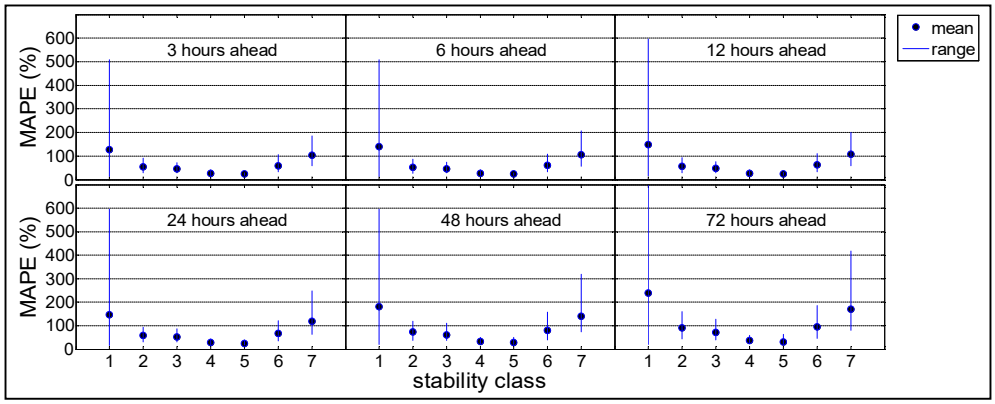
In Figure 3-6 the range and mean error statistics across all 15 sites are shown for 7 stability classes. From this, it can be seen that the MAE is typically slightly higher under neutral conditions than under slightly stable or unstable conditions. This could be due to the higher wind speeds included in this category. RMSE shows similar trends. MAPE is lowest for all time horizons either under neutral conditions or slightly stable conditions. Furthermore, MBE shows an over prediction outside of neutral conditions, which is more pronounced as the time horizon increases. This is in contrast to the under prediction seen in the overall NWP predictions seen in Figure 3-3(c). This is because there is a high frequency of neutral conditions observed, as shown in Figure 3-5, in which there is a negative bias. This indicates that separating the observations by stability class may allow the positive bias seen under non-neutral conditions to be corrected more effectively than if all data is considered together. Figure 3-7 shows a sample of the error statistics for one site so that the characteristics can be seen more clearly. The average error for the time

horizon (averaged over observations in all stability classes) are shown as a dashed line for reference. It can be seen that MAPE is smaller under neutral conditions, and increases significantly for stable conditions. It also shows that MAE is slightly higher under neutral conditions compared to slightly stable or unstable conditions at this site. This could be because under the PGT classification scheme observations with a high wind speed are classified as neutral, and as seen in Figure 3-4 at higher wind speeds the absolute error is higher. Furthermore, MBE approaches zero for neutral conditions and shows an over prediction in non-neutral conditions. It is unclear why this positive bias occurs under non-neutral conditions. However, this suggests that a model which subdivides data by stability class could reduce bias in these conditions. The trends seen at this site are representative of all 15 sites, as seen in Figure 3-6. The difference in error characteristics seen under different stability conditions suggests that considering this in a wind prediction model may be beneficial.

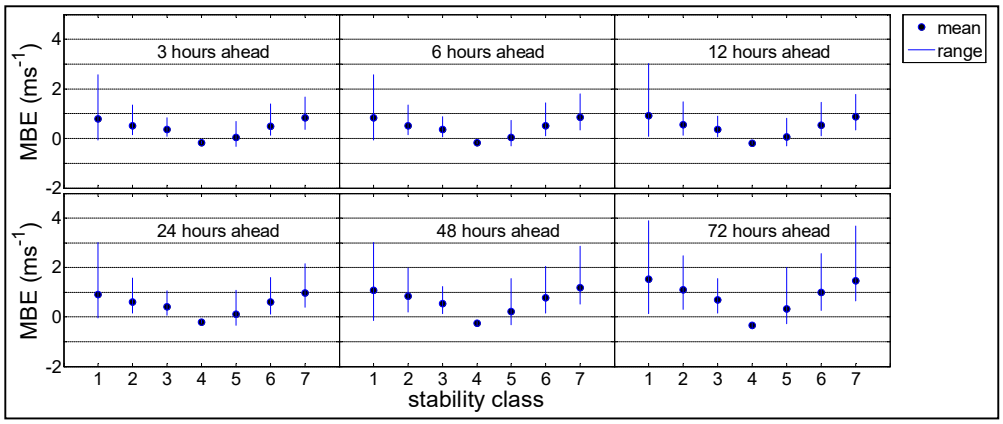
Figure 3-6: Mean, highest and lowest NWP prediction errors across all 15 sites for different stability classes. (a) Shows MAE, (b) MAPE, (c) MBE and (d) RMSE.



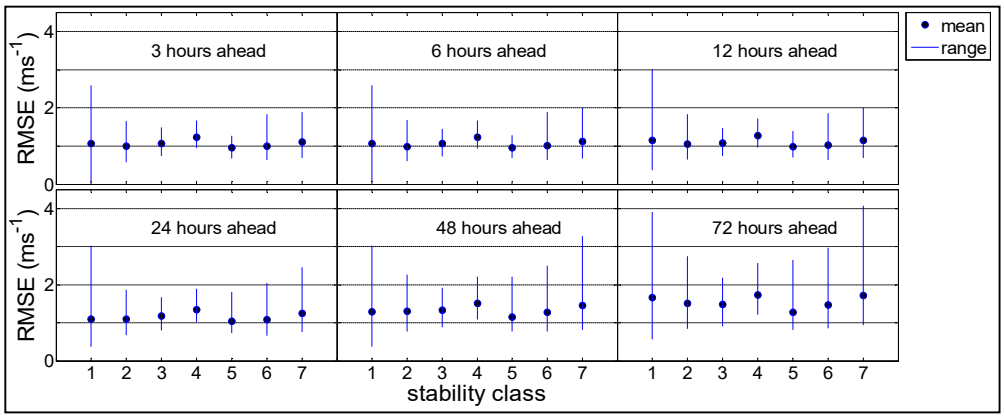
(a)


3 hours ahead
6 hours ahead
12 hours ahead
24 hours ahead
48 hours ahead
72 hours ahead
MAPE (%)
stability class
mean
range

(b)

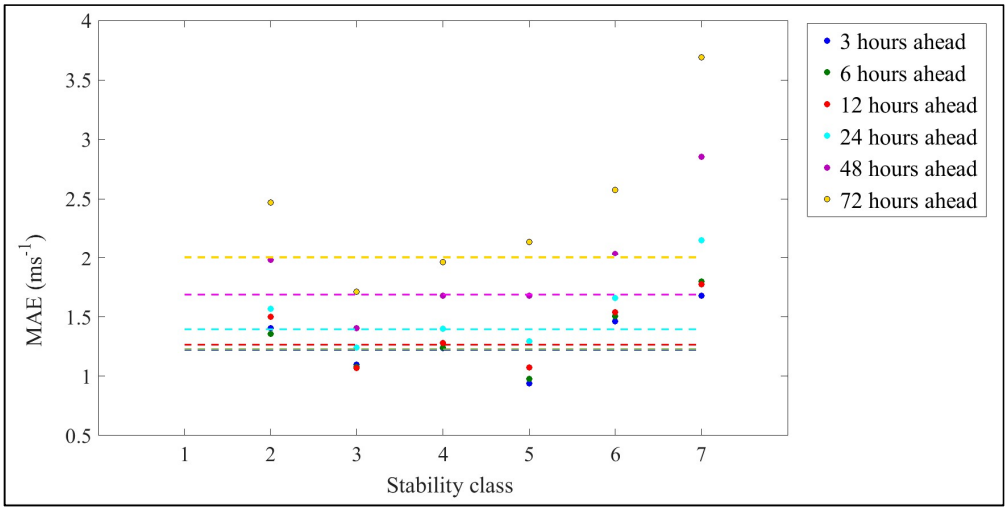

3 hours ahead
6 hours ahead
12 hours ahead
24 hours ahead
48 hours ahead
72 hours ahead
MBE (ms-1)
stability class
mean
range

(c)

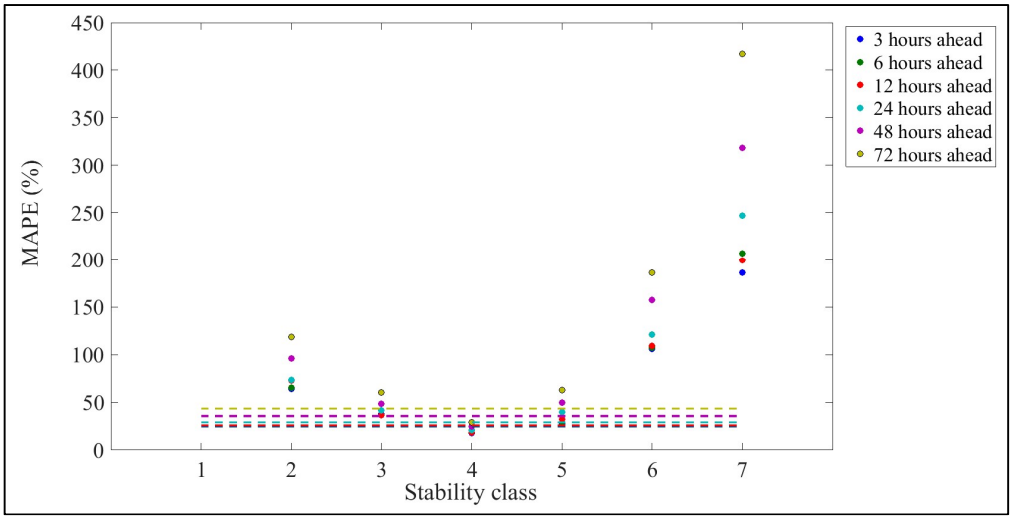

3 hours ahead
6 hours ahead
12 hours ahead
24 hours ahead
48 hours ahead
72 hours ahead
RMSE (ms-1)
stability class
mean
range

(d)

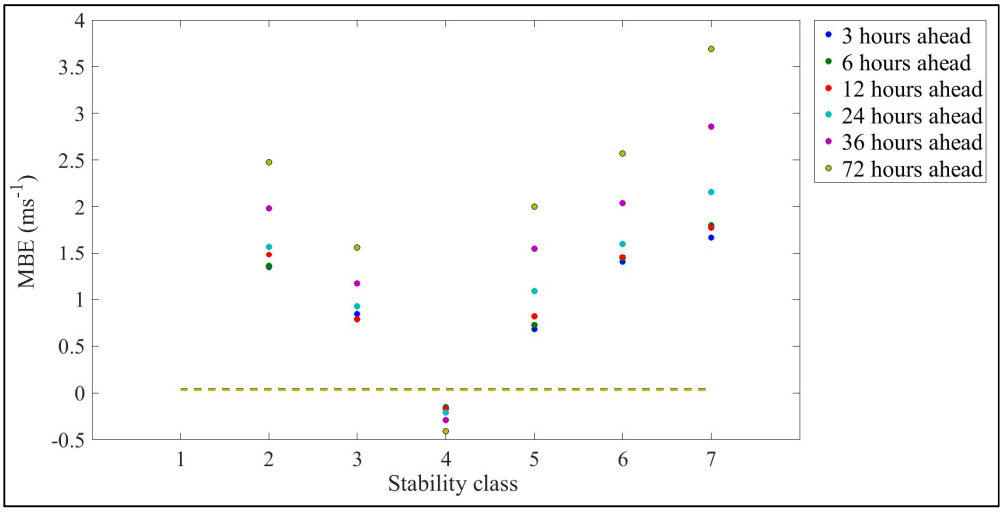
Figure 3-7: The difference in error statistics for a sample site. (a) Shows MAE, (b) MAPE, (c) bias and (d) RMSE. In all four graphs, the dashed lines (of the same colour as in the legend) indicates the average error for that time horizon (averaged over observations in all stability classes).



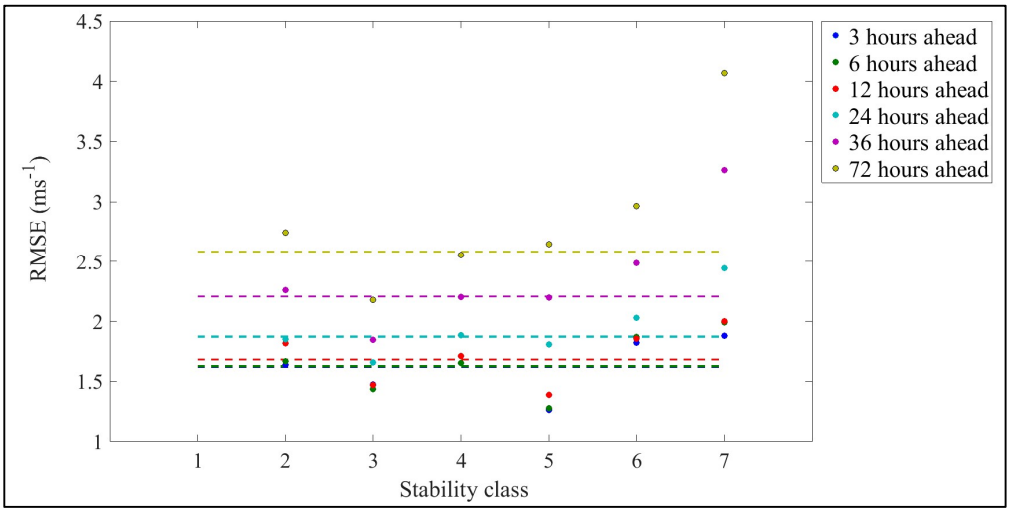
(a)



(b)



(c)



(d)

3.2.7 Model set-up

An introduction to the GPR model is given above, in addition to an overview of the data used in this chapter. However, the model inputs and outputs require further definition. As detailed in Section 3.2.1 the model develops a relationship between

target variable $\mathbf{y}$ and input variables $\mathbf{x}$ of the form $y_i = f(\mathbf{x}_i) + \varepsilon_i$. The model is a multivariate regression model with 4 predictor variables;

(i) The Met Office prediction of near surface wind speed.
(ii) 3 hours of observed hourly wind speed data prior to the beginning of the forecast.

Forecasts up to 72 hours in advance were considered, at 3 hour intervals. Hence the predictor variables are given by Equation 3-11.

$$\mathbf{X}_t = [\mathbf{m}_t, \mathbf{y}_{t-k-1}, \mathbf{y}_{t-k-2}, \mathbf{y}_{t-k-3}] \text{ for } t = 4, \dots, n \qquad \text{Equation 3-11}$$

where:

$t =$ time of observation
$\mathbf{m}_t =$ Met Office forecast of 10 m wind speed at time (ms^{-1})
$\mathbf{y}_t =$ observed wind speed at time t (ms^{-1})
$k =$ Forecast horizon

The regression model is used for the prediction of 10 m wind speeds and hub height wind speeds. For both applications the Met Office forecast of 10 m wind speeds used as input variable $\mathbf{m}_t$ are taken from the Met Office NWP model described in Section 3.2.4. The source of the observed wind speeds $\mathbf{y}_t$ differ depending on the application of the forecast.

Section 3.3.1 presents the results of the forecast model for predicting wind speed at 10 m above ground level for 15 MIDAS sites across the UK. In this case, the target variables, $\mathbf{y}_t$, are the MIDAS observations of 10 m wind speed at the site. Due to the variables required to estimate stability conditions, only sites where information on wind, cloud depth and coverage was available were considered. 15 sites were investigated and the location of these sites was shown in Figure 3-1.

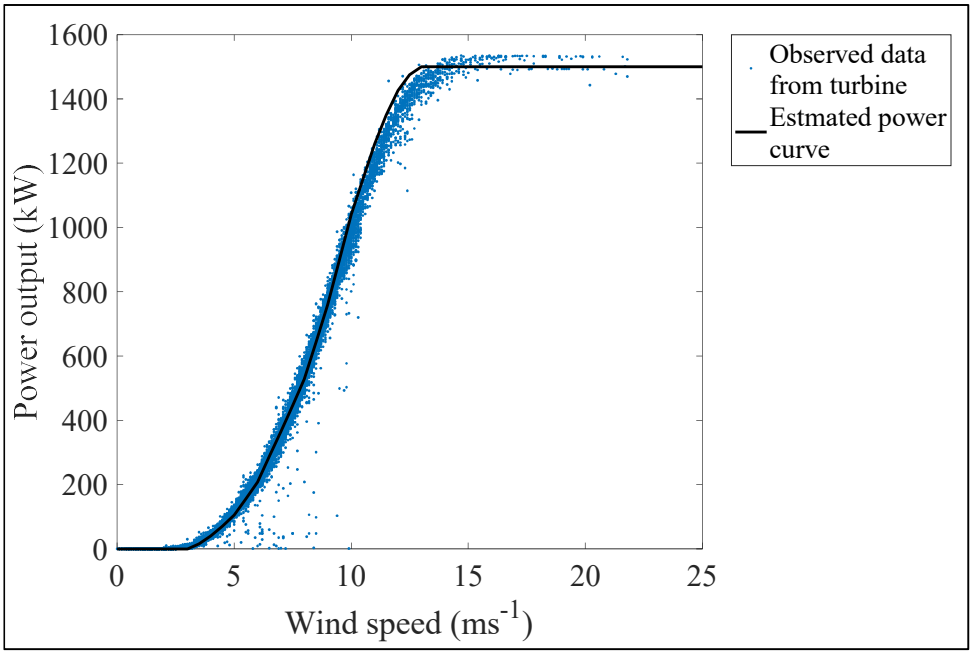
Further to this, for one turbine in the UK hub height wind speed and power output data was available. Section 3.3.2 presents results of the forecast model for wind speed prediction at hub height for a suburban location in the UK. In this case, the target variables, $\mathbf{y}_t$, are the hub height wind speed observations at the site. Hub height wind speeds were obtained from the wind farm operator at one site in the UK for a 1.5 MW turbine. Power output and hub height wind speed data for wind

turbines is generally difficult to obtain due to commercial sensitivity, hence only one dataset is used to show results for this work. Wind speed was measured at approximately 65 m above ground level. MIDAS data from an observation site located approximately 8 km from the turbine is used to calculate the stability class at the time of forecast, and the Met Office forecast data is taken from the same location as the MIDAS data.

Finally, Section 3.3.3 explores the potential impact of improved hub height wind speed forecasting on wind power forecasting. In this case, the wind speed predictions from Section 3.3.2 were translated into power output using a power curve. A power curve is a relationship between wind speed and power output, these were discussed in Section 2.1.3. The turbine in use at the location in question is an old model, for which the manufacturer's power curve is not available. Hence the power curve used has been chosen from a database of available power curves obtained from "The Wind Power" a wind energy market intelligence company [39]. A power curve was chosen for a turbine of the same size, for which the relationship between wind speed and power approximately matched the observed wind speed and power data for the turbine. The observed data and the power curve chosen are shown in Figure 3-8. Whilst it would be preferable to obtain the power curve for the actual wind turbine from which the data was obtained, the approximate relationship between wind speed and power observed in the power curve used is sufficient for this initial analysis.

To ensure an independent forecast, the data was split into a training dataset and a test dataset. This avoids an overestimate of the skill of a forecast which can occur when model performance is assessed on a data set which has been used to train the model. The training dataset was used to train model hyperparameters and the test dataset to assess the model performance. The training data are defined as the concurrent observations and Met Office forecast data for the first 9 months of 2014 and the test data are same data from the final 3 months of 2014.

Figure 3-8: Observed data for the 1.5 MW wind turbine used in this chapter and an estimated power curve for a similar turbine [39]



The observation data was available at hourly intervals and the forecast data at 3 hourly intervals. This describes the formulation of the GPR model with no stability data. To test the impact of using atmospheric stability to improve the model, the data was split into 7 stability classes and the model trained separately for each class. From the variables available in the Met Office forecast an indication of forecasted stability conditions is difficult to obtain, therefore currently this study uses the stability conditions at the time of the observation, as calculated from the MIDAS data using the method described in Section 3.2.5. Whilst in the case of an actual forecast scenario this information would not be available it gives an indication of the potential improvements possible using stability information in a GPR model.

3.3 Results

3.3.1 MIDAS site wind speed prediction

The hybrid NWP and GPR model was first used to predict 10 m wind speeds at 15 MIDAS locations. Wind speeds predicted by the GPR model are compared to the MIDAS observations with several criteria used to assess performance. Here three criteria are shown, mean absolute error (MAE), mean absolute percentage error (MAPE) and root mean squared error (RMSE), calculated using Equations 2-11 to 2-13 given in Section 2.2.4. The comparison of different error metrics allows a full overview of the model performance.

The model errors are shown for a GPR model in which the datasets were split by stability class and a GPR model using the full dataset (referred to as a simple GPR model). These results are compared to the NWP predictions made by the Met Office. In order to fully illustrate the model results, detailed results are shown for 4 of the 15 MIDAS sites tested, and summary results are shown for the 15 sites. In Figure 3-9 and Figure 3-10 MAE and MAPE are shown for 4 of the 15 MIDAS sites. This shows how the errors increase as the forecast period increases, and also how the model error is reduced by using the GPR model with data subdivided by stability class. Overall, the simple GPR trained using the full dataset reduces the error in predicted wind speed compared with predictions made by the NWP. The improvement is site specific, with the greatest error reduction seen at a forecast period of 3 hours ahead for some sites, and further ahead for others. Figure 3-13 shows a summary of errors over the 15 sites. This shows an average reduction in MAPE of approximately 2% for the simple GPR model in comparison to wind speeds predicted by the NWP. For the GPR model with stability information there is a reduction in MAPE of 5% for a 3 hour forecast period, rising to 9% for a 72 hour forecast period. The simple GPR model also shows an average of 2% improvement in MAE and RMSE for all forecast periods, whilst the GPR model with stability information shows a 10% improvement in MAE and a 7% improvement in RMSE.

Figure 3-9: MAE (with 95% confidence interval) for 4 sample MIDAS sites shown for GPR models both with and without stability, and the NWP.

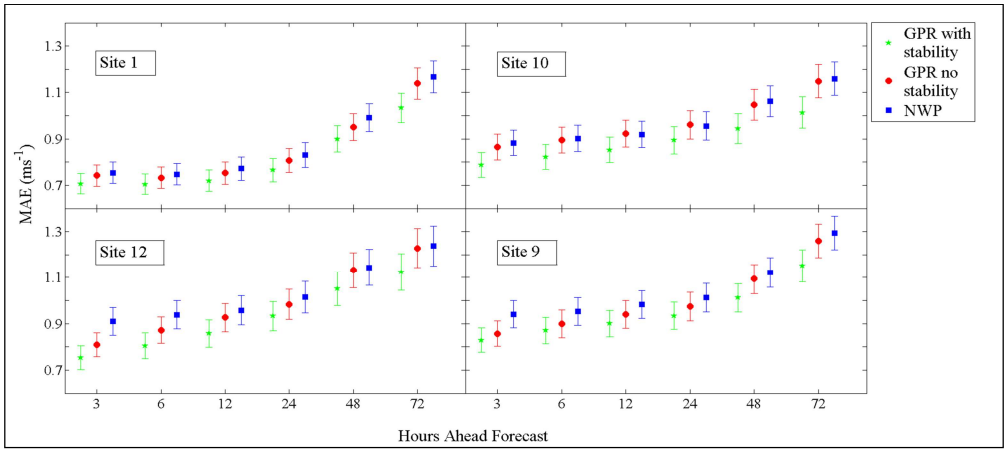
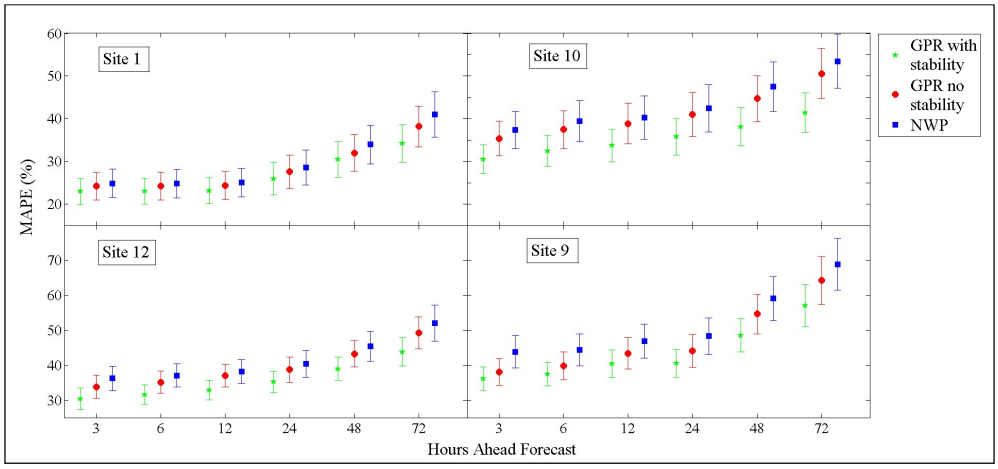


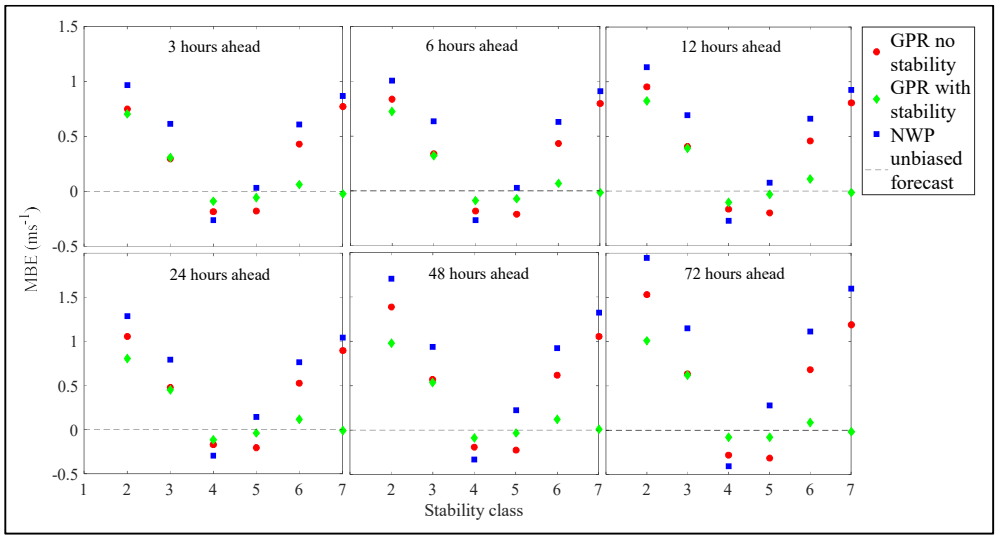
Figure 3-10: MAPE (with 95% confidence interval) for 4 sample MIDAS sites shown for GPR models both with and without stability, and the NWP.



In Figure 3-6 (c) MBE averaged over the 15 MIDAS sites was shown for the NWP model. This indicated that in unstable and stable conditions the NWP over predicted 10 m wind speeds and in neutral conditions MBE approached zero but was slightly negative. In Figure 3-11 MBE averaged over the 15 MIDAS sites is shown for the simple GPR model and the GPR model with input data subdivided by stability class. MBE is shown for each stability class. The results indicate that the GPR model with input data sub divided by stability class is able to reduce over prediction in stable

and extremely stable conditions much more than the simple GPR model. The simple GPR model reduced MBE in comparison to the NWP predictions. However, not as much as the GPR model with data sub divided by stability class. There was not as much difference between the two GPR models for slightly and moderately unstable conditions. This was because there was less data for these classes in the training data set. There was still some over prediction in these conditions when the GPR models were implemented. Both GPR models reduced the under prediction seen for predictions in neutral stability conditions, with the GPR model with data subdivided by stability class having MBE closer to zero than the simple GPR model.

Figure 3-11: Average MBE for all 15 MIDAS sites for different stability classes.

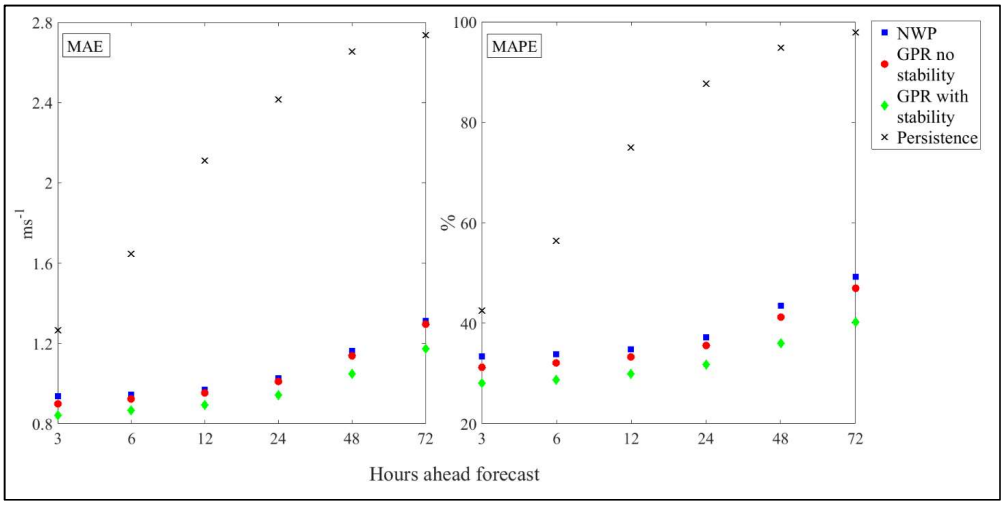


These results indicate that the hybrid NWP and GPR model with data subdivided by stability class is offers improved near surface wind speed predictions over the NWP predictions or a simple NWP and GPR model. This is largely attributed to the GPR models with data subdivided by stability class being able to correct the bias in predictions in non-neutral conditions in the NWP predictions.

In Figure 3-12 MAE and MAPE for the simple GPR model and the GPR with data subdivided by stability are shown in comparison to the persistence model. It can be seen from this that there is a significant reduction in error in comparison to the

persistence method for MAE, RMSE and MAPE. For the GPR model with stability, the reduction in MAPE over the persistence model is 14.5% at 3 hours ahead, increasing to 57.6% at 72 hours ahead.

Figure 3-12: Average MAE and MAPE for all 15 MIDAS sites shown in comparison to the persistence method.

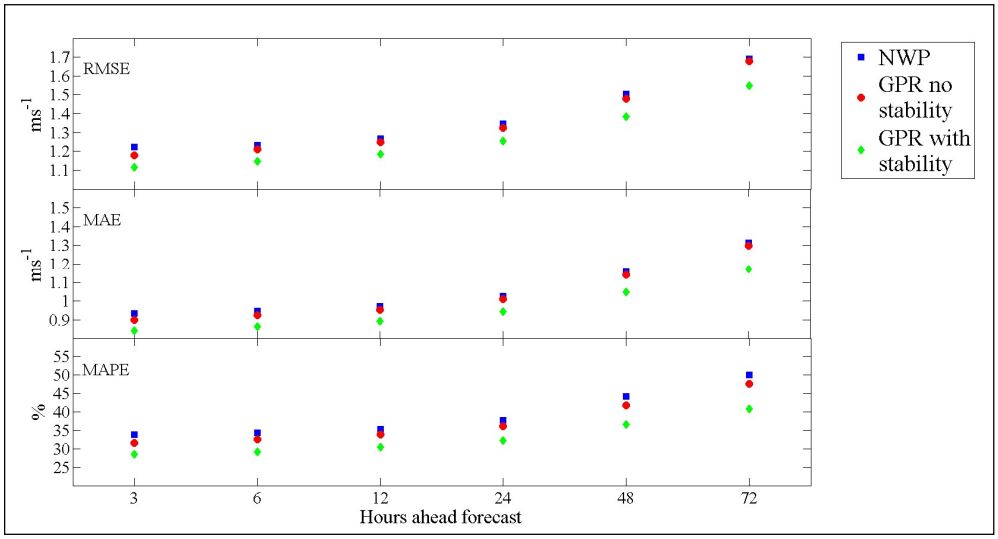


The average RMSE and MAE overall 15 sites can be compared with other wind speed prediction models using the results shown in Table 2-6. Figure 3-13 shows the average RMSE across the 15 sites for the GPR model with stability is 1.1 ms^{-1} at 3 hours ahead. At 1 hour ahead Li and Shi [60], Chen et al. [71] and Li et al. [61] give an RMSE of between 0.96 ms^{-1} and 1.5 ms^{-1}. Hence an RMSE of 1.1 ms^{-1} at 3 hours ahead is within the range of a good forecast. For the same three studies, an MAE of between 0.72 ms^{-1} and 1.13 ms^{-1} is reported for a forecast 1 hour ahead. Figure 3-13 shows an average MAE of 0.82 ms^{-1} at 3 hours ahead for the GPR model with stability, again falling within the range shown by other studies.

At 72 hours the average RMSE for all 15 sites for the GPR model with stability is 1.54 ms^{-1} which is smaller than the RMSE reported by Louka et al. [72] of 2.38 – 2.88 ms^{-1} and Chen et al. [71] of 2.04 ms^{-1}. Similarly, the average MAE over 15

sites for the same model at 72 hours ahead is 1.17 ms^{-1} compared to 1.75 – 2.04 ms^{-1} reported by Louka et al. [72] and 1.6 ms^{-1} reported by Chen et al.[73]. Furthermore, the MAPE shown in Figure 3-13 for the GPR model with stability at 72 hours is 42%, slightly lower than the 44% reported by Chen et al.[73]. Overall the model performs well in comparison to other similar models, and improvements are seen when stability is incorporated in the GPR model for predicting near-surface wind speed.

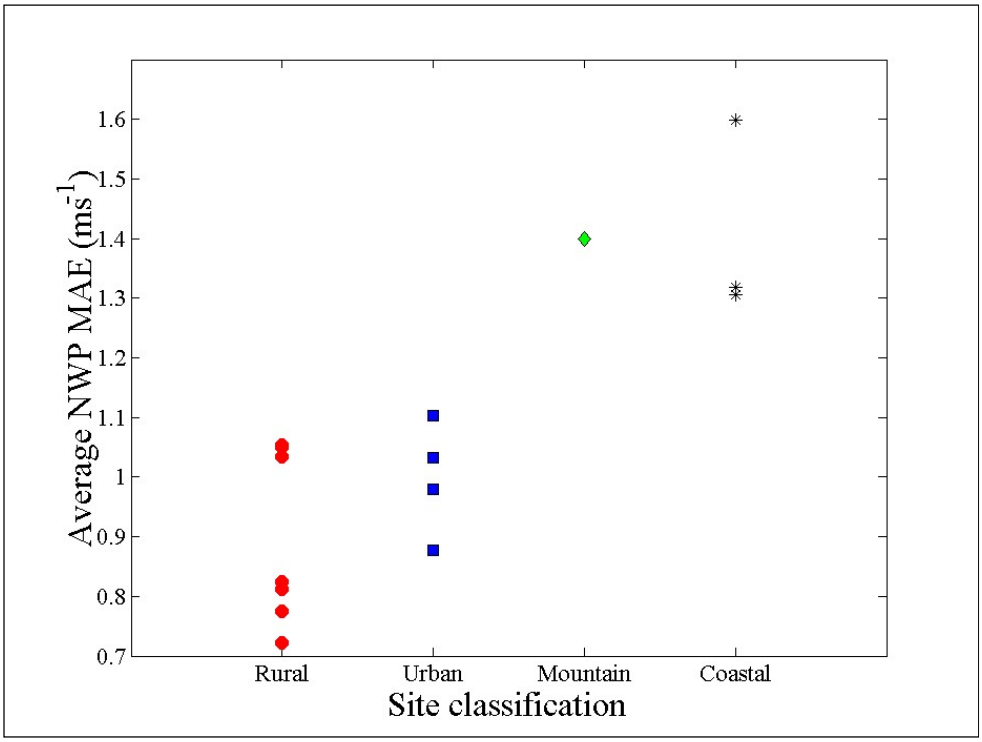
Figure 3-13: Average RMSE, MAE and MAPE for all 15 MIDAS sites. Error bars are not shown here to allow clarity.



The model performance was also considered for the 4 different site categories observed; rural, urban, mountain and coastal. Differing meteorological effects present different forecasting issues dependent on site characteristics. For example at coastal sites wind speed is affected by changes in surface roughness, and availability of heat and moisture [30]. In mountainous areas, complex orography and changes in temperature drive wind speeds, and within urban areas high densities of buildings can interfere with expected wind patterns. Taking this into account one might expect the model results to vary with different site characteristics. Within the 15 MIDAS sites considered there was 1 mountain site, 3 coastal sites, 4 urban sites and 7 rural

sites observed. Figure 3-14 shows how the Met Office NWP error varies within different classifications. Average NWP error is shown for each site, calculated as an average over the time periods considered (3 hours ahead, 6 hours ahead, up to 72 hours ahead). For the three coastal sites in the dataset, the average NWP model error is higher than for the rural and urban classifications. Only one mountain site is identified within the set, hence it is difficult to suggest whether the results seen at this site are representative of all mountain sites. The NWP errors observed for this mountainous site are higher than the errors seen at rural and urban sites. The difference between errors in rural and urban sites seems to be small. However, marginally higher errors are seen at the urban sites.

Figure 3-14: Average mean absolute error of the NWP forecast across all time periods considered (3 hours ahead – 72 hours ahead). A single average is shown for each site, with the sites split by site classification.

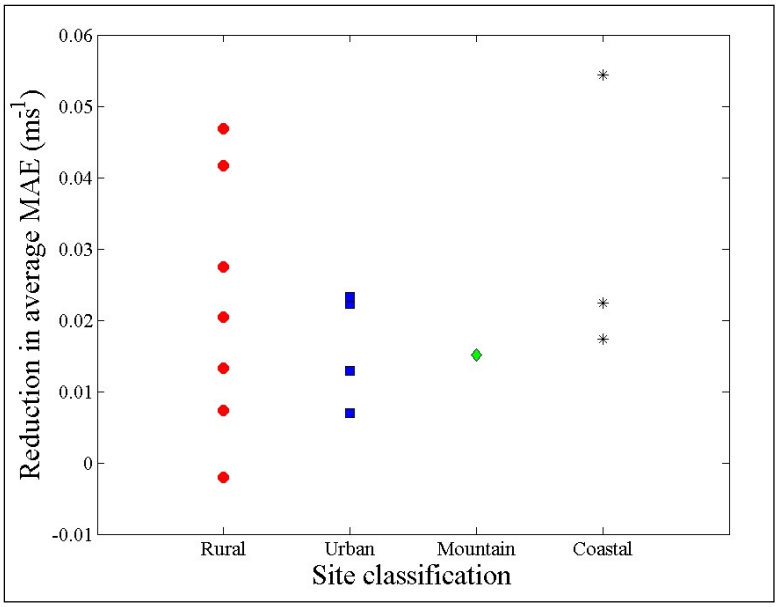


Having examined the NWP prediction error for different site classes, the reduction in error achieved using both GPR models is considered. This is calculated using Equation 3-12.

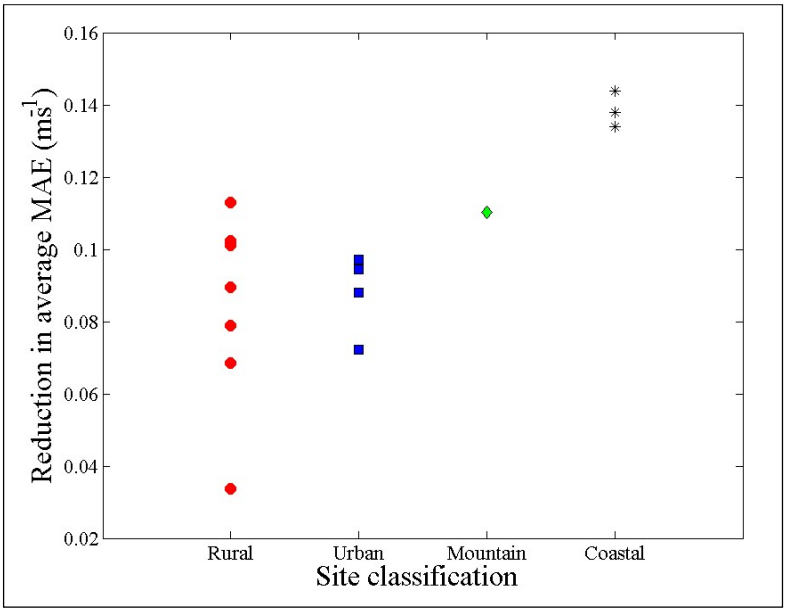
$$\text{Reduction in average MAE} = \sum_{t=3}^{72}\left(\text{MAE}_{t,\text{NWP}} - \text{MAE}_{t,\text{GPR}}\right) \qquad \text{Equation 3-12}$$

where t is the time ahead forecasted. This reduction in error is shown in Figure 3-15 for both the simple GPR model and the GPR model with data subdivided by PGT stability class. It can be seen that for the simple GPR model, lower errors are seen at all but one site. The site which did not achieve an improvement over the NWP was a rural site at which the NWP prediction error was the lowest of any sites considered, making it difficult to make enhanced predictions. Despite this, an improvement was seen when using GPR with stability at this site. Figure 3-15 shows that for the simple GPR the reduction in model error is not significantly different between site classes. However for the GPR model with data sub divided by stability classes a larger improvement is seen at coastal sites and at the mountain site. This is because the over prediction in the NWP model for stable and unstable conditions was higher at coastal and mountainous sites and the GPR model with data subdivided by stability class was better able to address this. Given that in Figure 3-14 it was observed that coastal and mountain sites had the highest prediction errors from the NWP model, this shows that the method improves upon sites where prediction accuracy is lower, which may be useful for wind farms located in regions with highly variable wind regimes. An improvement is also seen when using the GPR model with stability for prediction over both the simple GPR and the original NWP model in urban and rural areas. However, the achievement of the GPR model with stability at urban and rural sites is slightly less pronounced than for the coastal and mountain sites.

Figure 3-15: Reduction in error achieved by applying the GPR model compared to Met Office NWP model (A larger reduction indicates better model performance). (a) Shows results for simple GPR model, whilst (b) shows results for GPR model with data subdivided by PGT stability class.



(a)



(b)

3.3.2 Hub height wind speed prediction

Whilst looking at the prediction of 10 m wind speeds shows the potential of the GPR model and the importance of stability in the reduction of model error, for wind power prediction hub height wind speed prediction is more important. In Figure 3-16 MAPE and MAE are shown for hub height wind speed for both the simple GPR model and the GPR with data subdivided using PGT stability class. It shows a reduction in MAPE of between 1 and 2% and between 3 and 5% reduction in MAE using a GPR model with data subset by stability class. In Figure 3-16 the persistence results are omitted in order to show more clearly the difference between the two models. In Figure 3-17 the MAE, MAPE and RMSE for both GPR models are shown in comparison with a persistence model. It can be seen in Figure 3-17 that the GPR model shows significant improvements over the persistence model.

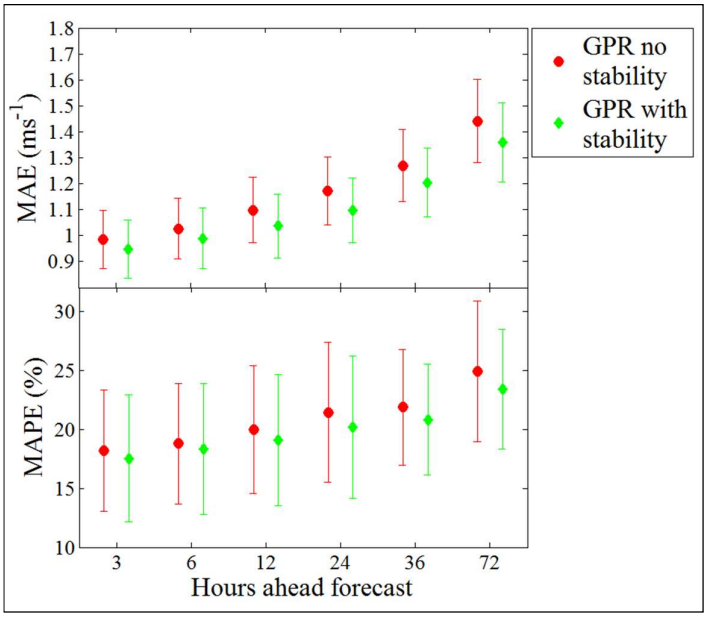
Taking the results for the GPR model with data subdivided according to stability class, comparisons can be drawn with other methods seen in Table 2-6. The GPR model with stability has an MAE of 0.95 ms^{-1} at 3 hours ahead, as shown in Figure 3-16 and Figure 3-17. This is lower than some of the results shown in Table 2-6. For example, Chen et al. [71] report an MAE of between 0.72 and 1.1 ms^{-1} for a forecast between 1 and 4 hours ahead, Li and Shi [60] between 0.9 and 1.05 ms^{-1} for a forecast 1 hour ahead and Li et al. [61] 1.137 ms^{-1} at 1 hour ahead. Similarly the GPR model with stability has an RMSE of 1.2 ms^{-1} at 3 hours ahead, compared to Chen et al. [71] who reported between 0.96 and 1.95 ms^{-1} for a forecast between 1 and 4 hours ahead, Li and Shi [60] who reported between 1.2 and 1.4 ms^{-1} for a forecast 1 hour ahead, and Li et al. [61] 1.5 ms^{-1} at 1 hour ahead.

At 72 hours ahead MAE for the GPR model with stability rises to 1.36 ms^{-1}, however this is still lower than the results presented by Louka et al.[72] and Chen et al. [73] for this timescale in Table 2-6. Similarly, RMSE rises to 1.7 ms^{-1}, again lower than the results from Louka et al. [72] and Chen et al. [73]. Figure 3-17 shows MAPE at 72 hours ahead for this model is 24%, which is 1% lower than for a GPR

model without using stability, and 31% lower than the persistence method. MAPE is not given for predictions 72 hours ahead for any other model shown in Table 2-6.

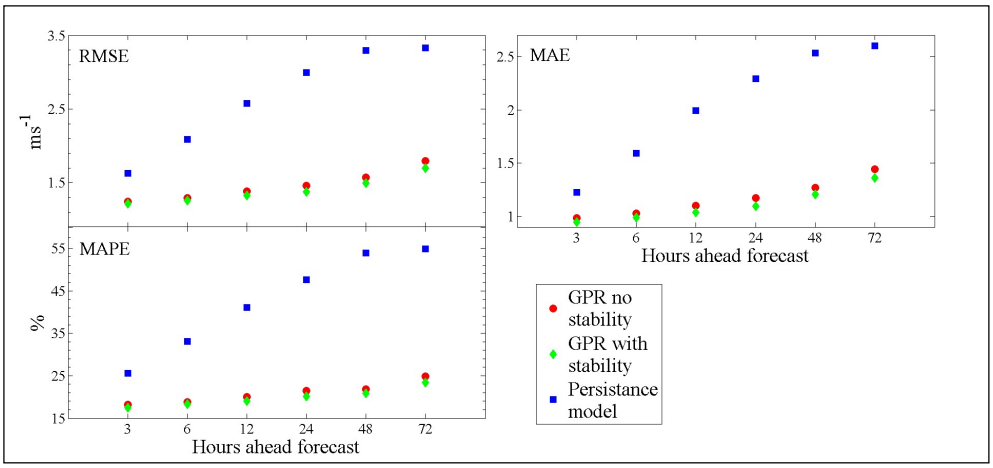
It is difficult to make direct comparisons of the results seen in this section with other models presented in literature as each study uses a distinct dataset. Because of this it is not possible to state that one model exclusively outperforms others. However, the results seen in this section are similar to results seen in literature indicating that this model is competitive with other models for hub height wind speed predictions.

Figure 3-16: MAPE and MAE (with 95% confidence interval) for hub height wind speeds predicted with a simple GPR model and a GPR model with stability



These results indicate that using a hybrid GPR and NWP model for hub height wind speed prediction can improve wind speed predictions over a benchmark method (the persistence model). Furthermore, hub height wind speed prediction errors are reduced when the input data is subdivided by stability class. This suggests that this method could be used to predict wind speeds in a way which would be valuable to users in the wind industry. It would be beneficial to use this model to predict hub height wind speeds at for other sites to see whether this result is replicated. However, at the time of this analysis sufficient data was not available to enable this.

Figure 3-17: RMSE, MAE and MAPE for hub height wind speed predicted using GPR model with and without stability information. The persistence model is shown for comparison. Error bars are not shown here to allow clarity. Error bars for the GPR models can be seen in Figure 3-16.



3.3.3 Significance of results in power output forecasting

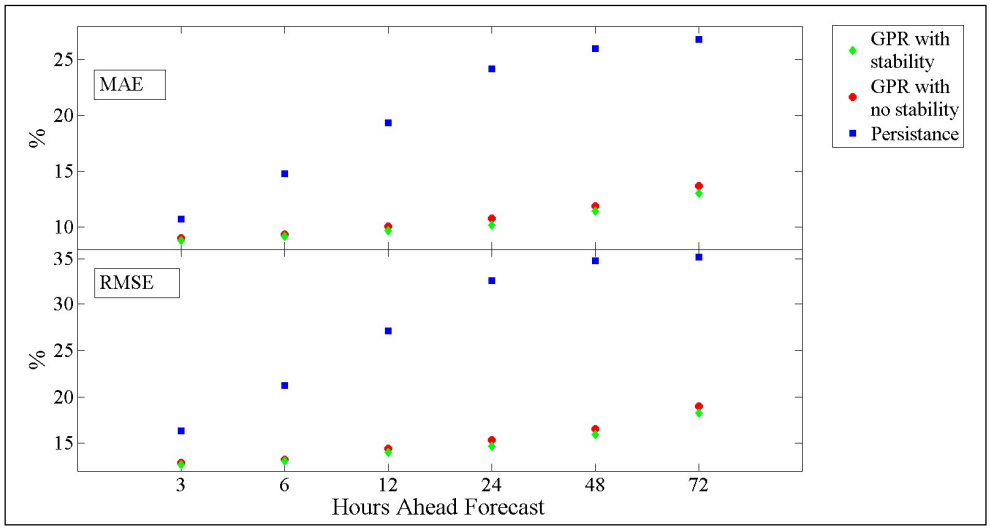
In Section 3.3.2, it was seen that using a GPR to predict hub height wind speed leads to a reduction in prediction error in comparison to using the persistence method. Additionally, a further reduction is seen when the data is split using PGT stability class at the time of observation. In order to establish whether the reduction in error seen in hub height wind speed prediction is sufficient to suggest a reduction in power output prediction error, predicted power output is calculated from the predicted wind speed using a wind turbine power curve.

In order to allow comparison between other models based on different datasets, the model errors are shown as a percentage of turbine capacity, giving normalised MAE (NMAE) and normalised RMSE (NRMSE). This allows model results from larger or smaller turbines to be compared in a meaningful way. The equations for NMAE and NRMSE were given in section 2.5.1.4.

Figure 3-18 shows the difference in NMAE and NRMSE between power output predicted from wind speeds using a persistence model, a simple GPR model and a

GPR model with stability classification. It shows a reduction in NMAE of between 2 and 12% for the power output predicted using wind speeds from the simple GPR model over a persistence model, and a further 0.5% for the GPR model with data subdivided by stability class. Additionally, a reduction of between 4 and 16% in NRMSE for the simple GPR model is observed in comparison to the persistence model with a further 0.5% improvement using the GPR model with data subdivided by stability classes.

Figure 3-18: MAE and RMSE normalised by turbine capacity, shown for simple GPR and GPR with stability information



Due to the power curve for this turbine not being available the method for predicting power output could be improved upon significantly. Furthermore, this is a very simplistic method power prediction which aims to see whether the improvement in hub height wind speed predictions seen in section 3.3.2 lead to improved wind power predictions. It shows that there may be some potential for improvement in power output forecast using the GPR model with data subdivided by stability classes to predict hub height wind speed. However, these improvements are quite small. This is because, looking at the power curve in Figure 3-8, a 1 ms^{-1} change in wind speed leads to, at most, a 225 kW change in power output. This is 15% of capacity. The difference in wind speed prediction error between the simple GPR model and

the GPR model with data subdivided by stability class is around 0.1 ms^{-1} which will lead to small differences in predicted power output using this method.

3.4 Conclusions

The motivation for this chapter has been to assess the performance of a hybrid numerical weather prediction model (NWP) and Gaussian process regression (GPR) model in predicting near surface wind speeds up to 72 hours ahead, and show how subdividing data using the PGT atmospheric stability class can improve model performance.

The results show that when the simple GPR model is used for 10 m wind predictions there is a reduction in MAPE for all forecast periods of 2% over the NWP wind speed predictions. When the GPR model is used with data partitioned by atmospheric stability there is a reduction in MAPE of 5% for forecasts made 3 hours ahead and 9% for forecasts made 72 hours ahead. This indicates that the GPR model with data partitioned by stability class leads to improved wind speed predictions over the NWP model. Particular improvements are seen at mountainous and coastal sites. Furthermore, using the GPR model using data partitioned by stability class for the prediction of hub height wind speeds lead to a reduction in MAPE of between 1 and 2% over the simple GPR model. It can also be seen that the improvements achieved using this model have a positive impact on wind power output predictions. Implementing the GPR model with data partitioned by stability class lead to a reduction in NMAE of 0.5% over the simple GPR model, and a reduction of between 2% and 12% in comparison to the persistence methods. In general the results seen for wind speed prediction are of comparable magnitude to those observed in other methods listed in Table 2-6, as discussed in Section 3.3.

The work so far demonstrates the potential of the method. If additional data was available it is possible that improvements could be made to the method shown. For example, as highlighted at the beginning of this chapter, due to the availability of forecasted weather variables, this work has so far relied upon the use of stability

class as calculated from observed weather variable rather than predicted stability class. In Chapter 4 this work is developed further to show a fully predictive model.

Whilst many methods for wind speed and power prediction exist, GPR has not been used widely for wind speed prediction. Furthermore, despite the numerous methods that exist, the impact of atmospheric stability on predictions is rarely considered. Because of this, the method provides a novel approach to forecasting and indicates promising results.

Chapter 4. Developing wind power predictions using a hybrid NWP and Gaussian process regression model.

4.1 Overview

In this chapter the hybrid numerical weather prediction model (NWP) and Gaussian process regression (GPR) model for near surface wind speed prediction developed in Chapter 3 is investigated using predicted atmospheric stability as an input variable, rather than subdividing data by atmospheric stability class observed at the time of prediction. Predicted atmospheric stability conditions are incorporated into the GPR model by calculating the Obukhov length using data obtained from the Met Office UKV and Euro4 models. The Obukhov length stability parameter is then used as an input variable in a hybrid NWP and GPR model. Results are shown for the prediction of near-surface wind speeds at 15 MIDAS sites across the UK and hub height wind speeds at one turbine site. To develop the method further the model is adapted to predict wind power output. The results show how this method can be used to predict wind power output for both an individual wind turbine and a whole wind farm. The results are shown for a single 1.5 MW turbine and for 22 wind farms across the UK.

In Chapter 3 a Gaussian process regression model for predicting near surface wind speeds was introduced. Furthermore, the potential benefit of subdividing data by PGT atmospheric stability class was demonstrated. Key atmospheric variables needed to predict stability class were not available in the publically available Met Office forecast data obtained from the UK government online data portal [151]. Because of this, the PGT stability class used to divide data in the GPR model in Chapter 3 was based on observed rather than predicted values. Whilst this meant that

the method did not give a true prediction, this allowed a preliminary investigation of the hypothesis that including subdividing data by atmospheric stability class within the GPR model could improve wind speed prediction accuracy. This chapter extends upon this work by calculating predicted atmospheric stability conditions based on data provided by the Met Office, and using this as an input variable in the GPR model for wind speed and power predictions. To calculate the predicted stability conditions estimates of other meteorological variables such as heat flux and frictional velocity are required, which are not routinely available. Further data is obtained from the Met Office to allow this. The work done in this chapter aims to show whether using predicted stability conditions calculated using a different stability parameter can also aid wind speed predictions. This is also developed further to show a model for prediction wind power output for both individual turbines and wind farms

There are few examples in literature of a GPR model being used for wind speed and power predictions. Examples include Chen et al. [71] and Hu et al. [160] which use a hybrid GPR and wavelet transform models to predict wind speed, and Chen et al. [73] where hybrid NWP and GPR model was used to predict wind power output. However, to the best of the author's knowledge there no examples wind speed or power forecasting literature where the Obukhov length stability parameter has been included as a model input. Therefore this presents a novel approach to wind speed and power forecasting.

The main objectives of this chapter are:

(i) To investigate whether using a predicted atmospheric stability parameter as an input variable in the GPR model leads to improved wind speed predictions over a GPR model without this input.
(ii) To develop the GPR model for wind power output prediction and test performance on both a small scale (1.5 MW turbine) and a larger scale (12 MW to 322 MW wind farms).
(iii) To consider how the GPR model for wind speed and power prediction, both with and without stability, compares to other models presented in literature.

This chapter will be presented as follows: In Section 4.2 the method employed in this chapter is introduced. This starts with an introduction to the hybrid GPR and NWP model focusing on the model inputs and outputs. This is followed by a discussion of the method used to predict atmospheric stability conditions and finally a discussion of the data used. In Section 4.3 the results are presented. This section is subdivided into 4 sections; near surface wind speed predictions, hub height wind speed predictions, power prediction for a single turbine and power prediction at a wind farm. Section 4.4 gives conclusions and highlights how this work will be continued.

4.2 Methodology

The model used for predicting wind speed and power output is a hybrid NWP and GPR model as outlined in detail in Chapter 3. In Chapter 3 the mathematical basis for the GPR model was introduced. The mathematical concepts remain the same in this chapter. The GPR model develops a relationship between the input variables **x** and target variable *y* of the form $y_i = f(\mathbf{x_i}) + \varepsilon_i$. The target variables and predictor variables differ depending on the application of the GPR model.

In Section 4.3.1 the hybrid NWP and GPR model is used for the prediction of 10 m wind speeds at 15 MIDAS sites across the UK. These sites are the same as those detailed in Section 3.2.4. In Section 4.3.2 the same model is used to predict hub height wind speeds for one wind turbine in the UK. Again, this turbine is as used in Chapter 3 and described in Section 3.2. The predictor variables for these prediction models are given by Equation 4-1 for the model including atmospheric stability, and by Equation 4-2 for the model without.

$$\mathbf{X}_t = \left[\mathbf{m}_{t,k}, \mathbf{y}_{t-k-1}, \mathbf{y}_{t-k-2}, \mathbf{y}_{t-k-3}, \frac{1}{\mathbf{L}_{t,k}}\right] \text{ for } t = 4, \dots, n \qquad \text{Equation 4-1}$$

$$\mathbf{X}_t = \left[\mathbf{m}_{t,k}, \mathbf{y}_{t-k-1}, \mathbf{y}_{t-k-2}, \mathbf{y}_{t-k-3}\right] \text{ for } t = 4, \dots, n \qquad \text{Equation 4-2}$$

where:

$t =$ time of prediction

$k =$ forecast horizon

$\mathbf{m}_{t,k}$=Met office 10 m wind speed forecast for time, t at k hours ahead

$\mathbf{y}_t =$ Observed wind speed at time t (ms^{-1})

$\mathbf{L}_{t,k} =$ Stability parameter predicted for time t, k hours ahead.

In each case the target variable at time *t*, is the observed wind speed $\mathbf{y}_t$. The observed wind speed is dependent on the application of the model. In Section 4.3.1 the observed wind speed is 10 m wind speed at 15 MIDAS sites across the UK. In Section 4.3.2 the observed wind speed is hub height wind speed. This data was obtained from the wind turbine operator for a 1.5 MW wind turbine. The hub height was 65 m above ground and the Met Office forecast site used is located 8 km from the turbine in question. The stability parameter, $\mathbf{L}_{t,k}$ is the Obukhov length which is discussed in Section 4.2.1. The Met Office 10 m wind speed forecasts, $\mathbf{m}_{t,k}$ are taken from the Met Office NWP model described in Chapter 3, available from the UK governmental public data website [151].

Next, in Section 4.3.3, the hybrid GPR and NWP model was used for the prediction of hub height wind speed. The wind power output data used here was for the same turbine used in Section 3.3.2, a 1.5 MW turbine located in the UK. For this turbine, observed wind speed data is available in addition to power output data therefore this is incorporated into the model. A two-stage model is used to first predict hub height wind speed and then use this to predict power output. This allowed uncertainties to be minimised. In stage one the hub height wind speeds are predicted using the model described by Equations 4-1 and 4-2. The inputs for stage 2 are given by Equation 4-3 and 4-4. The target variable, $\mathbf{y}_t$ in this case is power output for the 1.5 MW wind turbine.

$$\mathbf{X}_t = \left[\mathbf{g}_{t,k}, \mathbf{y}_{t-k-1}, \mathbf{y}_{t-k-2}, \mathbf{y}_{t-k-3}, \frac{1}{\mathbf{L}_{t,k}}\right] \text{ for } t = 4, \dots, n \qquad \text{Equation 4-3}$$

$$\mathbf{X}_t = \left[\mathbf{g}_{t,k}, \mathbf{y}_{t-k-1}, \mathbf{y}_{t-k-2}, \mathbf{y}_{t-k-3}\right] \text{ for } t = 4, \dots, n \qquad \text{Equation 4-4}$$

where:

$t =$ time of prediction

$k =$ forecast horizon

$\mathbf{g}_{t,k} =$ GPR hub height wind speed forecast for time, t at k hours ahead

$\mathbf{y}_{\boldsymbol{t}} =$ Observed wind power at time t (ms^{-1})

$\mathbf{L}_{\boldsymbol{t,k}} =$ Stability parameter predicted for time t, k hours ahead.

Finally, in Section 4.3.4 the hybrid NWP and GPR model is used for the prediction of wind power output at wind farms. Wind power output data was obtained for 22 sites across the UK as shown in Figure 4-1. This data was obtained on a half hourly basis was obtained from the Balancing Mechanism Reporting System (BMRS) [161]. At these sites observed wind speed data was unavailable, therefore forecasts were based on predicted wind speeds and prior power output. The inputs used are given in Equation 4-5 and Equation 4-6. The wind farms used are described in Section 4.2.2. In this case, the Met Office 10 m forecasts are taken from the closest available MIDAS site to the wind farm. Only wind farms where there was a MIDAS sites within 10 km were considered.

$$\mathbf{X}_t = \left[\mathbf{m}_{t,k}, \mathbf{y}_{t-k-1}, \mathbf{y}_{t-k-2}, \mathbf{y}_{t-k-3}, \frac{1}{\mathbf{L}_{t,k}}\right] \text{ for } t = 4, \dots, n \qquad \text{Equation 4-5}$$

$$\mathbf{X}_t = \left[\mathbf{m}_{t,k}, \mathbf{y}_{t-k-1}, \mathbf{y}_{t-k-2}, \mathbf{y}_{t-k-3}\right] \text{ for } t = 4, \dots, n \qquad \text{Equation 4-6}$$

where:

$t =$ time of prediction

$k =$ forecast horizon

$\mathbf{m}_{t,k} =$ Met office 10 m wind speed forecast for time, t at k hours ahead

$\mathbf{y}_{\boldsymbol{t}} =$ Observed wind power at time t (ms^{-1})

$\mathbf{L}_{\boldsymbol{t,k}} =$ Stability parameter predicted for time t, k hours ahead.

This gives the model inputs and outputs used in each section of this chapter. Next, the process for calculating a stability parameter, the Obukhov length is discussed, before returning to discuss in more detail some of the data used for predictions.

4.2.1 Atmospheric stability using Obukhov length

As outlined in Section 3.2.2 atmospheric stability is a measure of the atmospheric tendency to encourage or deter vertical motion [158]. This vertical motion can be driven by convective or mechanical processes, and most methods for parameterising atmospheric stability use an estimate of convective turbulence and mechanically driven turbulence. In Chapter 3 this was done using the Pasquill-Gifford-Turner stability classification, which uses solar insolation and cloud cover to estimate convective turbulence and wind speed to estimate mechanical turbulence. This method gives a rudimentary estimation of an atmospheric stability class, from highly unstable to extremely stable. The PGT method has the advantage that it can be calculated based on commonly available meteorological parameters. Because of this, the PGT stability was used to give an initial indication of the merit of including stability class in a prediction of wind speed. Alternatively, the Obukhov length is a parameter which is proportional to the height above the surface at which convective factors first dominate over the mechanical production of turbulence [25]. Above this height convection dominates, below this shear forces dominate. The Obukhov length is denoted by L and is given in metres. L can be calculated using a buoyancy term and a shear production term, given by Equations 4-7 and 4-8. Through additional data obtained from the Met Office UKV and Euro4 NWP models, the Obukhov length was calculated and used to obtain a predicted stability parameter.

$$b = \text{bouyancy term} = \frac{g}{\theta}\frac{H_{sen}}{\rho C_p} \qquad \text{Equation 4-7}$$

$$s = \text{shear term} = \frac{u_*^3}{k} \qquad \text{Equation 4-8}$$

where:

g = Acceleration due to gravity = 9.81 ms^{-2}
θ = Surface temperature (K)
H_{sen} = Sensible surface heat flux (Wm^{-2})
C_p = Specific heat of air ($kJ\ kg^{-1}K^{-1}$)

ρ = Air density (kg m^{-3})

u_* = Friction velocity (ms^{-1})

k = Von Karman constant = 0.4

Using Equations 4-7 and 4-8 L can be calculated using Equation 4-9 [162]

$$L = \frac{s}{b} = \frac{\theta \rho C_p \, u_*^3}{g \, H_{sen} k} \qquad \text{Equation 4-9}$$

Within this, the specific heat of air is 1005 Jkg^{-1}K^{-1} and air density is 1.225 kg m^{-3} [25]. The von Karman constant is a dimensionless parameter, an appropriate range is generally given by $0.35 \leq k \leq 0.42$ [25]. Sensible heat flux is a measure of heat transfer between the earth's surface and the atmosphere. From the Obukhov length a non-dimensional stability parameter is given by Equation 4-10, where z is height above ground and L is the Obukhov length.

$$\zeta = z/L \qquad \text{Equation 4-10}$$

A relationship between L and PGT stability classes defined by Gryning et al. [163] is shown in Table 4-1. However, Golder [164] states that the relationship between Obukhov length and stability classes is site specific. In the hybrid NWP and GPR model the parameter $1/L$ is used as the stability parameter as for each application z will be constant and can be omitted.

Table 4-1: Obukhov length and stability classes

Stability class	Obukhov length
Highly unstable	-100<*L*<-50
Moderately unstable	-200<*L*<-100
Slightly unstable	-500<*L*<-200
Neutral	|*L*|>500
Slightly stable	200<*L*<500
Stable	50<*L*<200
Extremely stable	1-<L<50

4.2.2 Training and validation data

As in the previous chapter, wind speed predictions are used to inform the hybrid model. These are taken from the Met office's NWP model, as outlined in Section 3.2.1. Further predictions of atmospheric variables are obtained from the Met Office to allow a prediction of atmospheric stability via the Obukhov length, as outlined in Equation 4-9. These included sensible surface heat flux, surface temperature and the x, y components of surface wind stress. For shorter prediction horizons (up to 36 hours ahead) the prediction of these variables are obtained from the Met Office UKV model. For longer prediction horizons (up to 72 hours ahead) the predicted values are obtained from the Met Office Euro4 model. The Euro4 model is run 4 times per day in contrast to the 8 times per day that the UKV model is run. The data used for analysis was for 1st Jan 2015 – 31st Dec 2015. For the model development the data was split into a training and a test dataset to ensure the forecast is independent. The training set contains data from the first 9 months of the data set and the test set contains data from the final 3 months.

Near-surface wind speed observations used in the model are the same as in Chapter 3, with the same 15 sites used to explore the model results. Identical hub height wind speed and power data are used to that in Chapter 3. That is power output and hub height (65 m) wind speed at one 1.5 MW turbine located in the UK. This chapter also investigates the use of the hybrid NWP and GPR model for predicting power output for wind farms. Power output data at 22 wind farms across the UK is obtained from the publically available Balancing Mechanism Reporting System (BMRS) [161]. Data is available in half hourly settlement periods for individual generation units. The location of these wind farms is shown in Figure 4-1. The site details are given in Table 4-2. Overall there are 18 standard sites (onshore sites with 1 year data set), 2 offshore sites and 2 sites where only a short data set was available.

For the offshore sites, the NWP wind speed predictions are taken from the nearest available sites. These sites are onshore, coastal locations. The difference in surface

roughness when moving from land to sea and the availability of heat and moisture can affect wind speeds. It might be expected that wind speed would increase with distance offshore due to a reduction in surface roughness. However, Barthelmie et al. [165] state that vertical and horizontal changes in wind speed in coastal regions are significantly influenced by atmospheric stability conditions. For example, in stable conditions wind speeds may only increase lightly (or even decrease) with increasing distance from the shoreline whilst in unstable conditions the increase in wind speed is larger than would be expected for a reduction in surface roughness [30]. These effects can be important for the prediction of wind power output at wind farms up to around 70 km offshore [165]. Because of this, the prediction results for offshore wind farms are separated from onshore wind farms.

Figure 4-1: Wind farm locations.

Table 4-2: Wind farms location details

Wind farm name	Capacity (MW)	Location
An Suidhe	19	Argyll & Bute
Baillie	52	Highland
Berryburn	66	North Cairngorms
Black Law	134	North Lanarkshire
Braes of Doune	74	Stirling
Clyde South	128	South Lanarkshire
Dalswinton	30	Dumfries and Galloway
Edinbane	42	Skye
Fallago rig	144	Scottish Borders
Farr 2	47	Invernesshire
Glens of Foundland	26	Aberdeenshire
Gordonstown	12	Aberdeenshire
Griffin 2	102	Perth and Kinross
Hadyard Hill	130	South Ayrshire
Minsca	38	Dumfries and Galloway
Toddleburn	28	Scottish Borders
Whitelee	322	East Ayrshire
Whitelee Extension	238	East Ayrshire
Offshore Sites		
Wind farm name	Capacity (MW)	Location
Burbo Bank	90	Offshore - off Merseyside
Barrow Offshore	90	Offshore - off Cumbria
Short data sets		
Wind farm name	Capacity (MW)	Location
Beinn Tharsuinn	30	Ross & Cromarty
Mark Hill	56	South Ayrshire

For the short datasets, only 3 months of data was available. In this case, 2 months was used for training and 1 month for testing. These short datasets are of interest as for newly established wind farms limited data may be available. In this case, generating a wind power forecast based on limited data is important. This can allow

new wind farms to bid in electricity markets or schedule maintenance. These two datasets are used to assess how the hybrid NWP and GPR model performs for a short dataset.

In addition to the lack of wind speed data available at the wind farms under investigation, there are some limitations to the power data which are worth considering when evaluating the predictive ability of this model. The wind power data provided is an aggregate wind power figure given on a half hour basis. There is no additional information regarding turbine operation, maintenance or any other conditions which may lead to suboptimal wind farm operation in any specific half hour period. In order to eliminate some issues associated with data some filtering is applied. Where long periods (longer than 24 hours) of zero power output were recorded, data was removed, particularly if power recorded immediately before or after was high. Further information detailing operational conditions at the wind farm could eliminate some of these issues. However, at the time of analysis, this data was not available.

4.3 Results

In this section, the results for the prediction of near surface wind speeds, the prediction of hub height wind speed, the prediction of power output for a single turbine and the prediction of wind farm power output are shown. The models employed in this chapter builds on the GPR model introduced in Chapter 3.

4.3.1 Near-surface wind speed prediction

In this section the results for a GPR model with and without stability are shown for the prediction of near-surface wind speed. The model set-up was described in Section 4.2 by Equations 4-1 and 4-2. Predicted wind speeds are compared to MIDAS observations of wind speeds at 15 sites across the UK. RMSE, MAE and MAPE are shown for these sites, the equations for which were provided by Equation 2-12, Equation 2-13 and Equation 2-14 given in Section 2.2.7. The results for a

simple GPR model and a GPR model with Obukhov length as an input parameter are shown alongside the Met Office NWP performance. This allows the results to be compared to a benchmark model. In Figure 4-2 it can be seen that on average over the 15 sites under evaluation there is some reduction in error seen when the standard GPR model is implemented. However, there is no additional reduction in error achieved when implementing the GPR model with Obukhov length. The errors seen are largely the same for both GPR models. In Figure 4-2 it can be seen the reduction in error for the GPR models is larger for MAPE than for RMSE and MAE. In terms of MAE and RMSE, there is a bigger difference in the errors between the NWP model and the GPR models at 3 hours ahead than for further in advance. As highlighted in Chapter 3, the model performance is highly site specific.

Figure 4-2: Average RMSE, MAE and MAPE for all 15 MIDAS sites.

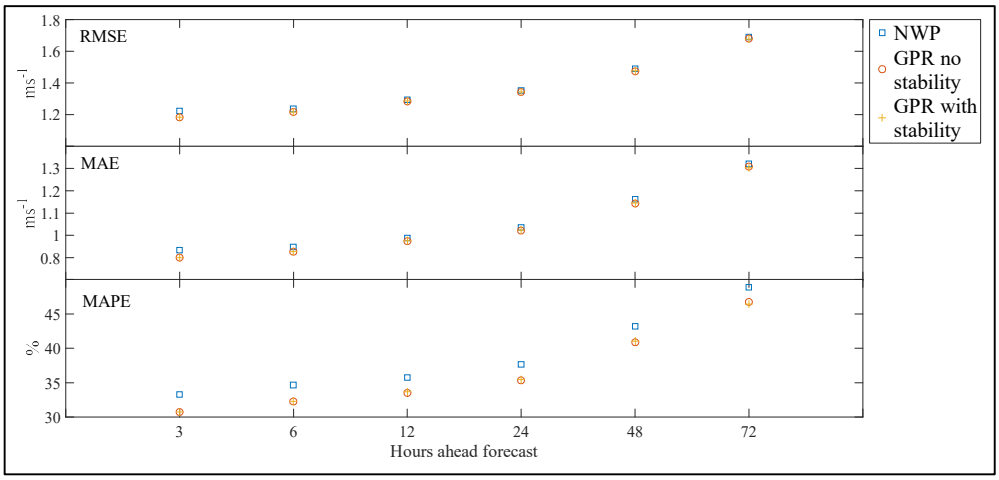


Figure 4-3 and Figure 4-4 show the errors seen for the NWP model and both GPR models at 4 of the 15 sample sites. It can be seen in Figure 4-3 and Figure 4-4 that the difference in errors between the NWP model and the GPR models is different at each of the sites. For example, at site 1 there is not much difference between the models and at site 9 the difference is more pronounced.

Figure 4-3: MAE (with 95% confidence interval) for 4 sample MIDAS sites shown for GPR models both with and without stability, and the NWP.

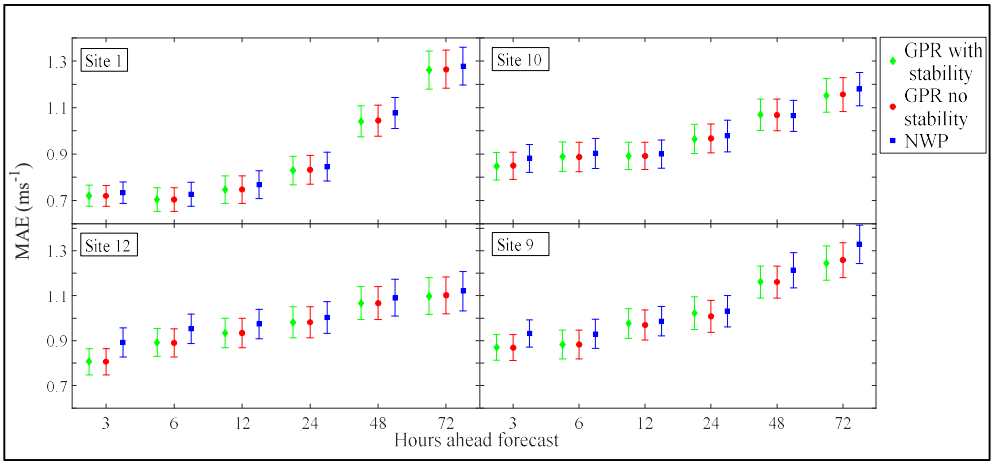
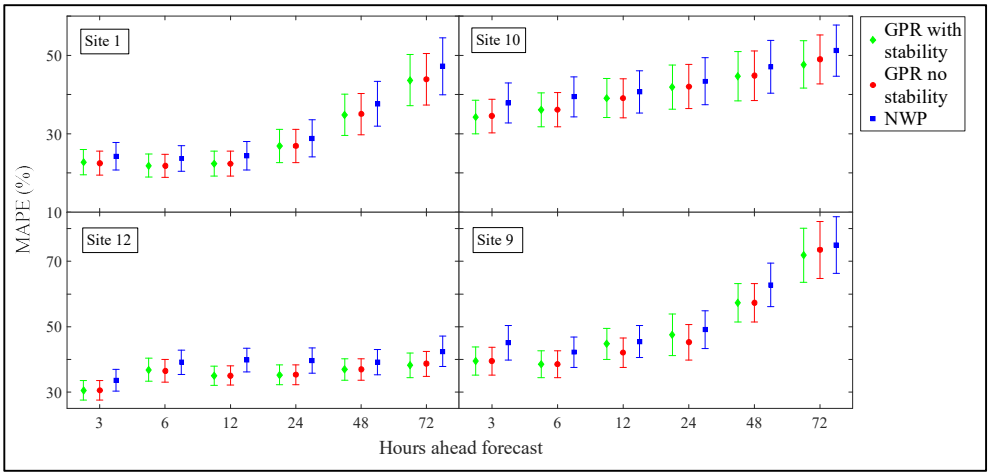


Figure 4-4: MAPE (with 95% confidence interval) for 4 sample MIDAS sites shown for GPR models both with and without stability, and the NWP.



In addition to considering the model performance, the GPR model hyperparameters can be used to show the relevance of the predictor variables used in the two models. As shown in Equation 4-1 and Equation 4-2 NWP predicted wind speed and the previous three hours of wind speed observations were used as inputs for both models, and the Obukhov length was used in the second model. The length scale parameters of the covariance function can be used to see the relevance of each parameter. It was seen that the NWP predicted wind speed and the previous 3 hours

wind speed were the most relevant predictor variables for forecasts 3 hours ahead, whilst for forecasts further in advance the NWP wind speed prediction becomes the most relevant parameter. In particular, for forecasts further than 6 hours ahead the wind speed 3 hours ago has little relevance. In addition, the Obukhov length was less relevant that that the NWP wind speed predictions and the previous 3 hours of observed wind speed.

Whilst the preliminary results shown in Chapter 3 suggested that including atmospheric stability in the predictions of near-surface wind speed lead to improved model results, these results are not replicated here. When using Obukhov length no difference is seen in the results from the GPR model with and without stability. This could be because the predictions of $1/L$ are inaccurate. The variables required to calculate $1/L$ are not measured at the MIDAS sites hence it is not possible to establish whether the predictions of $1/L$ reflect the observed conditions. This could impact the results. It still holds that whilst using the GPR model to predict wind speeds without stability there is some reduction in error over the wind speeds predicted using the Met Office NWP model. The improvements over the NWP model are modest. However, even small improvements in wind speed predictions could result in improvements in wind power modelling.

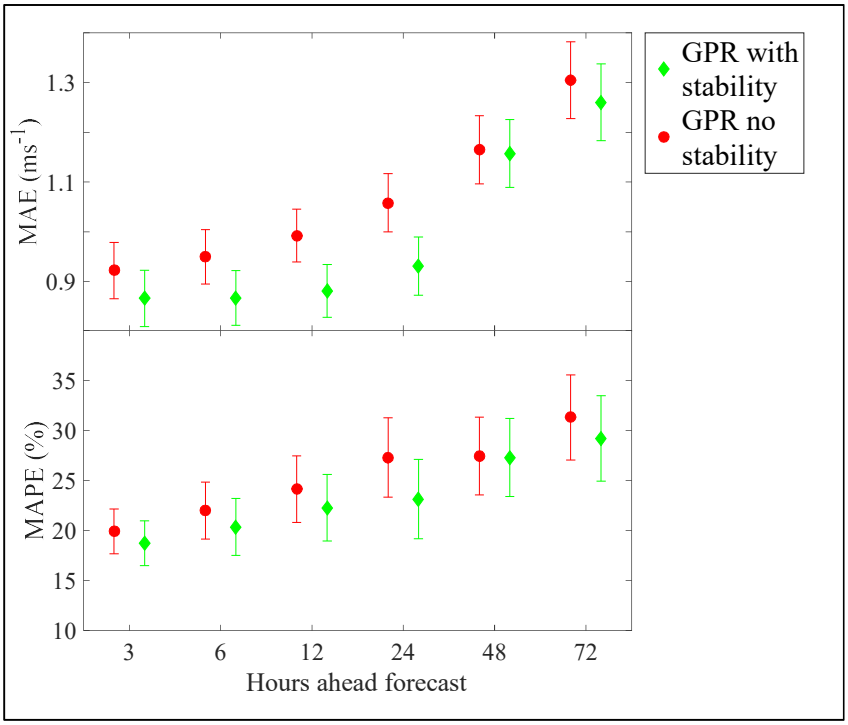
4.3.2 Hub height wind speed prediction

In this section, the results of a wind speed prediction model for hub height wind speeds are shown. As in Chapter 3, the results are only shown for one turbine as commercial sensitivity limited the data which was available. The model set-up was described in Section 4-2 by Equations 4-1 and 4-2. Figure 4-5 shows MAE and MAPE for both GPR prediction models. It can be seen for the prediction of hub height wind speeds significantly lower errors are seen when the GPR model including Obukhov length is implemented. The difference between the two models increases from around 2% at 3 hours ahead to 5% at 24 hours ahead and reduced to around 1% difference above 24 hours in advance. This reduction could be due to changes in the accuracy of stability parameter predictions with the transition

between UKV predictions and Euro4 predictions seen between 24 and 48 hours. It is difficult to verify this as the variables needed to calculate Obukhov length are not included in observation data sets. Figure 4-6 shows the results in comparison to the benchmark model, this time the persistence model is used as a benchmark. Both models achieve a significant reduction in errors over the persistence method in terms of RMSE, MAE and MAPE. The improvements when using the GPR models are smallest at 3 hours ahead and gradually increase as the forecast horizon increases. This is largely due to the persistence method quickly reducing in accuracy as the forecast horizon increases. However, there is still an improvement at 3 hours ahead where the persistence model can be seen as a valid model.

The results shown here indicate that the GPR model with Obukhov length as an input parameter provides hub height wind speed predictions with higher accuracy than the simple GPR model. This is despite the fact that this model does not outperform the simple GPR model for 10 m wind speed predictions. This could be because the inclusion of the Obukhov length stability parameter aids the translation from 10 m met office predicted wind speeds up to hub height wind speed. The change in wind speed with height and the effect of stability on this is well documented and was discussed in Section 2.1.2.

Figure 4-5: MAPE and MAE (with 95% confidence interval) for hub height wind speeds predicted with a simple GPR model and a GPR model with stability.

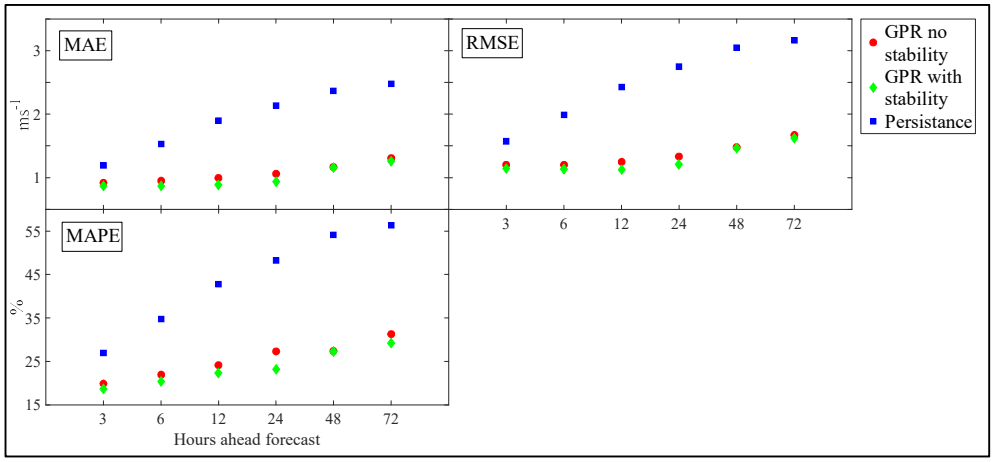


In comparison to the model for predicting near-surface wind speeds, the length scale hyperparameters for the covariance function in this model indicated that the Obukhov length was more relevant than the observed wind speed 2 and 3 hours before the forecast, particularly for forecasts up to 12 hours ahead. However, the Obukhov length parameter was less relevant than the NWP forecast, particularly for forecasts further in advance.

In addition to the comparison to the persistence method seen in Figure 4-5 it is beneficial to compare the results seen here to other methods for wind speed prediction. Table 2-6 provides a summary of deterministic wind speed and power prediction methods. RMSE, MAE and MAPE seen here for the GPR model with stability is lower than that seen for predictions at 1 hour ahead given by Li and Shi [60], Li et al. [61] and Torres et al. [56]. MAE in listed for these methods at 1 hour ahead ranges from 0.9 – 1.14 ms^{-1}, in contrast to 0.86 ms^{-1} seen at 3 hours ahead for

the GPR method with stability. RMSE ranges from 1.2 – 1.5 ms^{-1} at 1 hour ahead for the methods listed, in contrast to 1.14 ms^{-1} seen at 3 hours ahead for the GPR method with stability. The results shown by Chen et al. [71] indicate lower errors at 1 hour ahead. However, the results seen at 3 hours ahead are higher than those shown for the GPR model in the same time frame.

Figure 4-6: RMSE, MAE and MAPE for hub height wind speed predicted using GPR model with and without stability information. Persistence model shown for comparison. Error bars are not show here to allow clarity.



Further ahead (24-72 hours ahead) the results seen for the method presented here show lower MAE and RMSE than the results shown by Chen et al. [71], Kavassseri and Seetharam [57] and Louka et al. [72]. The results shown by Kavassseri and Seetharam indicate much larger errors in the persistence forecast, indicating a complex site therefore the results may not be comparable.

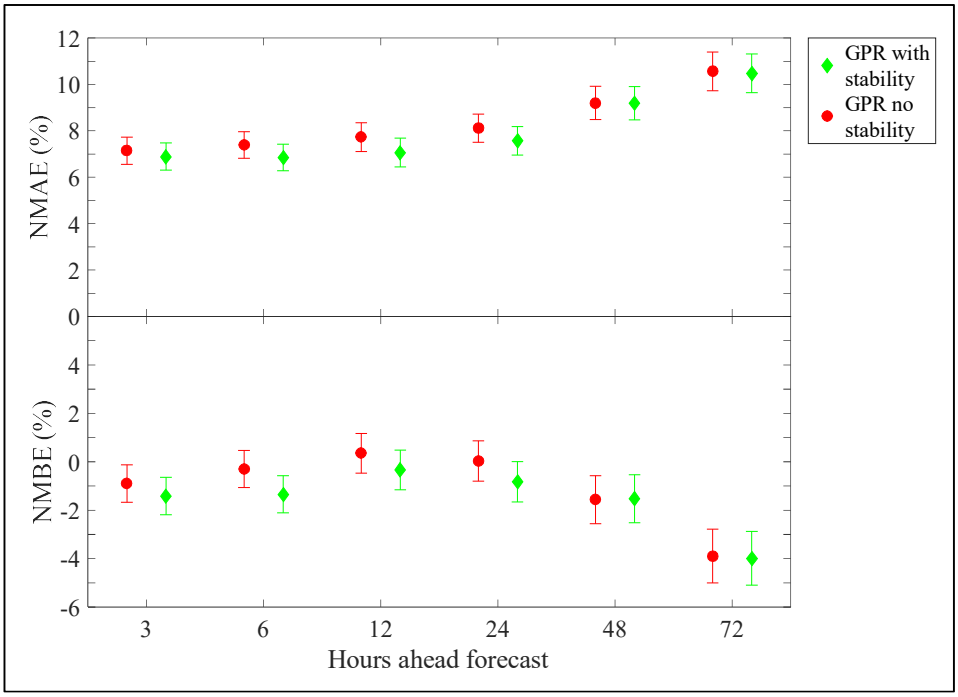
It is impossible to use these results to show one models superiority over another as different data sets are used. However, it does indicate that the GPR model provides predictions with a similar level of accuracy to other state of the art wind forecasting techniques.

The results suggest that including Obukhov length to parameterise stability conditions can improve predictions of hub height wind speed. However, it is difficult to establish whether this is a universal or a site-specific result as there is only one data set available. If the additional predicted atmospheric variables needed to calculate Obukhov length were more widely available these results could be tested for hub height wind speed prediction at additional locations. Given the advantages seen here of using atmospheric stability in hub height wind speed predictions, the next sections continues by developing a method for wind power prediction at an individual turbine.

4.3.3 Power prediction for a single turbine

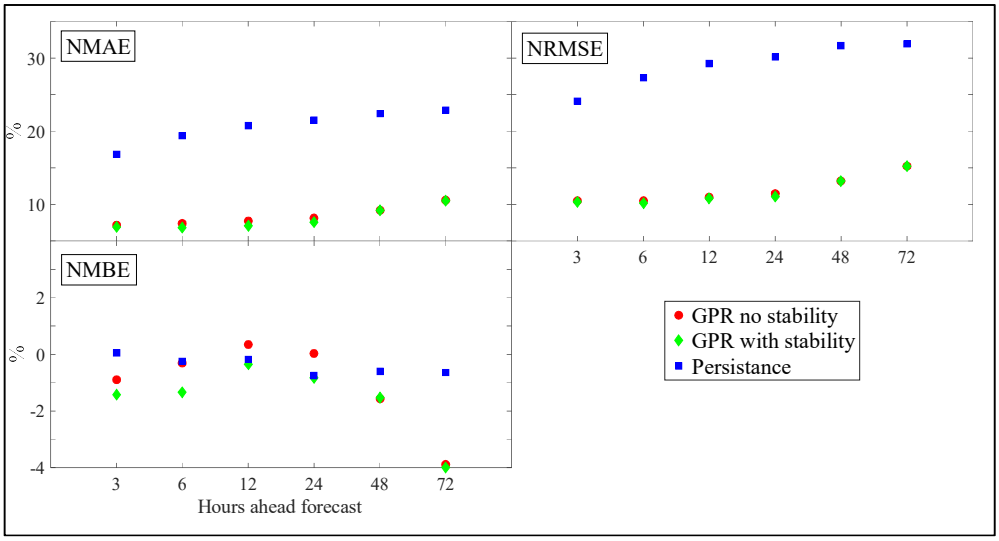
The results shown are for the prediction of wind power for one turbine in the UK. The model set-up was described in Section 4-2 by Equations 4-3 and 4-4. Results are compared using MBE, MAE and RMSE normalised by turbine capacity (equations found in Section 2.2.4). Figure 4-7 compares NMAE and NMBE for the GPR models both with and without the Obukhov length being included. NMAE is reduced when using the GPR model with Obukhov length in comparison to the GPR model with no stability. The reduction in error is more significant for forecast horizons between 3 and 24 hours in comparison to predictions further ahead. The reduction in error is modest: between 3 and 24 hours ahead a reduction in NMAE of between 0.5% and 0.7% of capacity is seen. For 48 and 72 hours ahead this is reduced to 0.1%. The reduction in difference in the predictions with the two models is a result of the changes in accuracy observed in hub height wind speed prediction between 24 and 48 hours in advance. This was observed in Figure 4-4. NMBE is closer to zero for the GPR model with stability. However, both GPR models show some under prediction overall. In Figure 4-7 NMAE, NMBE and NRMSE are shown in comparison to the persistence model. Both GPR models show between 10% and 14% reduction in NMAE over the persistence method and between 14% and 19% reduction in NRMSE.

Figure 4-7: NMAE and NMBE (with 95% confidence interval) for single turbine power predicted with a simple GPR model and a GPR model with stability.



The results here can also be compared to results seen in literature. In Table 2-6 results are shown for 4 studies which attempt to predict wind power on a variety of timescales. 3 of these show NMAE. Chen et al. [73] show results for 1-24 hours in advance, giving NMAE of 7.5% – 11.1% and NRMSE of 11.69% – 15.96%. This is larger than the 6.8% – 7.5% NMAE and 10.2 – 11% NRMSE seen for this method. The improvement over persistence is also more significant. The results shown in Figure 4-7 and Figure 4-8 also show lower NRMSE and NMAE than the results shown by Louka et al.[72], Ramirez-Rosado et al. [62] and Shu et al. [74]. NMBE is not presented by most of these authors and is therefore not compared.

Figure 4-8: NMAE, NMBE and NRMSE for single turbine power prediction. Results are shown for simple GPR and GPR with stability in comparison to a persistence prediction.



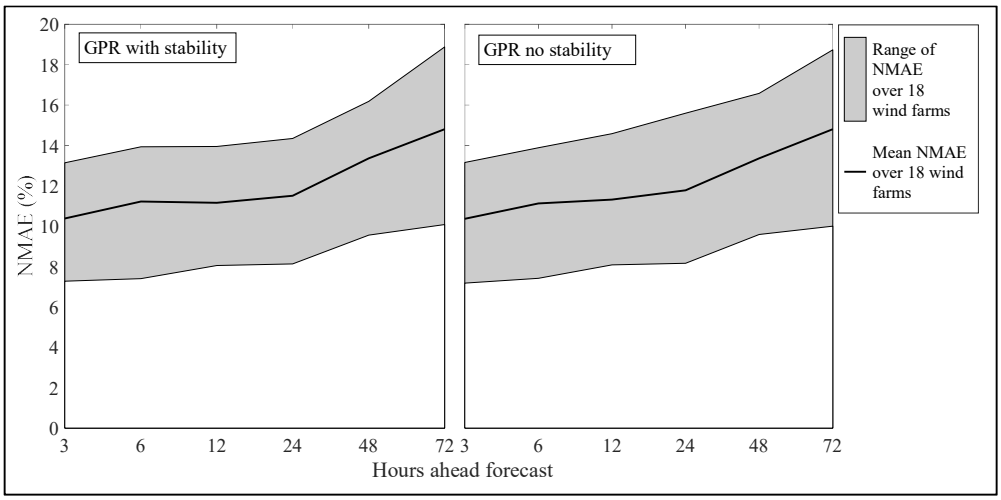
For individual wind turbine predictions, the length scale hyperparameters of the covariance function indicate that the Obukhov length stability parameter is more relevant as the observed data 2 and 3 hours before the prediction for forecasts over 6 hours ahead. For forecasts 3 hours ahead the Obukhov length stability parameter is more relevant than the observed data from 3 hours prior to the forecast but less relevant than the observed data 2 hours prior to the forecast. In all forecasts, the NWP prediction is the most informative predictor variable. This suggests that the Obukhov length has some value in the prediction of individual turbine wind power output.

It can be seen here that there is some benefit in including atmospheric stability in wind power predictions. However, the difference between the models is small and it would be useful to obtain further data to test this result more rigorously. Prediction power output from an individual turbine can be useful for small businesses and other stakeholders in renewable energy projects. However, the prediction of power output for a full wind farm is of greater commercial value. Therefore, this work is extended in the next section to predict wind power output for a selection of UK wind farms.

4.3.4 Wind farm power prediction

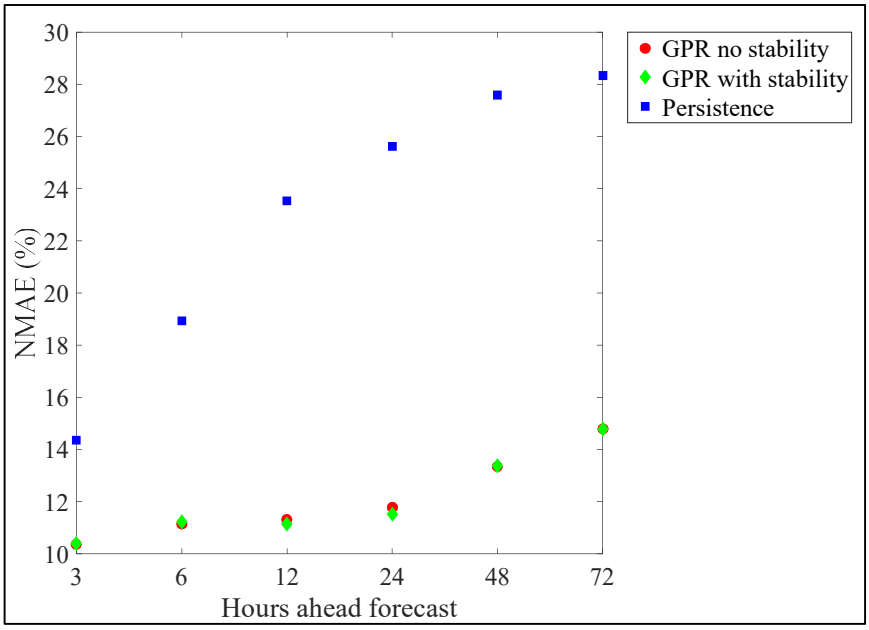
The results presented in this section evaluate the GPR model for 22 wind farms across the UK, with the locations shown in Figure 4-1. Figure 4-9 shows NMAE for 18 standard sites. The model set-up was described in Section 4-2 by Equations 4-5 and 4-6. It can be seen in Figure 4-9 that there is limited difference in the NMAE seen using a GPR model with stability in comparison to the GPR model without stability. The average NMAE over 18 sites is shown to be 10.2% of capacity at 3 hours ahead rising to 14.8% at 72 hours ahead. The difference ranges from no difference (at 3 hours ahead) to a 0.2% improvement when atmospheric stability is included (seen at 24 hours ahead). The lowest and highest errors seen at any one site are also shown in Figure 4-9. At 3 hours ahead NMAE ranges from 7.2% -13.1% of capacity, whilst at 72 hours ahead NMAE ranges from 10.0% to 18.9% of capacity. In Figure 4-10 NMAE for the 18 standard sites is shown in comparison to the persistence model. It can be seen here that the GPR models (both with and without stability) are around 4% better than persistence at 3 hours ahead, rising to 12% better at 12 hours ahead and 14% better at 72 hours ahead. At 3 hours ahead there is a smaller improvement over persistence than seen for a single turbine power forecast, most likely because the wind farm power predictions do not use hub height wind speed predictions. This introduces additional uncertainty and difficulties into the prediction model.

Figure 4-9: NMAE for 18 standard sites wind farms



These results can also be compared to the methods listed in Table 2-6. Chen et al. [73] present a hybrid GPR and NWP model reporting results for forecasts from 1 – 24 hours ahead, with an NMAE of between 7.5 and 11.1%. The model given by Shu et al. [74] is a two-stage hybrid model with Bayesian clustering and support vector regression (SVR). It presents results for forecasts from 1 – 48 hours ahead, with an NMAE of between 7% and 15%. The model presented by Louka et al. [72] applies a Kalman filter to post process NWP predictions and shows results for forecasts between 24 and 120 hours ahead. This shows an NMAE of 11% at 24 hours ahead and 12% at 72 hours ahead. In comparison to these three methods the GPR model shown here performs well. The average NMAE for both GPR models at 3 hours ahead is 10%, which is around 3% higher than that shown at 1 hour ahead by Chen et al. and Shu et al. The results shown in Figure 4-9 are similar to those presented by Shu et al. at 48 hours ahead, however are 2% worse than those shown by Louka et al. at 72 hours ahead. There are some sites at which the GPR model outperforms all the models detailed above. However, without using the same dataset it is difficult to show the overall superiority of one model over another.

Figure 4-10: NMAE for 18 standard sites wind farms in comparison to the persistence model



The results shown in Figure 4-9 and 4-10 indicate that a hybrid NWP and GPR model can offer good predictions of wind power output even without wind speed predictions at the exact site or observed wind speed at the wind farm. This could be of value for wind farm operators as these forecasts do not require large amounts of external data which could be expensive to obtain. However, it is worth noting that this model would most likely perform much better if wind predictions which were more finely tuned to the specific site were used. In addition, the availability of hub height wind speeds at a wind farm would enable better analysis of results and further tuning of the model.

As detailed in Section 4.2.3 to predict wind power output wind speed predictions obtained from the Met Office are used. The wind farms where predictions are made are matched with the closest Met Office prediction site. Only wind farms where there is a Met office site within 10 km were used. Whilst it might be better to use onsite 10 m wind speed predictions to inform the power output prediction model in some cases these are unavailable or would be costly to obtain. In which case it is

valuable to consider whether the location of the Met office prediction site in comparison to the wind farm might impact upon model performance. Figure 4-11 explores the relationship between the distance between Met office sites and the wind farm and average NMAE in wind power predictions. This figure and the R-squared value of 0.082 suggests that there is no linear relationship between distance and error. This suggests that the distance between the wind farm site and the Met Office site does not result in excessive deterioration of model results and good power output predictions can be made using offsite wind speed predictions. This is useful as site specific wind speed predictions may be costly to obtain. In Figure 4-12 the relationship between Met office sites elevation and NMAE is explored. In general, predictions can be less accurate at mountainous sites or those with complex terrain. Here it can be seen that there is no increase in model error at mountainous sites in comparison to flatter, less complex sites. This suggests that good wind power predictions can be made from forecast locations on complex terrain, allowing a large range of sites to be considered for future predictions.

Figure 4-11: Correlation between NMAE and distance between prediction-observation sets

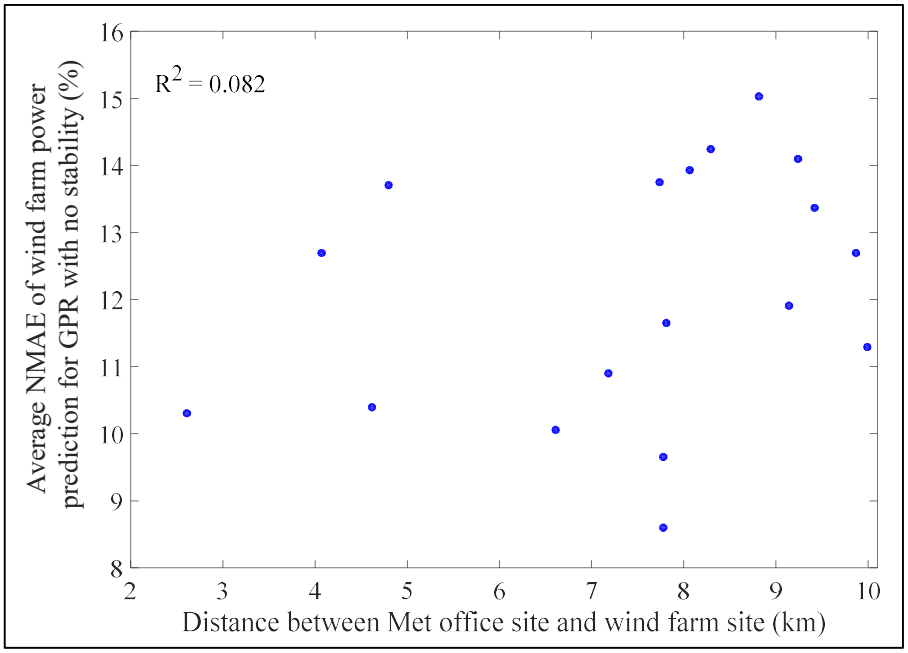
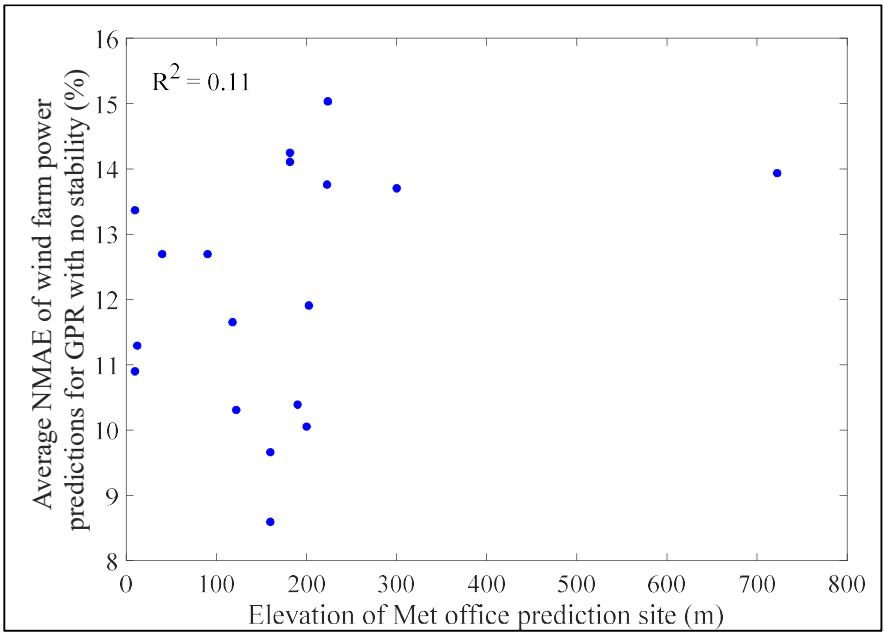


Figure 4-12: Correlation between NMAE and elevation at Met office prediction sites



In Figure 4-13 NMAE is shown for the 2 offshore sites considered. For both offshore sites the Met office predictions used were the nearest onshore site. In this figure NMAE is shown for both GPR models and the average NMAE over 18 standard sites is shown for comparison. It can be seen here that whilst there is limited difference between the GPR model with and without stability for the 18 standard sites, for both offshore sites the difference is larger, with the GPR model with Obukhov length stability parameter better predictions.

For Burbo Bank wind farm there is between 0.6% and 1.9% reduction in NMAE for all forecast horizons when the GPR with atmospheric stability is used in comparison to the GPR model with no atmospheric stability. For Barrow offshore wind there is between 0.5% and 1.5% reduction in NMAE. The GPR model with stability has higher NMAE than the average seen for all 18 standard sites, indicating some difficulty in making predictions at offshore locations. However, the difference between the GPR model with and without stability was more pronounced. In Figure 4-14 NMAE for the GPR models is shown alongside the persistence method for the two offshore sites. The improvement over persistence is quite similar to that seen in Figure 4-10 for the standard sites. In addition, the length scale parameter in the covariance function associated with the Obukhov length indicated that the Obukhov length is more important when predicting wind power output at offshore sites. Previous work by Barthelmie et al. [165] suggested that atmospheric stability conditions impact the difference in wind conditions between onshore and offshore sites. The results shown here indicate that using the Obukhov length stability parameter as an input parameter enables the difference in wind conditions between onshore and offshore sites to be modelled more accurately. It would be beneficial to obtain data from further offshore sites to see if this trend is seen at other sites. This result is important given the increasing number of offshore wind farms currently planned in the UK.

Figure 4-13: NMAE for offshore wind farms

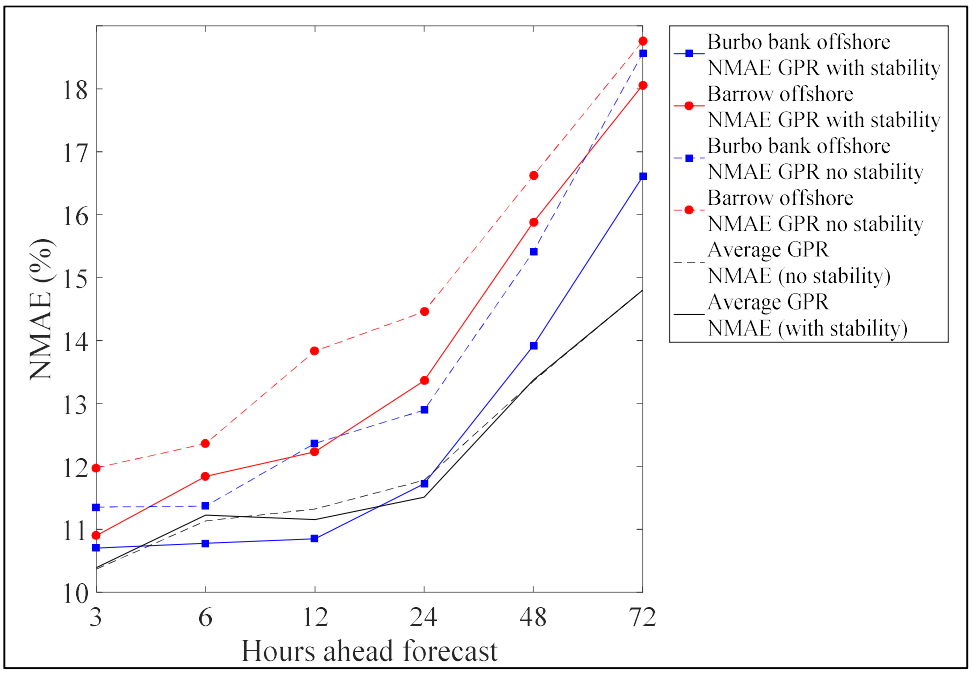
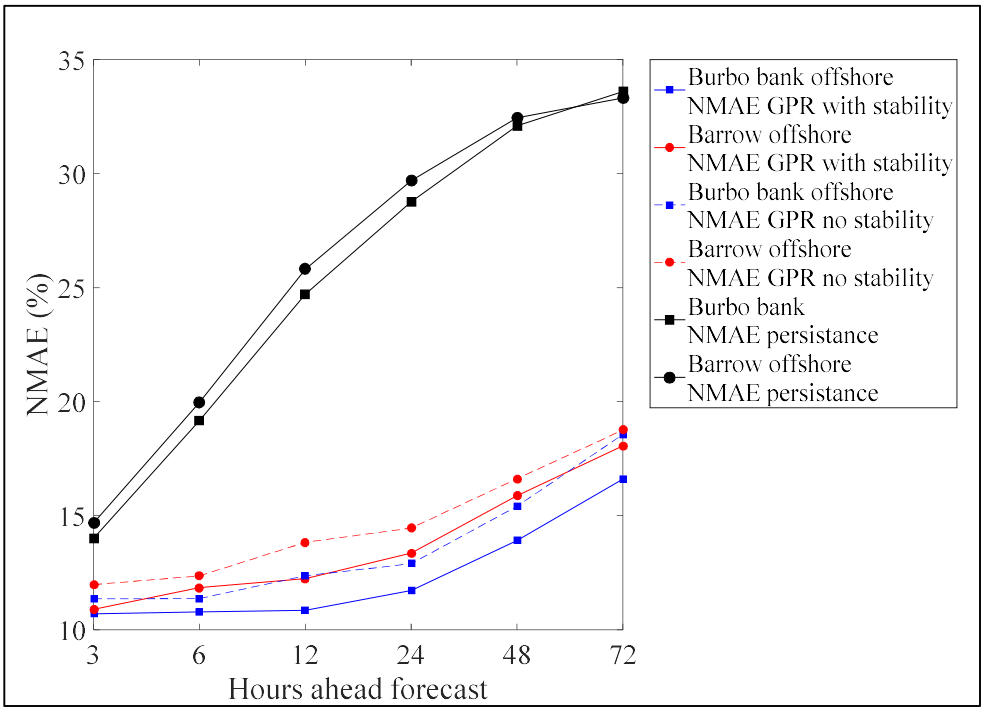


Figure 4-14: NMAE for offshore wind farms



In Figure 4-15 NMAE is considered for 2 sites where only a short data set was available. Both sites were onshore wind farms with a 3 month data set available. 2 months of data was used in the training set and 1 month for the test set. NMAE for the two sites with a short dataset are shown alongside average NMAE over 18 standard sites. It can be seen in Figure 4-15 that there is little difference between the GPR models with and without stability at these two sites. Furthermore, NMAE is higher than the average for the 18 standard sites. However it can be seen in Figure 4-16 that the hybrid NWP and GPR model can offer improved predictions over the persistence method. This indicates that whilst the method can be used with short data sets with good results, the model may not be able to predict quite as well with a shorter dataset. This agrees with some of the work performed by Chen et al. [73] who tested a hybrid NWP and censored GPR model for wind farm prediction at a site with 2.5 months of data. A short training data set could be useful in practice as less time is needed to gather the dataset allowing the method to be used to provide predictions at newly established wind farms. This enables new wind farms to be competitive in the electricity markets from an early stage, thus encouraging investment. Furthermore it can help in electricity system planning. In order to establish how long a data set is needed to provide reliable predictions further work is required.

Figure 4-15: NMAE for wind farms with a short data set

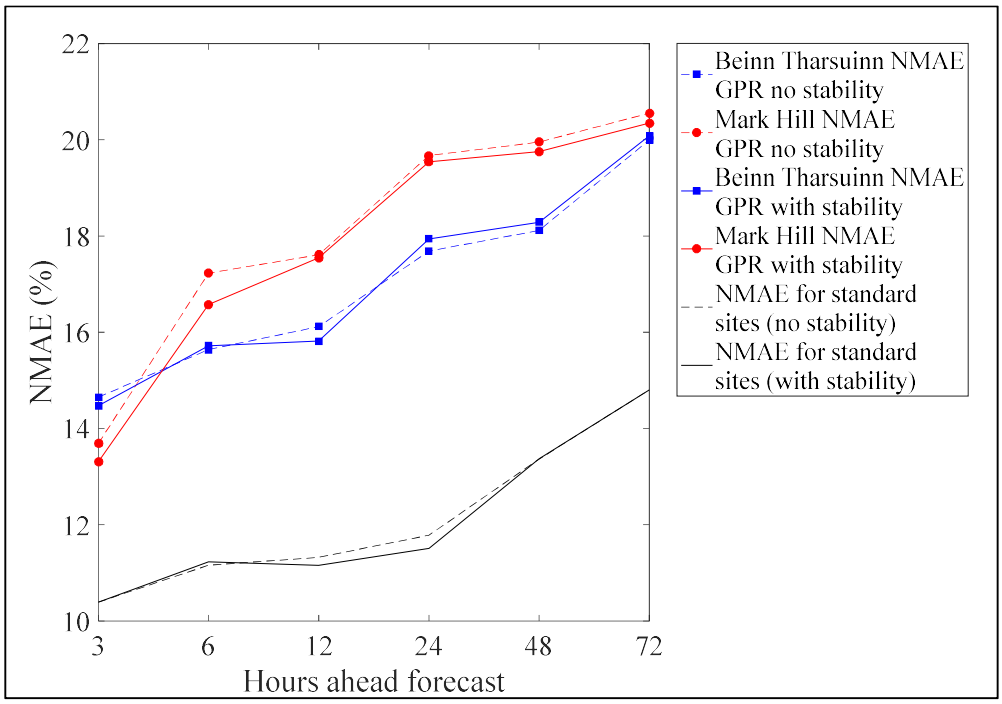
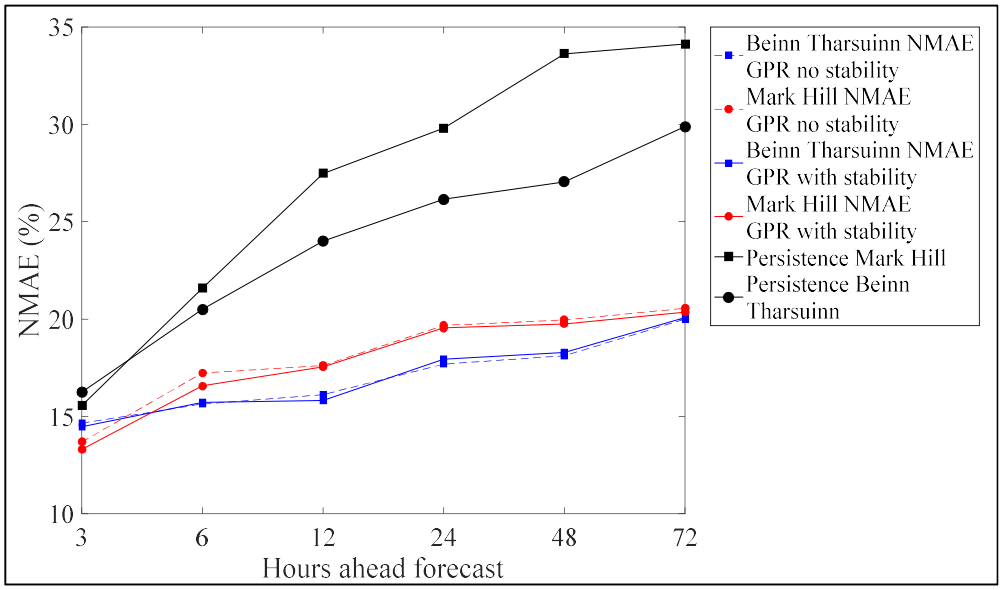


Figure 4-16: NMAE for wind farms with a short data set in comparison to persistence method



4.4 Conclusions

This chapter has further developed the hybrid NWP and GPR model first introduced in Chapter 3 by using the Obukhov length stability parameter as an input in the model. This allows predicted stability conditions to be incorporated into models for the prediction of wind speed and power output. The model was tested for the prediction of near-surface wind speeds, hub height wind speeds, individual turbine power output and wind farm power output. Overall it was seen that the hybrid NWP and GPR model produces competitive forecasts for all 4 applications. However, including stability in predictions did not prove useful in all applications.

For the prediction of near-surface wind speeds atmospheric stability was found to have little impact on predictive performance. However, the hybrid NWP and GPR model both with and without the Obukhov length stability parameter lead to a reduction in MAPE of approximately 2% over the Met Office NWP predictions across all time periods investigated.

For hub height wind speed predictions some improvement in wind speed prediction performance was observed when the hybrid NWP and GPR model was implemented with the Obukhov length stability parameter as an input in comparison to the model without. For predictions from 3 hours to 24 hours ahead there was between 2 and 5% reduction in MAPE and for predictions 48 or 72 hours in advance there was an improvement of 1%. In addition, this improvement lead to slight improvements in wind power predictions for an individual turbine. In order to improve this work it would be beneficial to test the model for an individual turbine for other locations. This is because, as seen for the other model applications, model performance can be highly site specific so it would be valuable to see if the results would be replicated at other sites.

The hybrid NWP and GPR model was also used to predict wind power output. This was done for 18 onshore sites, 2 offshore sites and 2 sites with a short data set. Whilst the GPR model produced wind farm power output forecasts which were competitive in performance to other methods in literature it was found that in

general including atmospheric stability did not lead to reduced prediction errors. However, at the offshore sites it was seen that including stability in predictions did lead to improved model performance. This indicates that the model could be a good method for predicting power output for offshore sites if limited wind speed predictions are available. To verify this result further it would be useful to obtain data for other offshore wind farms. At the 2 sites where only a short data set was available the errors were higher than for the 18 other sites. However, there was still a significant reduction in NMAE in comparison to the persistence method. At all sites investigated, it would have been valuable to have observed wind speed data at the wind farms investigated to see if the inclusion of wind speed data would lead to improved wind power output predictions. This data was unavailable at the time of research, however, could provide valuable additional insights.

In order to further explore the difference between the models in each application, an ANOVA (analysis of variance) test could be used to analyse the difference between the mean error in each model. ANOVA test can be used to determine whether the difference in errors seen between two models is statistically significant. This tests the null hypothesis that both group means are the same against and alternative hypothesis that there is a difference between the means. In this case, it could be used to check that the difference in MAE between the hybrid NWP and GPR model with Obukhov length and the model without Obukhov length is statistically significant.

It can be seen here that the GPR model captures the complex relationships between variables necessary for the prediction of wind speed and power output. In the deterministic framework the GPR model performs well in comparison to other methods available. However, the GPR model is able to give more information which has not been exploited here. The GPR model is a probabilistic forecasting technique which provides a density forecast. It gives the predictive probability distribution for each forecast point. In Chapter 5 the GPR model for wind farm power prediction is evaluated in a probabilistic framework. In addition to this, the choice of likelihood function in the GPR model is discussed.

Chapter 5. Probabilistic forecasting of wind farm power output

5.1 Overview

In Chapters 3 and 4 the hybrid NWP and GPR model showed promising results for both wind speed and power prediction. This demonstrates the potential of the model, particularly for wind farm power prediction which has value in a commercial setting. The results presented in Chapters 3 and 4 focused on the interpretation of predictions from the GPR model as a deterministic forecast. This is because this allows the results to be easily compared to other models, to show the benefit of using a GPR model. However, one of the key benefits of the GPR model is that it gives the user more information than simply point prediction. It gives a predictive distribution, which allows the user to obtain an expected value at the point of prediction and an estimate of the prediction uncertainty. This estimate of uncertainty has a number of practical benefits which were discussed in Section 2.3.4. For example, suppose a system operator wishes to use a wind power forecast to define an appropriate level of reserve capacity to ensure that demand for electricity is met. To evaluate this in a simplistic sense, consider a single generating unit and the expected power output at this unit. Suppose at time t the forecasted power output is 60 MW. There is a chance that the actual power output may exceed 60 MW and a chance that the actual power output may be lower than 60 MW. If the actual power output is lower than 60 MW, power must be procured from an alternative generator to meet demand. Where this is not possible, demand will not be met and will need to be curtailed. This scenario is generally avoided where at all possible. Alternatively, if the actual power output is greater than 60 MW, the wind farm may be curtailed or a different generating unit within the system will need to be curtailed to ensure supply of electricity matches demand. In order to plan in advance for either eventuality, the system operator must set aside a level of capacity (reserve capacity)

to increase and decrease production or demand at short notice. This level of reserve capacity within the system must be set such that the probability of failing to meet demand is low whilst keeping the cost of reserve capacity reasonably low. If a deterministic forecast is used, the level of uncertainty in the forecast is not predicted. In order to estimate the level of reserve required in this case, system operators usually consider the historical forecast errors of the particular method used [166]. Whilst this can give some indication of the reserve required it is a crude estimate and can lead to higher system costs [131, 167]. Alternatively, if a probabilistic forecast is used the level of reserve required can be set using predictive uncertainty specific to the forecasted time period in question. This enables a more adaptable representation of uncertainty and can lower costs whilst maintaining an appropriate probability of not meeting required demand. This use of probabilistic forecasts is becoming increasingly common with system operators moving away from a stochastic method for defining reserve to considering uncertainty in forecasts of both generation and demand. For example ERCOT (Texas independent system operator) and REE (Spanish system operator) have included some variety of forecast error into their definition of reserve requirements [167]. Improving forecast accuracy allows this system balancing to be performed with greater certainty and is hence valuable to a system operator. This is particularly important as non-dispatchable renewable energy generators form an ever increasing proportion of the electricity mix. This is just one example of why improving probabilistic forecasts is important.

In this chapter, the hybrid NWP and GPR model for wind farm power prediction developed in Chapter 4 is evaluated in a probabilistic framework and compared to other probabilistic wind power forecasting techniques. This chapter focuses on wind farm power prediction, as probabilistic forecasts have most relevance for larger scale wind power forecasting. In addition to evaluation of the GPR model in a probabilistic framework this chapter discusses appropriate likelihood functions for the GPR model. Thus far a Gaussian likelihood function has been used. However, in this chapter an alternative likelihood function is introduced. This chapter explores how, from a probabilistic perspective, an alternative likelihood function may be appropriate for wind power predictions. A number of authors have employed a GPR

model for wind speed or power forecasting. For example, Chen et al. [71] and Hu et al. [160] use hybrid GPR and wavelet transform models to predict wind speed and Chen et al. [73] use a hybrid NWP and GPR model to predict wind power output. However, to the best of the author's knowledge there no examples of hybrid NWP and GPR used in a wind power forecasting model which is evaluated from a probabilistic perspective. Additionally, no current literature explores the impact of using different likelihood functions in a GPR model for wind power prediction.

The main objectives of this chapter are:

(i) To evaluate the GPR model as a probabilistic forecast for wind power predictions at 22 UK wind farms.
(ii) To compare the hybrid NWP and GPR model for wind power forecasting to other available probabilistic forecasting techniques.
(iii) To consider the validity of a Gaussian likelihood function in the GPR model and introduce an alternative likelihood function.

In this chapter, predictions of wind farm power output made using the hybrid NWP and GPR model are evaluated in a probabilistic framework for the 22 wind farms detailed in Chapter 4. This includes 18 onshore sites with a 1 year dataset, 2 offshore wind farms with a 1 year dataset and 2 onshore sites with a short dataset. Section 5.2 explores how a more appropriate likelihood function might be chosen whilst the results are discussed in Section 5.3.

5.2 Methodology

GPR is a non-parametric Bayesian regression technique, as discussed in Section 3.2. Regression is used to learn a mapping from an input space to an unbounded observation space of the dependent variable [168]. In the prediction model presented here, the dependent variable is wind power. However in this case, particularly for wind power predictions, the value predicted at time t, $\hat{y}_t$, should lie within the range $0 \leq \hat{y}_t \leq C$ where C is the capacity of the turbine or wind farm for which

predictions are made. The Gaussian distribution for a random variable $X \sim N(\mu, \sigma^2)$ is defined on a continuous range $-\infty \leq X \leq \infty$. This implies that on any interval of real numbers X has a probability greater than zero. In the case of the predictive distribution for wind power this means $P(\hat{y}_t > C) > 0$ and $P(\hat{y}_t < 0) > 0$. Whilst using a well-trained model probability of $\hat{y}_t > C$, $\hat{y}_t < 0$ will be very small, it is still infeasible from a probabilistic perspective [168].

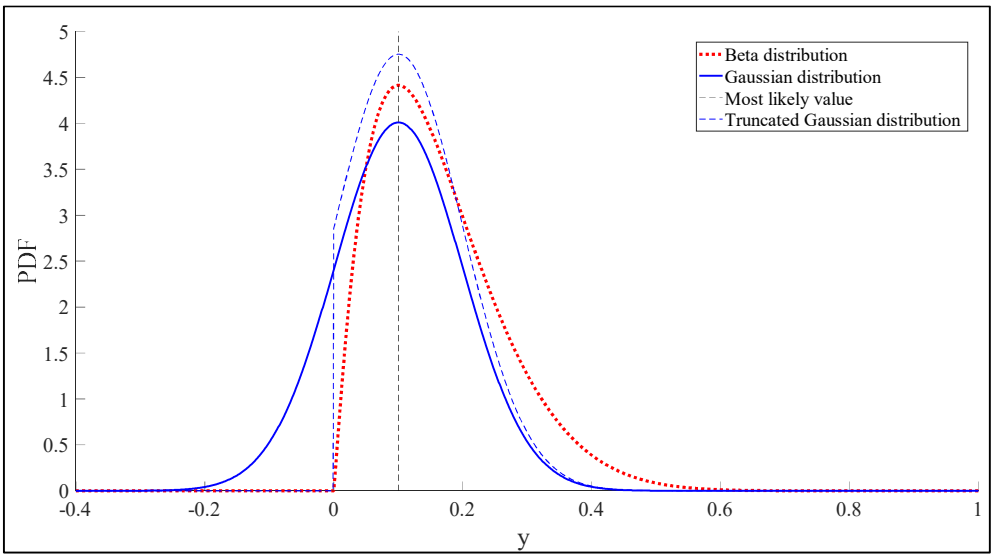
Equation 5-1 first given in Chapter 3 and repeated here gives the formula for posterior distribution, obtained using Bayes formula.

$$p(\mathbf{f}_*|\mathbf{y}, \mathbf{X}, \mathbf{X}_*) = \frac{p(\mathbf{y}|\mathbf{f}, \mathbf{X})p(\mathbf{f}|\mathbf{X})}{p(\mathbf{y}|\mathbf{X})} \qquad \text{Equation 5-1}$$

$$posterior = \frac{likelihood \times prior}{marginal\ likelihood}$$

where $\mathbf{y}, \mathbf{X}$ are the dependent and predictor variables in the training set, $\mathbf{X}_*$ are predictor variables in the test set and $\mathbf{f}_*$ is the underlying function at $\mathbf{X}_*$. If the likelihood is Gaussian then, the posterior distribution and the predictive distribution will be Gaussian. From a deterministic standpoint, this can still be used to obtain the most likely point value for a prediction. This is demonstrated in Figure 5-1. Here three distributions are shown: a Gaussian distribution, a beta distribution and a truncated Gaussian distribution. All three distributions give the same value if the most likely value is required, shown as the maximum value of the pdf. However, the Gaussian distribution allows $P(\hat{y}_t < 0) > 0$. In the case of wind power forecasts, the predictive distributions are bounded as wind power will always be positive, but less than the wind farm capacity. This indicates that whilst the Gaussian likelihood function can be used to find the most probable value for a prediction it is not appropriate from a probabilistic perspective.

Figure 5-1: Example predictive distributions a random variable, *y*.



Other distributions which could be used for predicting bounded data are a warped Gaussian distribution, a beta distribution and a truncated Gaussian distribution. A warped Gaussian distribution applies a transformation from the bounded variable to an unbounded variable so it can be modelled by a Gaussian process [169]. The beta distribution is a continuous probability distribution defined on the bounded domain [0,1] [170]. The truncated Gaussian distribution is a Gaussian distribution which is altered to have a bounded domain [157]. These are discussed by Jensen et al. [168] who conclude that the beta distribution and the truncated normal distribution provide better model fits for bounded data than a warped Gaussian distribution. At the time of writing, no stable function was available for the implementation of the truncated likelihood. The beta likelihood function for implementation using the GMPL toolbox [146] is described by Rasmussen and Nickisch [170]. They state that interval data $y \in [0,1]$ can be modelled in the Gaussian process framework using the beta likelihood function. For a random variable, *y*, the beta likelihood function is given by Equation 5-2.

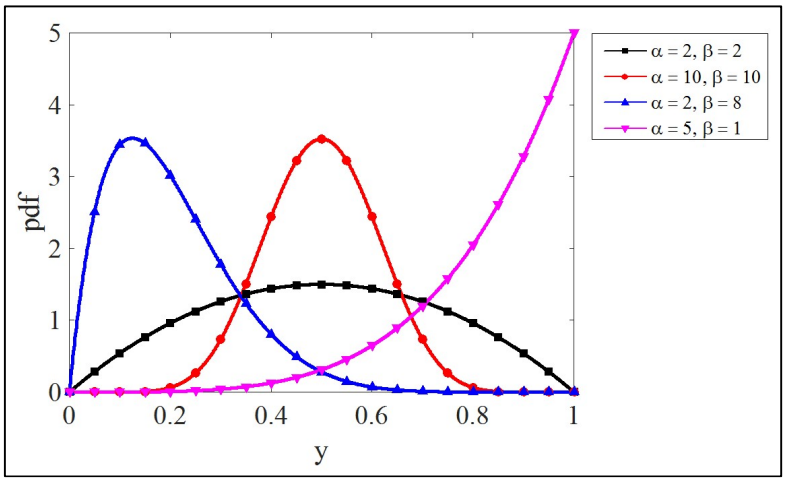
$$p(y) = \frac{y^{\alpha-1}(1-y)^{\beta-1}}{B(\alpha,\beta)} \text{, for shape parameters } \alpha,\beta > 0 \qquad \text{Equation 5-2}$$

where the beta function, $B(\alpha, \beta)$ is given by Equation 5-3.

$$\frac{1}{B(\alpha,\beta)} = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \qquad \text{Equation 5-3}$$

In Equation 5-3, Γ refers to the standard mathematical gamma function. The beta distribution has mean $\mu = \frac{\alpha}{\alpha+\beta}$ and variance $\sigma^2 = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$ and is multimodal for $\alpha, \beta < 1$. It is a flexible distribution which allows different characteristics of probability functions to be modelled. Some examples of the beta distribution with different shape parameters are given in Figure 5-2.

Figure 5-2: Example beta distribution for different shape parameters.



As introduced in Section 3.2.4, the GPR model predicts values $\mathbf{y}_*$ for a test set of inputs $\mathbf{X}_*$. The posterior distribution first introduced in Chapter 3 and shown in Equation 5-1 is required to obtain predictive distributions for test data points $\mathbf{D} = (\mathbf{X}_*, \mathbf{y}_*)$.

For the Gaussian likelihood function, the posterior distribution is analytically tractable. However, for the beta distribution approximate inference techniques must be used. In the work performed here Laplace approximation is used, implemented using the GPML toolbox [146]. Using a Laplace approximation the posterior distribution is approximated using a single Gaussian, $q(\mathbf{f})$ as shown by Equation 5-4 [168].

$$p(\mathbf{f}|\mathbf{y},\mathbf{X},\boldsymbol{\theta}) \approx q(\mathbf{f}) = N\left(\mathbf{f}\middle|\hat{\mathbf{f}},\mathbf{A}^{-1}\right) \quad \text{Equation 5-4}$$

where $\mathbf{f},\mathbf{y},\mathbf{X}$ are as used previously, $\boldsymbol{\theta}$ is the set of hyperparameters used in the GPR model, $\hat{\mathbf{f}}$ is the mode of the posterior distribution and $\mathbf{A}$ is the Hessian of the negative log posterior at the mode. The predictive distribution of a set of test points $\mathbf{y}_*$ is calculated using the integral shown in Equation 5-5.

$$p(\mathbf{y}_*|\mathbf{y},\mathbf{X},\mathbf{X}_*) = \int p(\mathbf{y}_*|\mathbf{f}_*)p(\mathbf{f}_*|\mathbf{y},\mathbf{X},\mathbf{X}_*)d\mathbf{f}_* \quad \text{Equation 5-5}$$

For the Gaussian likelihood function, Equation 5-5 has a closed form solution. However, when using the beta likelihood functions, the integral must be solved using numerical methods [168]. This is implemented using the GMPL toolbox [146].

In this chapter, the hybrid NWP and GPR model developed in Chapter 4 is implemented for the prediction of power output at 22 wind farms across the UK. The model is implemented twice, firstly with Obukhov length as an input parameter and then without. This allows the impact of Obukhov length on model performance to be evaluated. The predictor variables used for the model with and without Obukhov length are shown in Equation 5-6 and Equation 5-7 respectively. The dependent variable at time *t* for both models is power output at an individual wind farm. The wind farm power output was obtained from the Balancing Mechanism Reporting System (BMRS) [161]. This data was obtained for 22 sites across the UK as shown in Figure 4-1 in Chapter 4. At these sites observed wind speed data was unavailable, therefore forecasts were based on predicted wind speeds and prior power output. The Met Office forecast wind speed was obtained from the Met Office NWP model as described in Chapter 3. The data used for analysis was for 1st Jan 2015 – 31st Dec 2015. For the model development the data was split into training and test datasets to ensure the forecast is independent. The training set contains data from the first 9 months of the dataset and the test set contains data from the final 3 months.

$$\mathbf{X}_t = \left[\mathbf{m}_{t,k}, \mathbf{y}_{t-k-1}, \mathbf{y}_{t-k-2}, \mathbf{y}_{t-k-3}, \frac{1}{\mathbf{L}_{t,k}}\right] \text{ for } t = 4, \dots, n \quad \text{Equation 5-6}$$

$$\mathbf{X}_t = \left[\mathbf{m}_{t,k}, \mathbf{y}_{t-k-1}, \mathbf{y}_{t-k-2}, \mathbf{y}_{t-k-3}\right] \text{ for } t = 4, \dots, n \quad \text{Equation 5-7}$$

where:

$t =$ time of prediction

$k =$ forecast horizon

$\mathbf{m}_{t,k} =$ Met office wind speed forecast for time t at k hours ahead

$\mathbf{y_t} =$ observed wind power at time t (ms^{-1})

The key aim of implementing this model is to evaluate the performance of the hybrid NWP and GPR model for wind power forecasting in a probabilistic framework and to investigate the use of an alternative likelihood function in the GPR model. The model is implemented using the predictor variables in Equation 5-6 and Equation 5-7 with a Gaussian likelihood function and a beta likelihood function. The results are shown for an individual wind farm, on average over 18 standard wind farm sites, on average at 2 offshore sites and for 2 locations with short data sets. These sites were shown in Section 4.3.4. The results are discussed from a probabilistic perspective. Hence the measures used to compare results are those discussed in Section 2.2.6, reliability, sharpness, resolution and continuous ranked probability score (CRPS). An overview of these is given here to add clarity.

The first measure of a probabilistic forecast to be discussed is reliability. This was discussed in Section 2.2.6. Reliability is a measure of the statistical consistency between the predictive distribution and the observed values [76]. A prediction of a test point y^*_{t+k} made at time *t* with forecast horizon *k* is denoted by $\hat{y}^*_{t+k}$. The predictive distribution $\hat{y}^*_{t+k}$ is given by a cumulative distribution function (CDF), F, where $F_{\hat{y}^*_{t+k}}(x) = P(\hat{y}^*_{t+k} \leq x) = \alpha,\ x \in \mathbb{R}$ for a Gaussian distribution and $x \in [0,1]$ for a beta distribution. To allow forecasts to be compared at different sites, predictive distributions are given as a proportion total capacity. For a nominal probability, α, reliability compares the inverse CDF to the observed frequency of outcomes. The inverse CDF, ${F_{\hat{y}^*_{t+k}}}^{-1}$, is a function of α given by Equation 5-8.

$$F_{X_t}{}^{-1}(p) = inf\{x \in \mathbb{R} : F_{X_t}(x) \leq p\}$$ Equation 5-8

Subsequently, reliability, $\hat{a}_k^{\alpha}$, for a nominal probability α and forecast horizon k can be calculated using Equation 5-9 and Equation 5-10, first given in Section 2.2.7.

$$\varepsilon_{t+k}^{\alpha} = \begin{cases} 1 \text{ if } y_{t+k} < F_{\hat{y}_{t+k}^*}^{-1}(\alpha) \\ 0 \text{ otherwise} \end{cases}$$ Equation 5-9

$$\hat{a}_k^{\alpha} = \frac{1}{N}\sum_{t=1}^{N} \varepsilon_{t+k}^{\alpha}$$ Equation 5-10

Reliability diagrams are shown in this chapter. In these, the deviation between observed and nominal probability, $b_k^{\alpha} = \alpha - \hat{a}_k^{\alpha}$ is plotted against the nominal probability, α. Perfect reliability is achieved if $\hat{a}_k^{\alpha} = \alpha, \forall\alpha$, or alternatively if $b_k^{\alpha} = 0 \ \forall\alpha$.

Sharpness is measured by considering mean interval length given different probability intervals. Predictions with a lower level of uncertainty are more useful, provided reduced uncertainty does not lead to reduced reliability, as discussed in Section 2.2.6. For a chosen coverage rate, β sharpness is calculated using Equation 2-26 introduced in Section 2.2.7 repeated here for clarity (Equation 5-11). A lower value of $\overline{\delta_k^{\beta}}$ for any value of β indicates a more useful forecast.

$$\overline{\delta_k^{\beta}} = \frac{1}{N}\sum_{t=1}^{N}\left(F_{\hat{y}_{t+k}^*}^{-1}\left(1 - {}^{\beta}/_{2}\right) - F_{\hat{y}_{t+k}^*}^{-1}\left({}^{\beta}/_{2}\right)\right)$$ Equation 5-11

In addition to the mean interval length considered above, resolution is considered by calculating the standard deviation of interval size for different nominal probabilities. The standard deviation of interval size gives an indication of the model ability to differentiate between different circumstances. For a chosen coverage rate, β resolution is calculated using Equation 5-12. A higher value of σ_k^{β} for any values of β indicates a more useful forecast.

$$\sigma_k^{\beta} = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(F_{t+k}^{-1}\left(1-{}^{\beta}/_{2}\right)-F_{t+k}^{-1}\left({}^{\beta}/_{2}\right)-\overline{\delta_k^{\beta}}\right)^2} \quad \text{Equation 5-12}$$

The final measure of the predictive ability of a model is given by the continuous ranked probability score (CRPS). This gives an overall score which can be used to compare between models. This was introduced in Section 2.5.2 and is defined by Equation 5-13.

$$CRPS(F_t, y_t) = \int_{-\infty}^{\infty}(F_t(x) - H(x-y_t))^2 dx \quad \text{Equation 5-13}$$

where H is the Heaviside step function given by $H(z) = \begin{cases} 0, & z < 0 \\ 1, & z \geq 0 \end{cases}$

In order to implement CRPS in practice a closed form expression for Equation 5-12 must be established. Jordan [171] derives a closed-form expression for CRPS given different distributions and Lerch et al. [78] give software for calculating CRPS. For a normal distribution with mean μ and standard deviation σ CRPS is given by Equation 5-14 where φ denotes the pdf and Φ denotes the CDF of the standard normal distribution. For a beta distribution with shape parameters α, β CRPS is given by Equation 5-15.

$$CRPS\left(N_{\mu,\sigma}, y\right) = \sigma\left[\frac{y-\mu}{\sigma}\left(2\Phi\left(\frac{y-\mu}{\sigma}\right)-1\right)-2\varphi\left(\frac{y-\mu}{\sigma}\right)-\frac{1}{\sqrt{\pi}}\right] \quad \text{Equation 5-14}$$

$$CRPS\left(F_{\alpha,\beta}, y\right) = y\left(2F_{\alpha,\beta}(y) - \frac{2B(2\alpha,2\beta)}{\alpha B(\alpha,\beta)^2}\right) \quad \text{Equation 5-15}$$

where B is the beta function given in Equation 5-3. The CRPS is equivalent to MAE for a deterministic forecast. As for MAE, a lower value for CRPS indicates an improved performance.

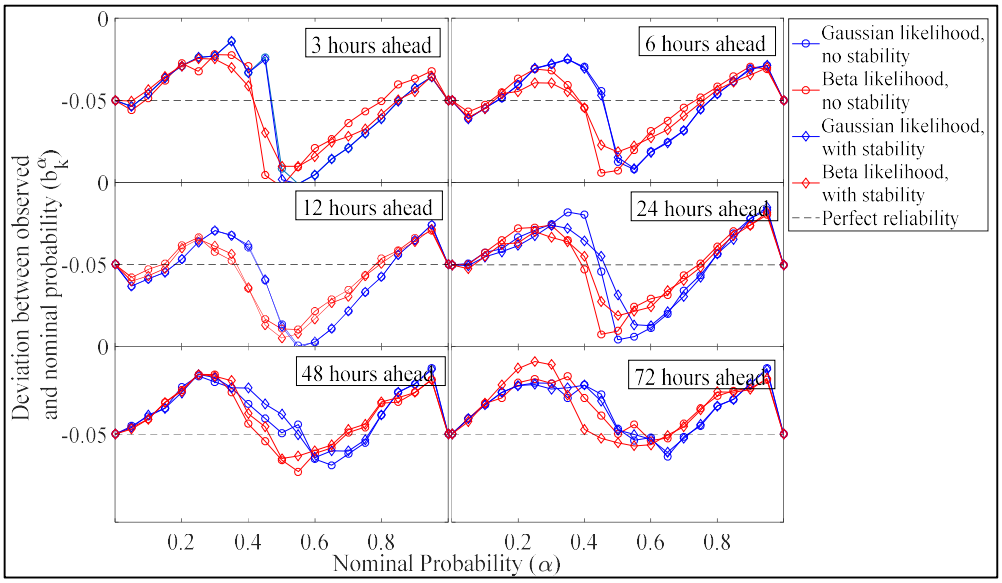
5.3 Results

5.3.1 Investigation of probabilistic power forecasting at one site

To begin with, the results are shown in detail for one site. This allows the key features of the GPR model with a beta likelihood function to be compared to the GPR model with a Gaussian likelihood function. In this section, the results for Toddleburn wind farm are discussed. Toddleburn is a 27.6 MW onshore wind farm located in Scotland. It consists of 12 2.3 MW turbines.

In Figure 5-3 reliability is shown for Toddleburn wind farm. The results are shown for $\alpha = 0, 0.05, 0.1, \ldots, 0.95, 1$ for forecast horizons from 3 hours ahead to 72 hours ahead. For comparison a horizontal line is shows reliability for a perfect forecast.

Figure 5-3: Reliability of GPR models at Toddleburn wind farm

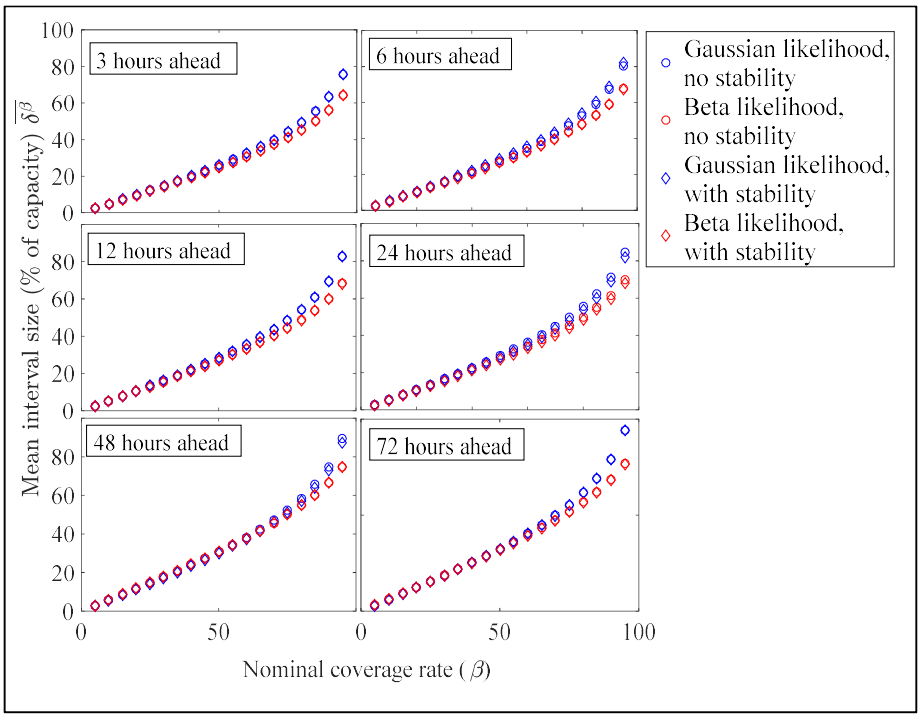


It can be seen here that the deviation between observed and nominal probabilities is usually closer to zero for the beta distribution. This indicates that the model which used the beta likelihood function gives more reliable predictions. Whilst the beta distribution provides more reliable predictions, there is limited difference between

the GPR model performance with stability and without stability for both likelihood functions.

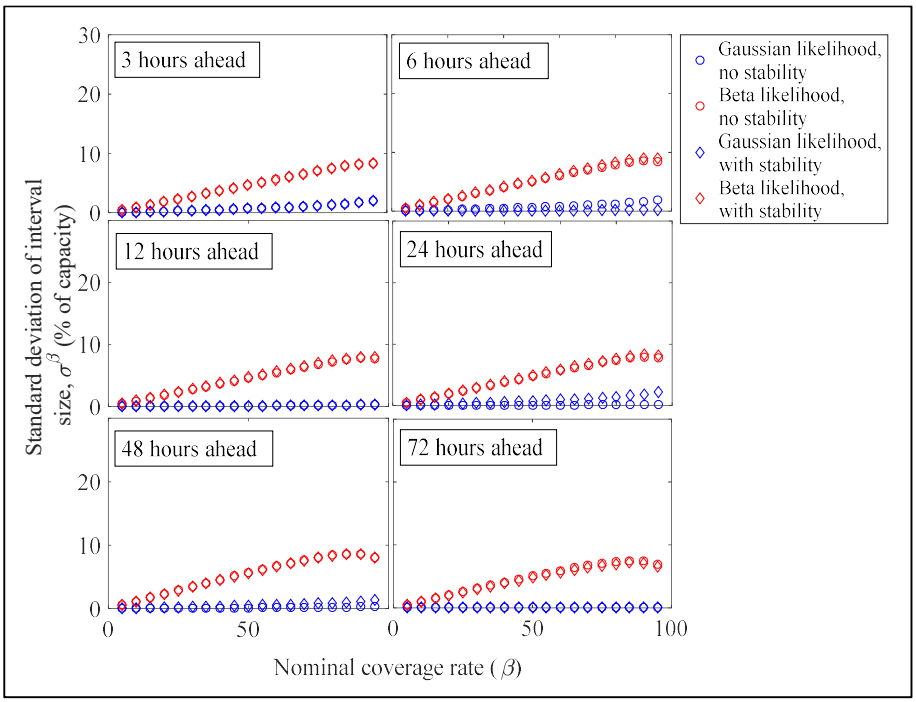
Alongside reliability, sharpness of predictive distributions is used as a measure of the level of uncertainty. In Figure 5-4 sharpness is shown for GPR with Gaussian likelihood and beta likelihood. For both likelihood functions, it can be seen that the GPR model with and without stability shows no difference in mean interval length. However, mean interval length is smaller for the GPR model with a beta likelihood for all forecast horizons. Mean interval size gradually increases as forecast horizon increases. As the nominal probability increases, the difference between the GPR model with a beta likelihood function and with a Gaussian likelihood function becomes larger. For a 90% nominal probability for a beta distribution mean interval size ranges from 56% of capacity at 3 hours ahead to 68% at 72 hours ahead. For a Gaussian likelihood function its ranges from 63% at 3 hours ahead to 79% at 72 hours ahead. The smaller mean interval size seen when using the beta likelihood function indicates a more certain prediction of wind farm power output. In order to understand why this might be useful, consider again the example given in Section 5.1 of a probabilistic forecast used to define the reserve requirement considering a forecast at time *t*. Comparing two scenarios, one in which there is a lower level of prediction uncertainty and an alternative with a higher level of uncertainty. In the first case, the range of output values expected with 99% confidence might be 58 – 62 MW. In the second scenario, for the same confidence level the range of output values might be 50 – 70 MW. In the first case it might be reasonable to suggest a reserve level of 2 MW whilst in the second scenario a higher level of reserve would be required. The higher level of reserve would come at an additional cost; hence a smaller mean interval size would be valuable in a forecast. Whilst this is a simplistic example, it gives some suggestion of the properties which make a good probabilistic forecast. In general, this type of reserve capacity definition would be done on a system wide scale considering power forecast uncertainty and load forecast uncertainty, however the reduction of uncertainty in a probabilistic forecast is still useful.

Figure 5-4: Sharpness of GPR models at Toddleburn wind farm



In addition to the mean interval length considered above, resolution is considered. The standard deviation of interval length for different forecast horizons is shown in Figure 5-5. The higher standard deviation of interval width seen for the GPR model with beta likelihood shows that this model is better able to distinguish between predictive points which have higher or lower uncertainty. For the GPR model with Gaussian likelihood the standard deviation of interval width reduces as forecast horizon increases, indicating that by the time the forecast horizon reaches 72 hours ahead the forecast uncertainty is mostly uniform over all the predictions made. This again indicates that the model with beta likelihood is superior to the model with Gaussian likelihood. However, as shown when considering sharpness in Figure 5-4, there is no difference in terms of uncertainty in predictions for the GPR model with stability in comparison with the GPR model with no stability.

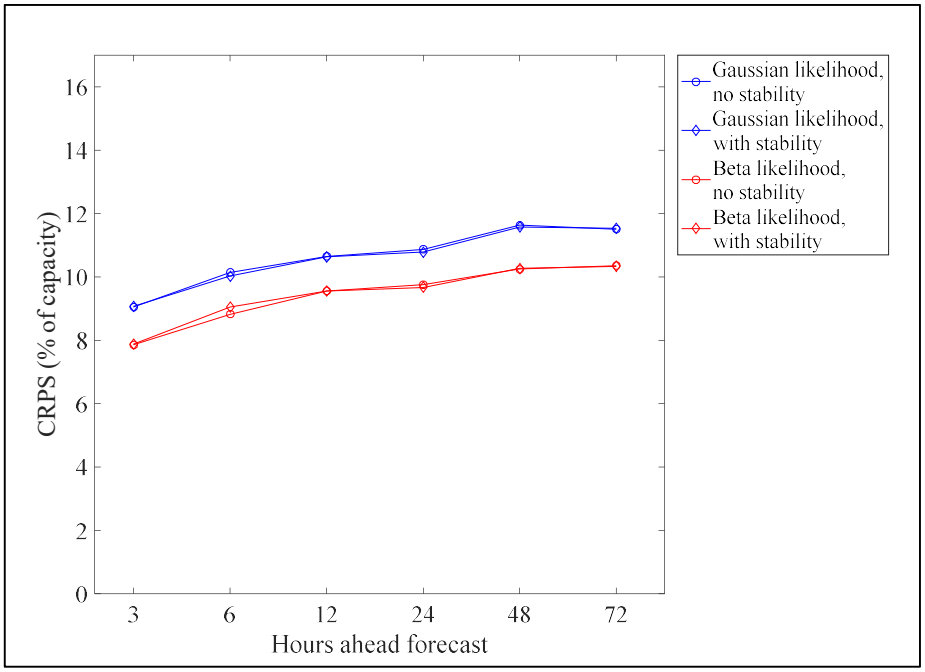
Figure 5-5: Resolution of GPR models at Toddleburn wind farm



CRPS for the GPR model with a Gaussian likelihood and a beta likelihood are shown in Figure 5-6. It can be seen here that there is largely no difference between the GPR model with stability and the GPR model without stability. However, it can be seen that CRPS is lower for the GPR model with a beta likelihood function. The lack of difference between the GPR model with and without stability mimics the results seen in the deterministic evaluation of the GPR model shown in Section 4.3.4. It is unclear why the inclusion of stability in wind farm power output does not improve model performance. This could be due to inaccuracy in the prediction of parameters used to calculate the Obukhov length stability parameter. From the observed variables available it is not possible to analyse the accuracy. Alternatively, it could be because the atmospheric stability conditions to do not explain a significant amount of the uncertainty seen when predicting wind power output from 10 m wind speed predictions. The 10 m NWP wind speed predictions used in the wind power prediction model were from sites up to 10 km away from the wind farm. Changes in wind speed between the sites, increases in wind speed with height and

converting wind speed to wind power amongst other things will all contribute to errors in the prediction of wind power output. It is likely that any uncertainty which might be explained by stability conditions is overshadowed by other errors in the prediction model. The lower CRPS for the GPR model with a beta likelihood function indicates that this is more appropriate for purpose of predicting wind power output.

Figure 5-6: CRPS for GPR models at Toddleburn wind farm



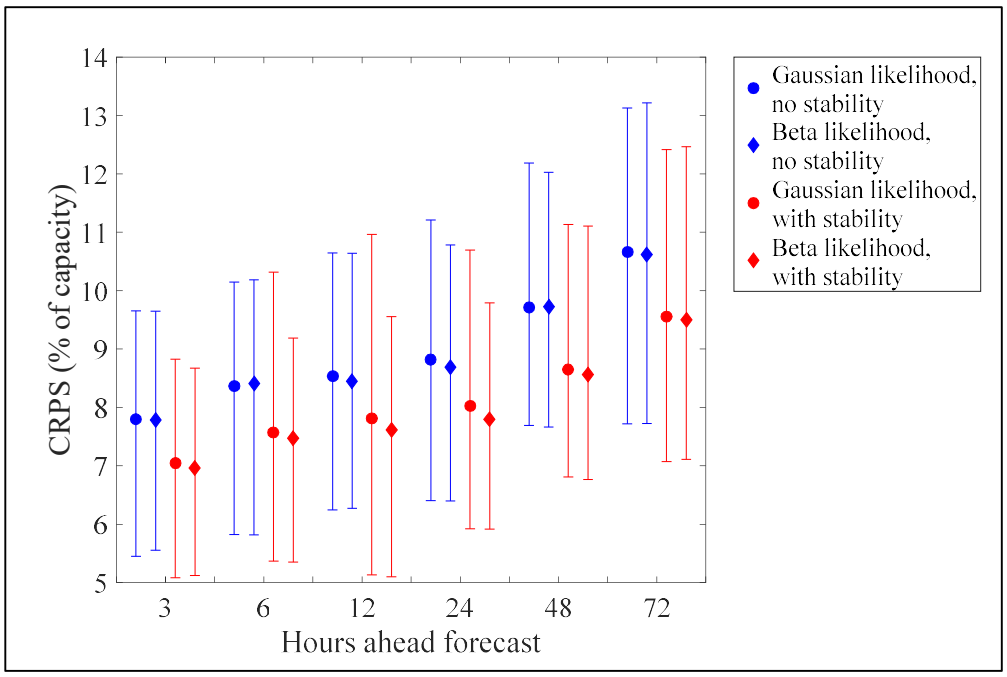
Overall, it can be seen from all the measures shown here that the hybrid NWP and GPR model with a beta likelihood function gives better probabilistic predictions than the same model with a Gaussian likelihood function. The reduced level of uncertainty in predictions and the increased ability to distinguish between different predictive conditions show that this model is useful for the probabilistic prediction of wind power output. This can have benefits in a number of practical applications of wind power forecasting. As mentioned in Section 5.1, using accurate probabilistic forecasts when predicting reserve levels can lead to lower system costs whilst maintaining system security [167]. Additionally, probabilistic wind power forecasts can be used alongside electricity price forecasts to manage financial risk whilst

bidding in electricity markets [172] or to decide when to charge or discharge an electricity storage system.

5.3.2 Probabilistic forecasting for 18 standard sites

Having evaluated the GPR models in a probabilistic framework at one site the next section looks at the overall results for 18 standard sites. This allows for a full evaluation of the model performance and allows comparisons with other probabilistic models to be made. In Figure 5-7 the average CRPS overall 18 standard wind farms is shown alongside the maximum and minimum CRPS at any of the 18 sites.

Figure 5-7: Average and range of CRPS for GPR models across 18 wind farms

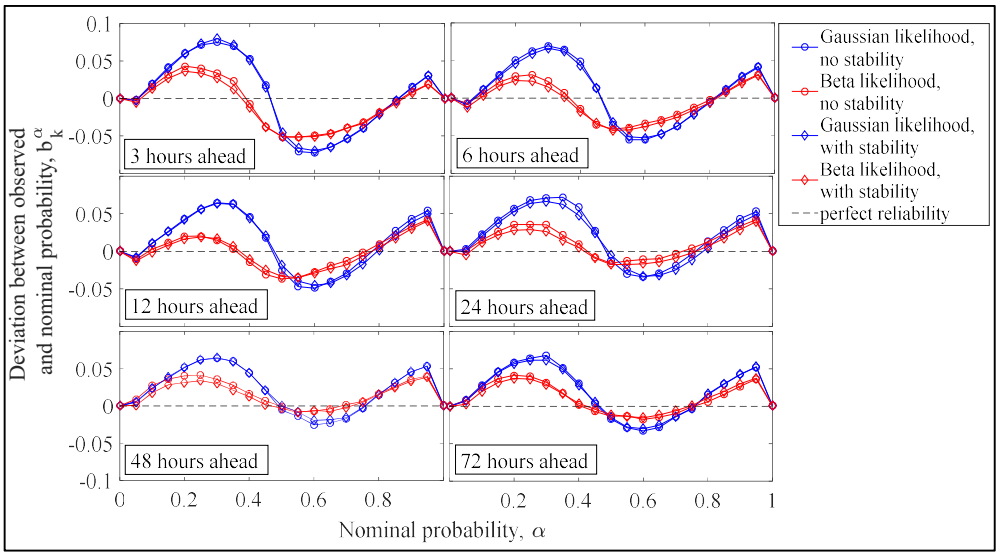


It can be seen here, as with MAE shown in Chapter 4 CRPS gradually increases as forecast horizon increases, indicating the erosion of forecast accuracy with prediction horizon. It can also be seen that CRPS is lower for the GPR model with a beta likelihood (both with and without stability) than for the GPR model with Gaussian likelihood. This is because using the beta likelihood function the predictive

distribution is limited to a feasible range which improves predictions. The difference between the models with different likelihood functions gradually increases with forecast horizon, from 0.85% at 3 hours ahead to 1.1% at 72 hours ahead. On average, the difference between the GPR model with and without stability is negligible, as shown in Chapter 4. The use of the beta likelihood function in the GPR model has not altered this result.

Next, the reliability over all 18 sites is discussed. Figure 5-8 shows the average deviation between observed and nominal probability for different nominal probabilities. This indicates that the deviation is closer to zero for the GPR model with beta likelihood in comparison to the GPR model with beta likelihood. This is true for all forecast horizons. This shows that the predictions made using a beta likelihood function are more statistically consistent with the underlying unobserved distributions. It can be seen that there are very small differences between results for the models with and without stability, however, these are not significant.

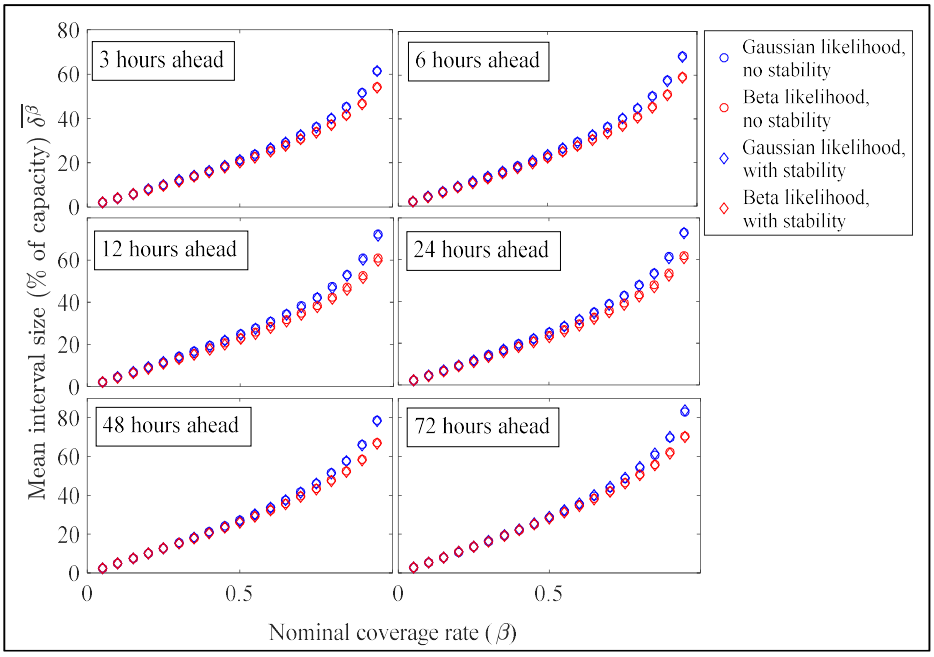
Figure 5-8: Average reliability over 18 standard wind farms



In Figure 5-9 sharpness is studied. Figure 5-9 shows the mean interval length over all 18 sites for different nominal probabilities. It can be seen that the mean interval length gradually increases over increasing forecast horizon, which is particularly notable for larger nominal probability. Furthermore, the mean interval size is lower

for the GPR model with beta likelihood function in comparison to the GPR model with Gaussian likelihood function. This indicates reduced uncertainty surrounding the predictions.

Figure 5-9: Average sharpness over 18 standard wind farms



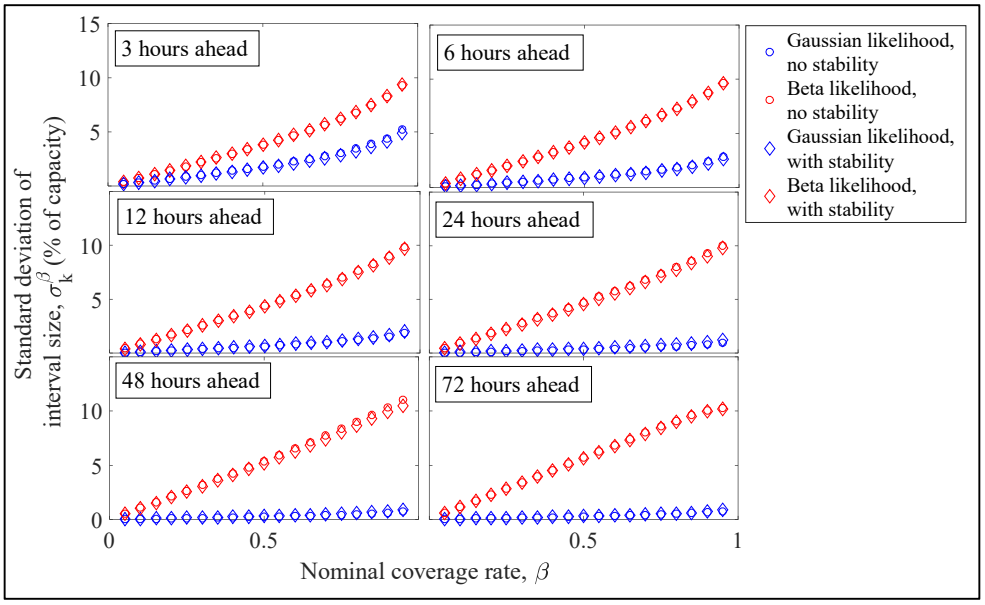
The difference between the two largely increases as the prediction horizon increase from 3 – 24 hours ahead, and is slightly lower for 48 and 72 hours ahead. For example, for a nominal probability of 90% the difference between models increases from 5% at 3 hours ahead to 9% at 72 hours ahead. Again, the difference between models with and without stability is small. This indicates that as the forecast horizon increases, the GPR model with a beta likelihood function is able to offer forecasts with less uncertainty, which are more useful.

In addition to sharpness, resolution is reviewed in Figure 5-10. The standard deviation of interval length is given as a measure of sharpness, and the average sharpness value for nominal probability β at forecast horizon k over all 18 sites is given by Equation 5-16.

$$\frac{1}{18}\sum_{s=1}^{18}\sigma_k^{\beta}$$ Equation 5-16

where σ_k^{β} is the standard deviation of interval length for nominal probability β at forecast horizon k is given by Equation 2-27 and s refers to the site indices. In this case a higher value for resolution indicates an increased ability to differentiate between levels of uncertainty under different predictive conditions. Figure 5-10 shows the mean value over 18 sites for resolution given difference nominal probabilities.

Figure 5-10: Mean resolution over 18 standard wind farms



This indicates higher resolution for the GPR with a beta likelihood function in comparison to the Gaussian likelihood function. Resolution for the GPR model with beta likelihood is largely constant irrelevant of the forecast horizon, whilst resolution decreases with forecast horizon for the GPR model with Gaussian likelihood. This indicates that the model with beta likelihood is better able to differentiate between levels of uncertainty under different circumstances. This also indicates that the ability to differentiate between circumstances is not eroded as forecast horizon increases for the model with beta likelihood. This indicates a better

forecast as the level of uncertainty in a prediction is dependent on the model inputs rather than being largely constant.

The results shown here indicate that the GPR model with a beta likelihood function performs better under a probabilistic framework in all terms of all 4 evaluation criteria: reliability, sharpness, resolution and CRPS. As described above, this is because the beta likelihood functions ensures the predictions made only have probability greater than zero when the predictions are in a feasible region.

In order to assess whether the results seen for here are reflective of other state of the art methods, some of the results seen in Table 2-7 are discussed. It can be difficult to draw comparisons between models, as results are frequently averages over different forecast horizons and given for different sites. However, some attempt is made at considering the performance of the GPR model in comparison with others. As discussed in Section 2.2.5 CRPS gives an overall scoring rule hence a discussion of CRPS is used to consider the achievements of other state of the art probabilistic wind power predictions. On average over 18 standard sites CRPS for the GPR model with a Gaussian likelihood ranges from 7.9% at three hours ahead to 11.8% at 72 hours ahead. For the GPR model with a beta likelihood function CRPS ranges from 7% at 3 hours ahead to 9.5% at 72 hours ahead. CRPS is only given for 4 of the methods discussed in Table 2-7. Juban et al. [87] give CRPS for forecasts from 0-60 hours ahead. CRPS ranges from 6.5% at 3 hours ahead to 11% at 48 hours ahead, higher than average CRPS for the GPR model with beta likelihood. Sideratos and Hatziargyriou [81] give CRPS for 2 sites. At the first site, CRPS is given as a near constant 10% over forecasts up to 60 hours in advance. At the second site, CRPS is given as around 6% at 3 hours ahead rising to 10% 42 hours ahead. CRPS given by the GPR model with a beta likelihood function is lower on average than at the first site, but slightly higher than at the second. Tastu et al. [90] give a density forecast generated based on average point forecast errors in the area. CRPS is shown for forecasts from 15 mins to 8 hours in advance. At 4 hours ahead CRPS is given as approximately 10%, staying as such for forecasts 6 hours ahead. This is again higher than CRPS seen for the GPR model. Finally, Zhang et al. [88] give CRPS for

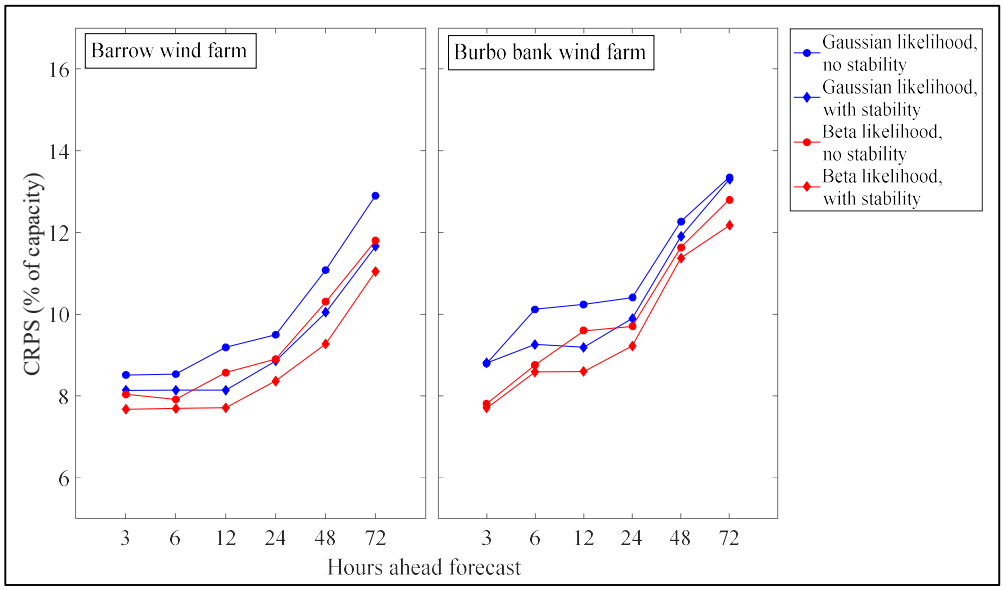
forecasts from 1 – 24 hours ahead. At 3 hours ahead CRPS of 2.5% at 3 hours ahead rising to 8% at 24 hours ahead. This is much lower than the results presented in this chapter at 3 hours ahead, but similar at 24 hours ahead. Overall the GPR model with a beta likelihood function performs well, with CRPS in the same range as other probabilistic models seen in Table 2-7. It is not possible from this to suggest that one model outperforms others overall as different data sets are used. However, this shows that the GPR model has the potential to make good probabilistic predictions for wind farm power output. From the discussion in Section 2.3.4 and the example given in Section 5.1 it can be seen that accurate probabilistic forecasts are valuable to a number of users. There is limited literature which considers the incremental benefit (economic or otherwise) of an improved probabilistic forecast. Because of this it is difficult to pinpoint the potential economic gain from improving forecast accuracy. However, many authors suggest that there is benefit in using probabilistic forecasts, therefore developing new methods and ensuring accurate probabilistic forecasts are available is important.

5.3.3 Probabilistic wind power forecasting at 2 offshore wind farms

Figure 5-11 explores CRPS for predicted wind power at 2 offshore wind farms. Reducing and anticipating uncertainty in wind power predictions for offshore wind farms is particularly important given the increase in offshore wind capacity. Decreasing commercial costs and higher offshore wind speeds has led to a significant increase in planned offshore wind capacity in the UK. These projects are usually large, with some planned wind farms having capacity up to 2.4 GW [17]. This means they can deliver a significant amount of power to the grid, but also have the potential to cause significant deviations from planned power output. In order to integrate this capacity in the electricity system without balancing costs increasing substantially, increasingly accurate wind power forecasts at these sites will be valuable. In addition to this, better forecasting of uncertainty in wind power output allows units to be scheduled effectively, again reducing system costs.

As it was seen in Chapter 4 where MAE was considered, in contrast to the onshore sites investigated the inclusion of stability reduces GPR error at these offshore sites. For Barrow wind farm the reduction in prediction error is approximately 1% from 12 – 72 hours ahead, for both likelihood functions. For shorter forecast horizons the improvement is reduced. At 3 and 6 hours ahead, the improvement for the GPR model with stability is 0.35% for both likelihood functions. Similar trends in CRPS are seen at Burbo Bank wind farm, also shown in Figure 5-11.

Figure 5-11: CRPS for GPR models for offshore wind farms



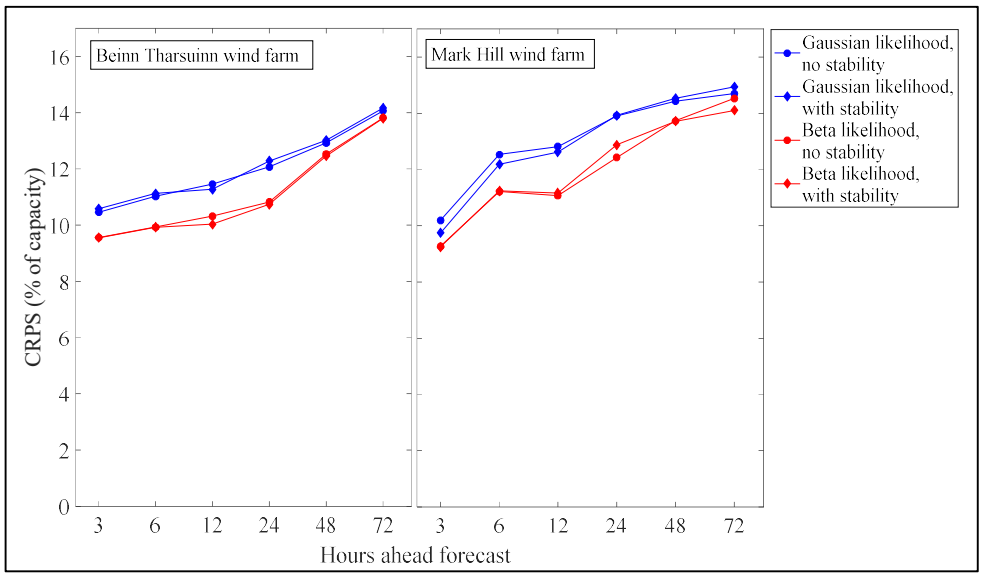
At 3 hours ahead there is no difference between the model with and without stability for either likelihood function. At 6 hours ahead using Obukhov length as an input parameter leads to between 0.3% and 0.4% improvement and for 12 hours to 72 hours the improvement is between 0.6% and 0.9% for both likelihood functions. In addition to the improvement seen when using Obukhov length as an input parameter within the GPR model for offshore wind prediction for both offshore wind farms the GPR model with a beta likelihood function offers improved predictions. At both sites, lower CRPS is seen using the GPR model with a beta likelihood function both with and without stability. In both cases the improvement ranges from 0.6% to 0.9%.

The results shown here indicate that the hybrid NWP and GPR model with a beta likelihood function can provide probabilistic forecasts at offshore wind farms with similar accuracy to onshore wind farms, particularly if the Obukhov length stability parameter is included as an input parameter.

5.3.4 Probabilistic wind power forecast for 2 sites with short datasets

Figure 5-12 considers CRPS for 2 wind farms where only a short data set was available. Probabilistic forecasts using short term data sets can be particularly useful to new wind farms

Figure 5-12: CRPS for GPR models for wind farms with a short data set



. Estimating uncertainty in power output allows a new wind farm to not only bid in the power markets, but to hedge against possible losses by understanding the uncertainty with the forecast. The results shown here mimic those seen in Chapter 4 for MAE, in that there are small differences between the GPR models with and without stability and the overall results are at slightly higher than those seen for the 18 standard sites. At Beinn Tharsuinn wind farm for the GPR model with a beta likelihood function CRPS is approximately 1% lower for forecasts 3-24 hours ahead,

reducing to 0.4% lower at 48 and 72 hours ahead. For the Mark Hill wind farm the improvement when using the beta likelihood function varies between 0.5% and 1.5%. Overall, the results suggest that the beta likelihood function offers better probabilistic predictions. In addition to this, the CRPS for these short sites is only between 1.5% and 2% higher than at the standard sites, showing that reasonable predictions can be made with shorter datasets.

The hybrid NWP and GPR model shows some success in probabilistic predictions for short term data sets. However, further work would be required to establish the length of data set required to produce predictions with confidence.

5.4 Conclusions

This chapter has developed the hybrid NWP and GPR model for probabilistic forecasts and investigated the potential use of a beta likelihood in comparison the Gaussian likelihood function used originally. The results are presented in a probabilistic framework and compared to other probabilistic prediction methods first presented in Table 2-7.

It can be seen from the results presented that the hybrid NWP and GPR model with a beta likelihood function gives improved probabilistic predictions in comparison to the hybrid NWP and GPR model with a Gaussian likelihood. This is due to the fact that the beta likelihood is more appropriate as it is based on a distribution which has upper and lower limits. This is important as wind power output predictions must be positive but less than the wind farm capacity. For the 18 standard sites the hybrid NWP and GPR model with a beta likelihood function outperforms the same model with a Gaussian likelihood in all 4 evaluation criteria used: CRPS, sharpness, resolution and reliability. There is a 0.85% decrease in CRPS for the predictions with a beta likelihood function in comparison to the Gaussian likelihood function at 3 hours ahead, rising to 1.1% at 72 hours ahead. Furthermore, the mean size of the 90% interval is 5% smaller at 3 hours ahead and 10% smaller at 72 hours ahead. This indicates that the hybrid NWP and GPR model with a beta likelihood function

provides a more accurate prediction of the predictive distribution and reduces uncertainty in the forecast. Overall the GPR model with a beta likelihood function is competitive with other probabilistic models for wind power prediction.

As seen in Chapter 4 there is limited difference between the power prediction models using atmospheric stability and the model which does not include stability. This result still holds when considering the models in a probabilistic framework. However, some difference is noted for offshore sites, at which the GPR model including stability gives more accurate predictions than the model without stability. This is true for both the Gaussian likelihood function and the beta likelihood function. This result is significant as the offshore wind capacity in the UK is increasing. The increased prediction accuracy when the Obukhov length is included in this forecast could allow offshore wind power to be more accurately scheduled, thus improving grid reliability.

As discussed in Chapter 4 the prediction of wind power output using the GPR model could possibly be improved if hub height or local wind speeds were available for the wind farms. In general, the use of a probabilistic forecast gives more information than an equivalent deterministic forecast. However, this is only true if the prediction intervals given are not too wide and the model is able to give different level of uncertainty dependent on the circumstances. It can be seen from the results here that the GPR model with a beta likelihood gives uncertainty information which gives added benefit to the user. It is difficult to evaluate specific economic gains which can be made by increasing probabilistic forecast accuracy. However, as probabilistic forecasts are increasingly used by system operators to define reserve requirements or to make economic decisions in electricity markets it is important to ensure accurate forecasts are available.

Chapter 6. An investigation of the economic benefit of employing wind power forecasts.

6.1 Overview

This chapter evaluates the economic value of wind power forecasting in the current UK electricity market. The economic benefit is evaluated by considering the price per MWh paid for electricity to generators participating in the UK electricity market. This chapter considers how forecast accuracy and changes in the imbalance pricing mechanism might impact the price per MWh obtained by wind generators. After studying the economic impact of wind power forecasts in general and the additional income which may be received from more accurate forecasts, the forecasts developed throughout this thesis are considered. In particular, this chapter considers whether the use of the hybrid NWP and GPR model for the prediction of wind power output leads to a higher price for electricity generated in comparison to the use of a simpler forecast in the current UK electricity system. In addition to this, this chapter considers how the deterministic and probabilistic forecasts developed in Chapter 4 and Chapter 5 of this thesis might be employed practically to choose an appropriate bid volume. In doing this, the value the forecasts developed throughout this thesis for a renewable generator participating in the UK electricity market is quantified.

Chapters 3, 4 and 5 have focused on evaluating the quality of predictions using a hybrid NWP and GPR model. Quality of the forecast is evaluated using a statistical framework. In this chapter the link between quality and economic value of a forecast is considered. The analysis performed considers how the price received per MWh generated for a wind farm participating in the electricity market may be affected by wind power forecast accuracy, and how this may change as further modifications are made to balancing system prices. As discussed in detail in Section 2.3.2, imbalance

pricing in the UK is currently undergoing a series of changes. These have been implemented through balancing and settlement code (BSC) modification P305. This modification has changed the imbalance pricing system from a dual pricing system to a single imbalance price and changed the way in which imbalance prices are calculated. One of the key objectives of this chapter is to assess the impact of these changes on price received per MWh generated for a wind farm where electricity output cannot be predicted perfectly in advance.

Whilst there is a wide body of literature covering the applications of wind power forecasting, to the best of the authors knowledge there are no attempts in the literature so far to assess the impact of the changes in the UK imbalance pricing system introduced through BSC modification P305, particularly for wind generators. Therefore, this chapter provides novel insights into the changes which are taking place in the UK electricity market and the importance of wind power forecasting.

As discussed in Section 2.3.4, there are three applications of wind power forecasting which are most frequently discussed in the literature. These are: using wind power forecasts to make unit commitment decisions, to set reserve requirements and to aid participation in electricity markets. The value of a power forecast in these situations can be financial or non-financial. As an example, consider a power forecast used to make unit commitment decisions. That is, a power forecast is used as part of the decision process when the number and type of generators required to provide electricity in a delivery period. If an accurate forecast of wind power output is given, an appropriate combination of units can be planned to match demand. This leads to a financial benefit of reducing costs (less balancing actions will be required) and a non-financial benefit of reduced emissions (reserve requirement can be reduced and fossil fuel plant can be used more efficiently leading to lower emissions).

The value of employing a forecast is quantified in different ways depending on the user and the application. Most frequently in literature the value is quantified from an economic perspective, though value is also discussed in terms of system reliability. Whilst wind power forecasting can lead to non-financial benefits such as reduced

emissions from electricity generation, there is no evidence in the literature of an attempt to quantify this.

One way economic benefit can be quantified is by evaluating the additional revenue achieved for electricity market participants when a forecast is used. Wind power forecasting has particular financial value where balancing markets are used as part of an electricity trading system as these provide financial incentives for accurate power predictions [23]. Economic benefit in terms of increased revenue is discussed by Barthelmie et al. [23], Zugno et al. [133] and Botterud et al. [134]. Barthelmie et al. [23] considered the imbalance penalty of deviating from a forecast and how forecasts can help to reduce this. Zugno et al. [133] used probabilistic forecasts of both wind power and market prices to develop a bidding strategy which optimises market revenue. The method developed was tested on the Eastern Danish price area of the Nordic Power Exchange for 10 months of 2008. They concluded that the use of a bidding strategy alongside forecasts lead to increased profits. Botterud et al. [134] proposed a methodology for optimal day-ahead bidding again concluding that the use of wind power forecasts lead to increased profit and decreased financial risk. This was due to the avoidance of penalties for deviating from the amount of electricity contracted.

Alternatively, economic benefit can be quantified by evaluating the reduction in system costs achieved when a forecast is used. These costs can include the fuel costs for thermal generators needed for reserve capacity, cost of unserved reserve, cost of unserved energy or the additional start-up costs of fast response plant. Pinson et al. [135] used a probabilistic wind power forecast to develop an optimal bidding strategy. When tested on a Dutch wind farm this lead to an average reduction in cost of regulation of €2.20 - €5.90 per MWh. Lowery and O'Malley [128] also considered the impact of wind forecasting error on unit commitment decisions and how forecasting can be used to reduce system costs. They concluded that a precise representation of forecast error can impact the optimal level of reserve a system should carry to reduce system costs.

When forecasting is used to make scheduling decisions or set reserve requirements value is often expressed in terms of system reliability. System reliability is usually quantified in terms of loss of load probability (LOLP) or loss of load expectation (LOLE) [173]. LOLE is a measure of the number of hours per year it is likely supply will not meet demand [174]. LOLP is the probability that supply will not meet demand for a given settlement period [175]. Matos and Bessa [129] developed a tool for setting adequate levels of reserve using a probabilistic wind power forecast. The authors considered how to balance cost of reserve with risk of loss of load. Doherty and O'Malley [127] showed how the use of wind power forecasts with a short forecast horizon can reduce the level of reserve needed whilst maintaining LOLE. Bessa and Matos [176] considered how wind power forecasting used in a probabilistic method for setting reserve requirement. In this case, the LOLP is used to assess whether the reserve requirements set are appropriate.

As identified by the examples above, there are a number of ways in which the value of a wind power forecast could be evaluated. In this chapter, the economic benefit of wind power forecasting is evaluated in terms of price received for electricity sold. This is an important measure of the value of a wind power forecast as in a liberalised electricity market wind farms must trade in electricity markets and generate sufficient income to be profitable. Whilst this work does not account for subsidies paid to wind power generators or other sources of income, it considers the price obtained through the electricity markets which is a major source of income. This is particularly relevant given the changes in support mechanisms for renewable technologies introduced through the electricity market reform (EMR) discussed in Section 2.3.3. As part of the EMR contracts for difference (CFD) replace renewable obligation certificates (ROC) in providing financial support for renewable technologies. However, onshore wind projects are not eligible for CFD auctions therefore in future will rely on income obtained from participation in the electricity markets. As a result of BSC modification P305, between 2015 and 2018 the UK electricity market is undergoing changes which will affect how generators are paid or charged for deviating from their contracted position in the electricity market. As wind power is less predictable than other forms of electricity generation, it is likely

that these changes will affect wind power producers. Because of this, it is interesting to evaluate how the price paid for electricity might be affected by these changes and how the economic benefit of employing a forecast might be affected.

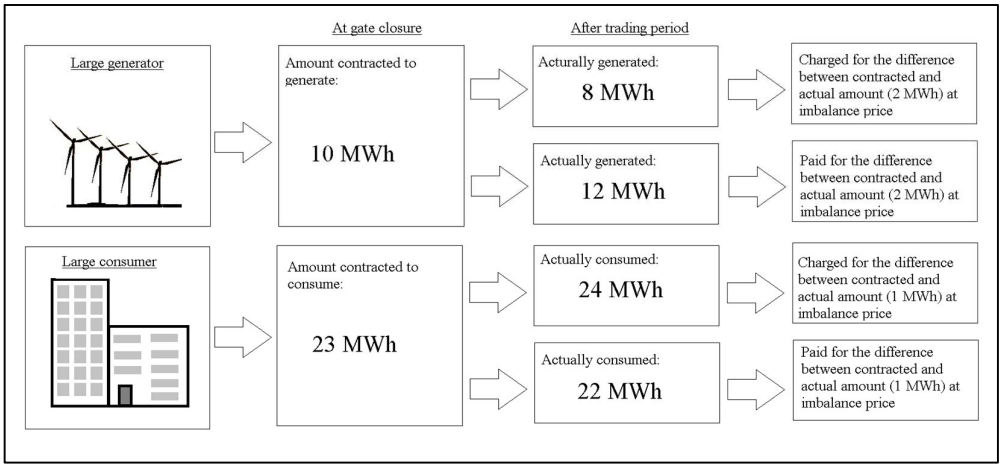
The main objectives of this chapter are:

(i) To evaluate the effect of forecast accuracy on price per MWh obtained for electricity generated.
(ii) To investigate how BSC modification P305 may affect price per MWh obtained for electricity generated and the value of wind power forecasts.
(iii) To assess the value added by the forecasts developed throughout this thesis.
(iv) To demonstrate how the deterministic and probabilistic forecasts developed through this thesis might be used to choose how much electricity to bid in the electricity markets.

6.2 Methodology

As described in Section 2.3.2 electricity is traded in half hour settlement periods ahead of delivery. Bids and offers for electricity demand and generation are made from months and years ahead, up to gate closure, one hour ahead of the trading period. After gate closure, the system operator (the National Grid) accepts bids and offers to increase or decrease demand and generation in order to balance electricity delivery on a short term basis. These are known as balancing actions. Finally, after the trading period, the actual volume of electricity used or generated by a market participant is calculated and payments or charges are received for deviations from the contracted level of generation or demand. These charges/payments are based on the average cost of the most expensive balancing actions, known as the imbalance price. An example of the process for a generator or a consumer participating in the electricity market is shown in Figure 6-1.

Figure 6-1: Diagrammatic example of the imbalance system for a generator or a consumer participating in the electricity market.



Prior to November 2015, the UK used a dual pricing system. Under this system the amount charged for generators producing less than their contracted volume of electricity or consumers using more than their contracted volume was charged at the system buy price (SBP) and the amount paid to generators producing more than their contracted volume of electricity or consumers using less than their contracted volume is paid at the system sell price (SSP). BSC modification P305, introduced in November 2015, changed the imbalance pricing system from a dual pricing system to a single pricing system. Under this system the same price is paid or charged for deviations in either direction. Alongside the change from a dual pricing system to a single imbalance price, BSC modification P305 changed how the imbalance price was calculated. Imbalance prices are calculated using the average cost of the most expensive balancing actions. Prior to November 2015 this was set using the most expensive 500 MWh of balancing actions (PAR500) which reduced to the most expensive 50 MWh of balancing actions (PAR50) in November 2015. This will change to the single most expensive MWh of balancing actions (PAR1) in November 2018 [104]. A full example of the imbalance price calculated under the new system is given in Section 2.3.2. The price obtained per MWh generated is a combination of the price agreed through bidding in the electricity market for

electricity produced as contracted, and payments/penalties for deviating from the contracted volume of electricity.

The work presented in this chapter is split into three sections. Throughout the chapter different forecasts are used to establish the volume of electricity bid in electricity markets, and thus the volume the generator is contracted to sell for the trading period. Firstly, the price received per MWh generated when a perfect forecast is used is discussed. In addition to this, the price per MWh received when a persistence forecast is used. This gives a benchmark at either end of the forecasting spectrum. The perfect forecast shows the price which could be obtained if a forecast with zero error was achievable, whilst a persistence forecast shows what would be obtained if state of the art forecasts were rejected and a simple forecast was used. Secondly, a simulated forecast is used to show the price obtained when a forecast with different levels of error is used to establish the contracted volume of electricity. Finally, an example of how a deterministic and a probabilistic forecast could be used to bid in the electricity market is shown. The price obtained per MWh of electricity generated using each of these methods is investigated. Throughout all three sections the impact of changing PAR volume is discussed.

To begin with, the method used to calculate the price received per MWh generated is introduced. The method is based on an adaptation of the simple economic model introduced by Barthlemie et al. [23]. Barthlemie et al. [23] used a simulated wind power forecasts to assess at what size of wind farm a wind power forecast became economically viable based on electricity prices in 2003. The authors concluded that the accuracy of the forecast and the cost of obtaining it were key factors in quantifying the economic value. The price obtained per MWh generated is a combination of the price agreed through bidding in the electricity market for electricity produced as contracted, and payments/penalties for deviating from the contracted volume of electricity. The price agreed through contracts in the electricity market is referred to as the trade price. This was paid for electricity produced as forecast. Under the system for balancing payments in place in the UK prior to November 2015 the system sell price (SSP) was paid for electricity produced in

excess of the forecasted amount and the system buy price (SBP) was the price which the wind energy generator must buy electricity at if they produced less than the forecasted amount of electricity. Under this system, the price obtained for electricity generated in trading period *t* is given by Equation 6-1.

$$R_t = \begin{cases} P_{for,t} \times TP_t - \left(\left(P_{for,t} - P_{obs,t}\right) \times SBP_t\right) & \text{if } P_{for,t} > P_{obs,t} \\ P_{for,t} \times TP_t + \left(\left(P_{for,t} - P_{obs,t}\right) \times SSP_t\right) & \text{if } P_{for,t} < P_{obs,t} \end{cases} \quad \text{Equation 6-1}$$

where $P_{obs,t}$ and $P_{for,t}$ are observed and forecasted power output for trading period *t* respectively, TP_t is the trade price, SSP_t is system sell price and SBP_t is the system buy price (for trading period *t*). In the work of Barthelmie et al. [23] the trade price was the annual average of SBP and SSP, a fixed value for all values of *t*.

As discussed in Section 2.3.2 BSC modification P305 was introduced in November 2015, changing balancing payments from a dual pricing system to a single pricing system. In this chapter the method outlined by Equation 6-1 for calculating electricity prices is adapted for an electricity system with a single imbalance price (SIP) rather than the dual pricing system used by Barthlemie et al. [23]. In an electricity system with a SIP, the price received for electricity generated in surplus to the forecasted amount is equal to the price paid by the generator when the amount of electricity forecasted is not met. The price received by electricity generators per MWh generated in trading period *t* in a system with a SIP is given by Equation 6-2.

$$R_t = P_{for,t} \times TP_t - SIP_t\left(P_{for,t} - P_{obs,t}\right) \quad \text{Equation 6-2}$$

where $P_{for,t}$ and $P_{obs,t}$ are observed and forecasted power output for trading period *t* respectively, TP_t is the trade price paid for electricity generated as contracted, and SIP_t is the SIP for trading period *t*. The data used throughout this chapter for the trade price, SIP and observed power is described next.

In this chapter market index price (MIP) is used as an estimate of the trade price for electricity. MIP is the average value of all contracts traded in a settlement period. It reflects the wholesale price of electricity at that time [104] and is obtained from

ELEXON [106]. This is slightly different to the strategy taken by Barthelmie et al. [23] which used an annual average of SSP and SBP to estimate the trade price of electricity. Using MIP allows the difference between MIP and SIP to be considered for each trading period, giving a more detailed analysis. However, it can be seen from Table 6-1 that the annual average MIP for 2003 – 2007 is very similar to the trade price used by Barthelmie et al. [23] indicating that both methods suggest similar values for the trade price.

Table 6-1: 'Trade' price for electricity used by Barthelmie et al. [23] in comparison to MIP.

	MIP (£/MWh) [106]	TP used by Barthelmie et al. [23] (£/MWh)
2003	18.37*	19
2004	20.80**	21
2005	36.55	36
2006	38.51	39

*data available for approx. 45% of trading periods. **data available for approx. 20% of trading periods.

The SIP data used in the chapter is based on recalculated imbalance prices published by ELEXON [177]. In order for interested parties to establish how the changes introduced by BSC modification P305 will affect them, ELEXON reproduced imbalance prices for January 2010 – May 2014 based on different PAR volumes [177]. Prior to November 2015 a PAR volume of 500 MWh was used to calculate SIP. SIP for a PAR volume of 500 MWh was not provided by ELEXON [177]. SIP was provided for PAR volumes of 350, 250, 100, 50 and 1 MWh.

The observed power data used in this chapter is from one of the wind farms used in Chapters 4 and 5. The wind farm chosen was Baillie wind farm, a 52 MW wind farm located in Scotland. Wind power output from a single wind farm is assumed to be independent of electricity system prices. This is because individual wind farms are not assumed to influence system prices, due to the fact they usually contribute a relatively small amount to the entire UK electricity system in a single trading period.

This assumption is commonly used in the economic analysis of wind farms for example by Pinson et al. [135]. This is currently a reasonable assumption but may need reassessing as the use of wind energy increases. Because of this assumption, and the lack of available concurrent wind power and pricing data, the analysis was performed using pricing data from 1st January 2013 – 31st December 2013 and wind power data from 1st January 2015 – 31st December 2015.

In each section, Equation 6-2 is evaluated using $P_{obs,t}, SIP_t$ and TP_t as defined above. Different forecasted values are used for $P_{for,t}$ in each section. To begin with, in Section 6.3.1, the impact of forecast accuracy is considered by using a benchmark forecast. Price received per MWh is calculated for a persistence forecast with different forecast horizons using Equation 6-2. The impact using different PAR volumes to calculate SIP is also considered. Subsequently, in Section 6.3.2, price per MWh is considered for simulated forecasts with different levels of systematic bias and random error. The use of simulations allows the effect of different levels of error to be investigated in a more statistically robust way. The forecasts used were simulated in a similar manner to that used by Barthelmie et al. [23]. Simulated forecasts with different levels of error were generated combining the observed wind power time series with a randomly generated error. Errors were assumed to be random variables from a normal distribution. Therefore the forecast power output at time *t* is given by Equation 6-3.

$$P_{for,t} = P_{obs,t} + \varepsilon_{t,\mu,\sigma} \qquad \text{Equation 6-3}$$

where $\varepsilon_{t,\mu,\sigma} \sim \mathrm{N}(\mu, \sigma)$ and $P_{for,t}$ is constrained such that $0 \leq P_{for,t} \leq \text{Capacity}$

The parameters μ, σ used to simulate forecast error are varied to change the level of systematic bias and random error in the simulated forecast. The values of μ, σ considered were $\mu \in [\pm 100\% \text{ of capacity}]$, $\sigma \in [0, 100\% \text{ of capacity}]$. Changing the value of μ changes the level of systematic bias in the forecast. For example if $\varepsilon_{t,\mu,\sigma} \sim \mathrm{N}(-2,0)$ then the forecasted power is 2% lower than the observed value with no random error. This is equivalent to systematically under predicting from a perfect forecast by 2% of capacity. If $\varepsilon_{t,\mu,\sigma} \sim \mathrm{N}(0,1)$, then the bias in the annual forecast

would tend to 0 (there is no under or over prediction). However, forecasts at an individual trading period *t* may deviate from the observed value. This would be observed as different levels of NMAE and NMBE in the forecast. NMBE and NMAE were defined in Section 2.2.4. In this case NMBE over the whole forecast would be close to zero but NMAE would increase with increasing values of σ. For each value of μ, σ 100 forecast time series simulations are made. This enables the impact on price per MWh to be evaluated in a statistically robust manner. Whilst this work does not investigate the actual distribution of forecast errors, nor the impact of error distribution on price per MWh, it allows the impact of systematic bias and random error to be investigated.

For the perfect forecast, the persistence forecast and the simulated forecast the impact on price per MWh is discussed by considering mean price per MWh and the standard deviation of price per MWh. These are calculated using Equation 6-4 and Equation 6-5 respectively.

$$\text{Mean price per MWh} = \frac{1}{n}\sum_{t=1}^{n}\frac{R_t}{P_{obs}} \qquad \text{Equation 6-4}$$

$$\text{Standard deviation of price per MWh} = \sqrt{\frac{\sum_{t=1}^{n}(\frac{R_t}{P_{obs}} - \bar{R}_{MW}\ )^2}{n-1}} \qquad \text{Equation 6-5}$$

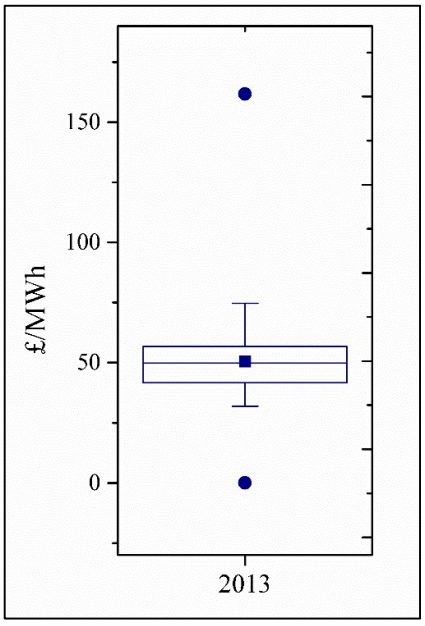
Mean price received for electricity generated per MWh gives an indication of the average level of income which would be received over the year. Standard deviation of the price paid for electricity generated per MWh allows the variation between different time periods to be analysed. This is equally important as it indicates the level economic risk a wind energy generator may be exposed to on a short term basis. Showing these results in terms of price per MWh allows the results to be generalised for other wind farms. The results are therefore independent of the size of wind farm in question. Whilst the size of wind farm will impact the point at which the cost of a wind power forecast is justified, this is not the focus of this work.

6.3 Results

6.3.1 Price obtained using a persistence forecast

To begin with, to set a benchmark position, consider the price per MWh which would be obtained if a perfect forecast was available. In this case, Equation 6-2 is used to calculate price per MWh with $P_{for,t} = P_{obs,t}$. Hence, from Equation 6-2, price received per MWh generated becomes $R_t = P_{for,t} \times MIP_t$. So the price received is a function of MIP and the average price (£/MWh) is the average MIP. It can be seen from Figure 6-2 the average price for the period 1st Jan 2013 – 31st December 2013 would be £50.70/MWh. The maximum price achieved in any trading period would be £161.73/MWh.

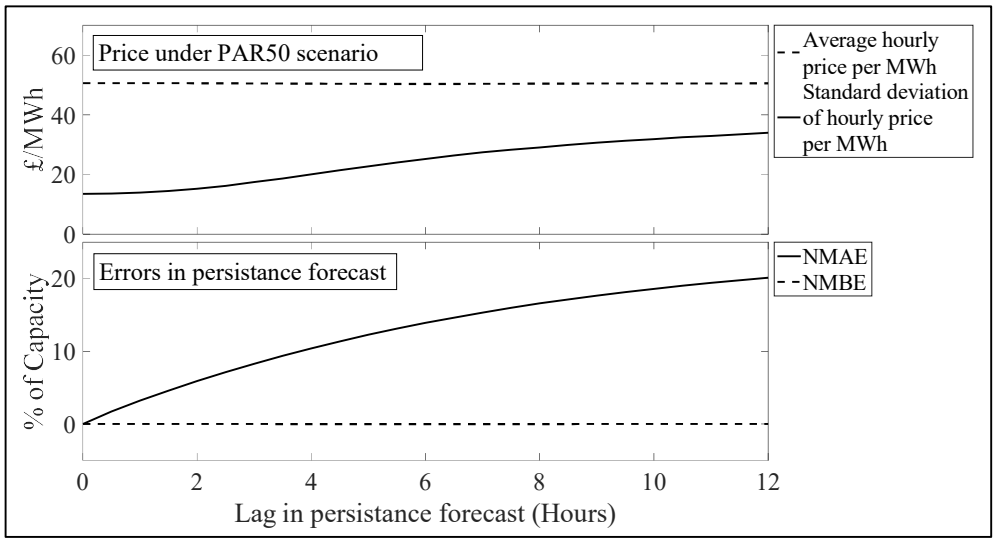
Figure 6-2: Box plot showing Market Index Price for 2013. Box shows 25, 50 and 75 percentiles. Circles show the maximum and minimum values and square shows the mean value.



Next, consider the price per MWh if the persistence forecast was used. The persistence forecast was used in Chapter 3 and Chapter 4 as a benchmark for forecast model performance. Figure 6-3 shows the mean and standard deviation of

price per MWh which would be received if a persistence forecast was used to predict wind power output for the test period. In Figure 6-3 the SIP used is for PAR volume of 50 MWh (PAR50). This allows average and standard deviation of price per MWh under the current (2017) market conditions to be considered. The price per MWh is calculated for increasing forecast horizon or lag in the persistence forecast. The mean and standard deviation of price per MWh are shown alongside the error metrics NMAE and NMBE for the persistence forecast. The equations for NMBE and NMAE were given by Equations 2-14 and 2-15 in Chapter 2. In Figure 6-3 it can be seen that NMAE rises from 3% of capacity at 1 hour ahead to 20% of capacity at 12 hours ahead. NMBE is roughly the same for all forecast horizons, staying at less than 1% of capacity. Because of this, it is difficult to assess the impact of NMBE on price per MWh. From Figure 6-3 it can be seen that there is very little change in average price paid per MWh as forecast horizon and NMAE increase.

Figure 6-3: Price per MWh achieved when the persistence forecast is used with different lag times.

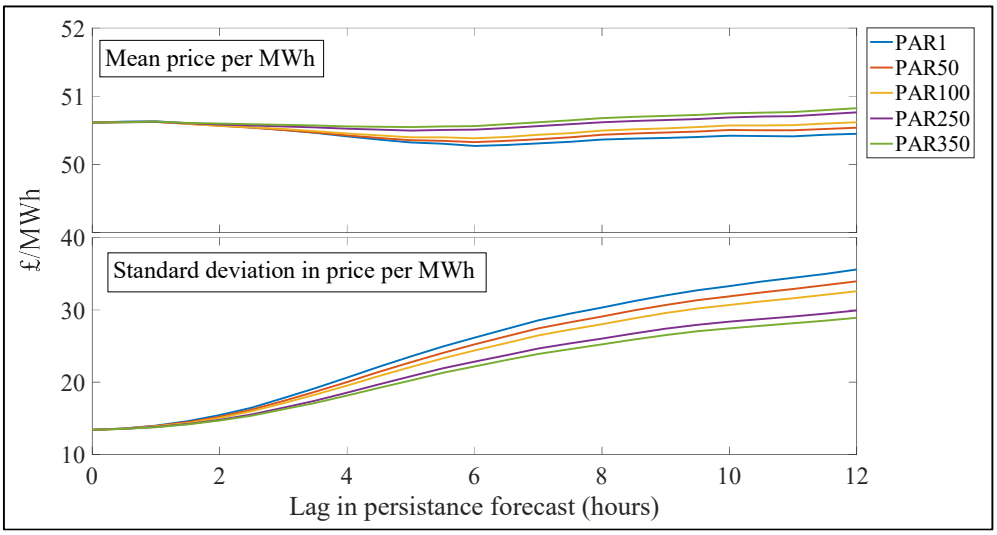


The average price is £50.70/MWh for a perfect forecast and only changes by £0.20/MWh across the forecast. However, the standard deviation varies significantly with increasing forecast error. For a persistence forecast at 1 hour ahead the standard

deviation in hourly price is £13/MWh, whilst at 24 hours ahead this rises to £34/MWh. This indicates that a forecast with higher NMAE may not have a significant impact on mean price per MWh but will impact upon the deviation in hourly prices. Because of this, the income for a wind generator will be more stable when a more accurate forecast is used to bid in the electricity market.

The results shown in Figure 6-3 show the price per MWh which would be received if a persistence forecast was used and the SIP was calculated using a PAR volume of 50 MWh. This reflects the SIP from November 2015 to November 2018. In Figure 6-4 the effect of changing PAR volume is investigated. From Figure 6-4 it can be seen that mean price obtained per MWh decreases with increasing forecast horizon (hence NMAE) when a PAR volume of 1 MWh is used, and remains mostly constant for higher PAR volumes.

Figure 6-4: Price per MWh achieved when the persistence forecast is used with different lag times using different PAR volumes.



The trend in mean and standard deviation of price per MWh seen in Figure 6-4 is the same as seen in Figure 6-3. Mean price per MWh does not vary much and standard deviation of price increases with increasing forecast horizon. However, the

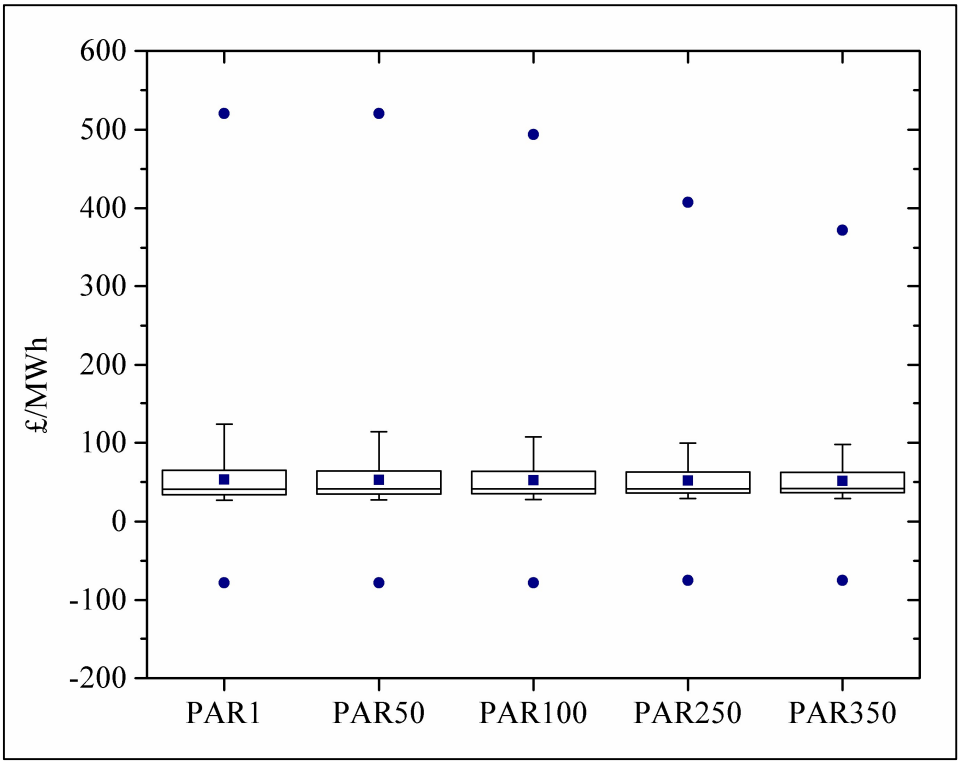
magnitude of change in both mean and standard deviation of price per MWh with increasing forecast error changes with PAR volume.

When the PAR volume is 50 MWh, as seen in Figure 6-3, mean price per MWh varies by £0.20/MWh when the forecast horizon increases from 1 hour to 24 hours. For all PAR volumes there is a decrease in mean price per MWh at around 6 hours ahead, which then increases up to 12 hours ahead. The decrease in mean price per MWh seen in Figure 6-4 at 6 hours ahead correlates with a larger NMBE of -1.5%, and the increase with a smaller NMBE of -0.2%. It is difficult at this point to properly assess the impact of NMBE of price per MWh. However, this will be investigated further in Section 6.3.2. The difference between mean price per MWh for different PAR volumes is most pronounced at 12 hours ahead. At this point for a PAR volume of 1 MWh, there is a decrease of £0.25 in mean price per MWh. For a PAR volume of 350 MWh there is an increase of £0.20 in mean price per MWh when the forecast horizon and NMAE increase. Overall, the effect of decreasing PAR volume on mean price per MWh obtained by a generator is small. Furthermore, Figure 6-4 shows that when the forecast error is larger, the standard deviation in price per MWh increases more rapidly for smaller PAR volumes. For a persistence forecast 12 hours ahead standard deviation in price per MWh is £29 when the PAR volume is 350 MWh and £36 when the PAR volume is 1 MWh. These results indicate that when the PAR volume is reduced to 1 MWh in November 2018 forecast accuracy will have a larger impact on deviation in hourly price. However, there will be very little impact on mean price per MWh when the PAR volume reduces from 50 MWh to 1 MWh in November 2018.

The difference in price per MWh can be explained by considering the distribution of SIP calculated using different PAR volumes. Figure 6-5 shows that the range of prices per MWh seen in the data set used here is larger when the PAR volume is lower, and the interquartile range is larger. This leads to the increase in standard deviation in price per MWh seen in Figure 6-4. This is because when forecast error increases the price paid per MWh is more dependent on the imbalance price, and so there is more variation in the price paid per MWh. The average difference between

MIP and SIP increases as PAR volume decreases. For PAR1 the imbalance price was on average £2.74 higher in 2013 than MIP. For PAR350 SIP was on average £0.84 higher. These differences influence the mean price per MWh. However, the analysis of the persistence forecast indicates that the decreasing forecast accuracy has little impact on mean price per MWh. This implies that while increasing forecast accuracy may not increase mean price per MWh achieved, it could reduce fluctuations in income. This creates a more secure income for wind farm generators. Therefore the increase in forecast accuracy is still valuable, despite not having a significant impact on mean price per MWh.

Figure 6-5: Distribution of SIP with different PAR volumes. Box shows 25, 50 and 75 percentiles. Circles show the maximum and minimum values and squares show the mean value.



Through analysis of the persistence forecast, we begin to understand the impact of a forecast on price per MWh. However, further analysis is required to separate the

impact of systematic bias and random error on price per MWh. Increased systematic bias leads to an increase in NMBE, whilst increase random error leads to an increase in NMAE. In the persistence forecast, NMAE increased with forecast horizon whilst NMBE remained fairly constant. In the next section, systematic bias and random error are explored through the use of simulated forecasts.

6.3.2 Impact of forecast error on price per MWh

Simulated forecasts with different levels of error were generated combining the observed wind power time series with a randomly generated error, as described by Equation 2-3 in Section 6.2. μ, σ are varied to change the level of systematic bias and random error in the simulated forecast. To begin with, consider simulated forecasts where $\mu = 0$ is used in Equation 6-3 to generate a simulated forecast. With $\mu = 0$ the forecast has no systematic bias and σ is varied to change the amount of random error in the forecast. In this case NMAE for the forecast is directly correlated to the level of random error in the forecast. Figure 6-6 shows mean and standard deviation of price per MWh for different PAR volumes in relation to NMAE. From Figure 6-6 it can be seen that as NMAE increases, there is very little difference in mean price per MWh. However, standard deviation in price per MWh increases. This mimics the results seen in Section 6.3.1 but attempts to remove the impact of systematic bias in the forecast on price. In Figure 6-6 mean price per MWh increases by less that £0.05 when NMAE increases up to 20% of capacity and there is no clear trend between different PAR volumes. This is smaller than the difference seen when using the persistence forecast in Figure 6-3. This perhaps indicates that the systematic bias in the persistence forecast had an impact on price per MWh. This follows mathematically from the formula for price and mean price. When the systematic bias approaches zero, $P_{for,t} - P_{obs,t}$ approaches zero in Equation 6-6, hence the impact of SIP on mean price also tends to zero. This results in a nearly constant mean price per MWh equal to annual mean MIP.

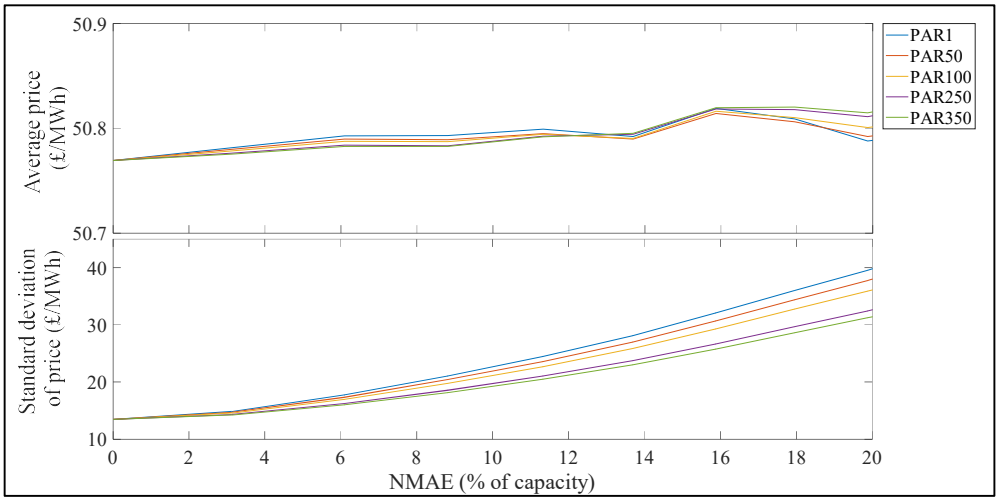
$$\text{Mean price per MWh} = \frac{1}{n}\sum_{t=1}^{n}\frac{R_t}{P_{obs,t}}$$

Equation 6-6

$$= \frac{1}{n}\sum_{t=1}^{n}\frac{P_{for,t} \times MIP_t - (P_{for,t} - P_{obs,t})SIP_t}{P_{obs,t}}$$

In addition, Figure 6-6 shows that the standard deviation in price is higher when the PAR volume is lower, due to the higher variation in SIP when the PAR volume is lower as seen in Figure 6-5.

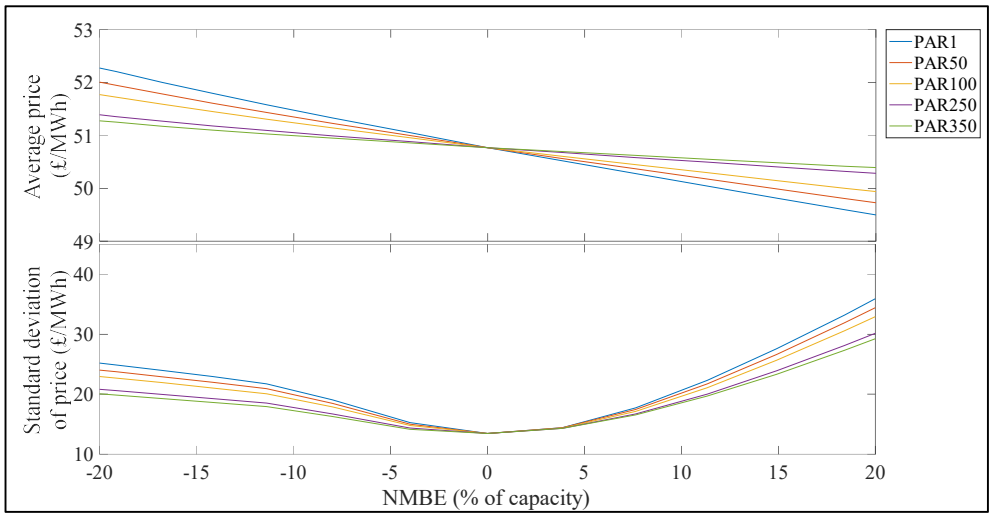
Figure 6-6: Average and standard deviation of price per MWh for simulated forecast with $\mu = 0$



In order to see the impact of systematic bias in a forecast on the price per MWh, next consider simulated forecasts where $\sigma = 0$ is used in Equation 6-3 to generate a simulated forecast. If $\sigma = 0$ and the value of μ is varied then there is no random error in the forecast but the systematic bias is varied. In this case the value of μ is equal to NMBE. A negative value for NMBE indicates an under prediction whilst a positive value indicates an over prediction. Figure 6-7 shows mean and standard deviation of price per MWh against NMBE. It can be seen here in comparison to Figure 6-6 that systematic bias has more of an impact on mean price per MWh than random error has. Also, Figure 6-7 indicates that under predicting leads to an

increase in mean price per MWh whilst over predicting leads to a decrease in mean price per MWh. This is because on average, for all PAR volumes MIP<SIP in 2013. This indicates that for wind farm generators under predicting can lead to a higher mean price per MWh and so it is better financially to under predict than to over predict. This may not hold for all years as the difference between MIP and SIP is variable. This is explored further in Table 6-2 and 6-3.

Figure 6-7: Average and standard deviation of price per MWh for simulated forecast with $\sigma = 0$



For mean price per MWh the difference for an under prediction or an over prediction is approximately symmetric. An under prediction of 20% of capacity leads to an increase in price of £1.45/MWh in comparison to a perfect forecast and an over prediction of 20% of capacity leads to a reduction of the same amount. However, this is not the case for standard deviation of price per MWh. For standard deviation in price per MWh, the increase with an over prediction or under prediction is equal up to NMBE = $\pm$10%. However, for higher or lower NMBE there is more of an increase in the standard deviation for an over prediction than for an under prediction. The trends seen in mean and standard deviation of price are the same for all PAR volumes. However, for both mean and standard deviation larger changes are seen when the PAR volume decreases. When the PAR volume is 1 MWh a 20% under

prediction leads to an increase in mean price per MWh of £1.50 in comparison to a perfect forecast. For a PAR volume of 350 MWh a 20% under prediction leads to an increase in mean price per MWh of £0.50 in comparison to a perfect forecast. This indicates that as the PAR volume decreases the value of a forecast will increase as the difference in income for a deviation from a perfect forecast increases. In addition to this, it will remain beneficial to under predict power output rather than over predict.

Figure 6-7 shows that a higher average price per MWh is obtained if an under prediction is made in comparison to a perfect forecast. This does not necessarily indicate that a forecast which consistently under predicts (has negative systematic bias) is more valuable than a perfect forecast, but rather that value may be obtained from bidding an amount which deviates from a perfect forecast. This is explored further in Section 6.4.

In reality a wind power forecast with no random error (i.e. where $\sigma = 0$) is infeasible. Figures 6-8 and 6-9 explore the combined impact of different levels of systematic bias and random error on mean and standard deviation of price per MWh. To do this μ, σ used in Equation 6-3 to generate a simulated forecast are varied simultaneously.

Figure 6-8 shows the change in mean price per MWh as μ and σ are varied in the simulated forecasts. It can be seen here that the change in systematic bias (μ) has more significant impact on mean price per MWh than random error (σ). The mean price per MWh increases to £53.67 per MWh if an under prediction of 100% of capacity (equivalent to bidding zero for all trading periods) is made and decreases to £47.93 if an over prediction of 100% is made (equivalent to bidding at capacity for all trading periods). Regardless of the systematic bias, increasing the random error by up to 100% of capacity reduces the average price per MWh by approximately £1 per MWh. This indicates that whilst average price per MWh can be increased by a negatively biased forecast, reducing random error in the forecast can increase mean price per MWh by a small amount. This indicates that reducing random error in a

forecast can be valuable and that the impact of systematic bias on mean price increases as random error reduces.

Figure 6-8: Average of price per MWh for simulated forecast with different values of μ and σ for PAR50. The average price is shown by both the vertical axis and the colour bar.

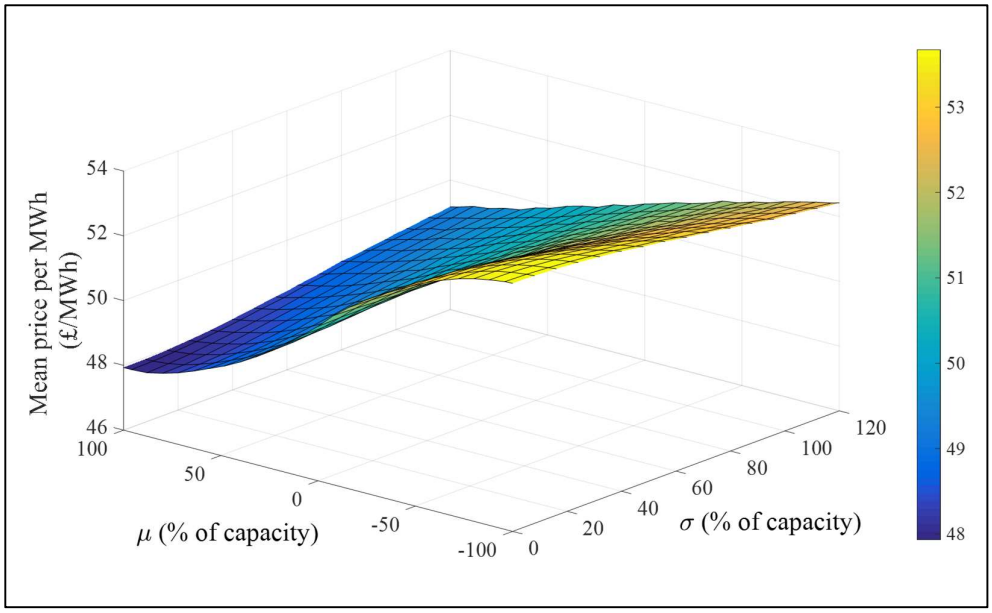
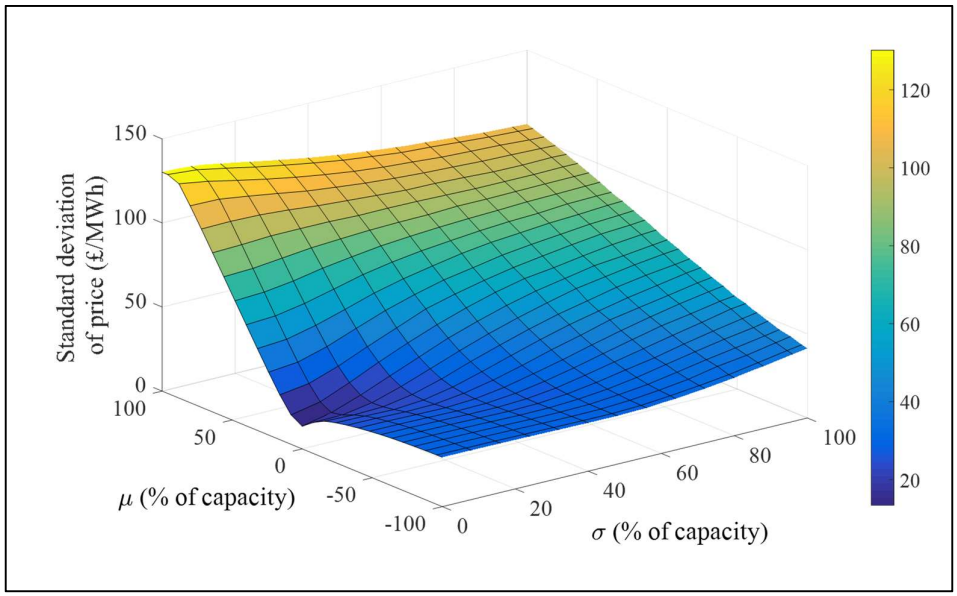


Figure 6-9 evaluates the standard deviation in price per MWh as μ and σ are varied in the simulated forecasts. This shows, as in Figure 6-6 that the standard deviation in higher when a forecast is positively biased and reduced when negatively biased. The standard deviation is lowest for a perfect forecast, indicating that a perfect forecast would create a more stable income than biased forecasts. Furthermore, for any amount of systematic bias the standard deviation is reduced if random error is reduced.

Figure 6-9: standard deviation of price per MWh for simulated forecast with different values of μ and σ for PAR50



Overall this section indicates that both systematic bias and random error impacts mean and standard deviation in price per MWh. The results indicate that despite the change to a single imbalance price, in the absence of a perfect forecast it is still better to under predict than to over predict. This is the same as the results shown by Barthelmie et al. [23]. In addition to this, when the SIP is used, a higher mean price per MWh can be obtained by under predicting. However, this comes with a higher standard deviation in price per MWh. This indicates that whilst the price received over the year might be higher, there will be a larger difference in day to day income which might be an issue to some generators. If this was taken to the extreme then the highest mean price per MWh would be achieved by bidding zero power output for all trading periods. However, this strategy would expose the market participant to significant fluctuations in income. Furthermore, in a competitive market this behaviour is likely to lead to changes in prices which would make the strategy non-profitable.

Whilst a forecast with negative bias might increase mean price per MWh, this could also be achieved by deviating from a perfect forecast. Hence a perfect forecast is still the desired result, as this can be used to derive an appropriate way of choosing an amount to bid in the electricity market.

As discussed in Section 6.2 the analysis performed uses MIP and SIP data from 2013 to calculate price received for each trading period *t* using Equation 6-2. Now consider how the results shown so far might be affected by the use of pricing data from different years. This gives an insight into whether the same trends might be seen using pricing data from other years. To do this, first consider Equation 6-2, which can be rewritten as Equation 6-7.

$$R_t = P_{obs,t} \times MIP_t - (SIP_t - MIP_t)\left(P_{for,t} - P_{obs,t}\right) \quad \text{Equation 6-7}$$

In this, the first term $P_{obs,t} \times MIP_t$ is the price which would be received if a perfect forecast was used. The second term, $(SIP_t - MIP_t)\left(P_{for,t} - P_{obs,t}\right)$ is therefore the difference seen if an imperfect forecast is used. It can be seen that this term is dependent on both the forecast error and the difference between MIP and SIP for each trading period. Therefore, the price per MWh received is dependent on both the forecast error and the difference between MIP and SIP. In order to consider whether the results seen in this section are likely to reflect prices in subsequent years consider the annual average MIP and SIP. One particular issue to bear in mind is the SIP used in the analysis are recalculated imbalance prices generated by ELEXON to reflect historic market prices under the changing PAR volume. ELEXON note that this dataset will not model any behavioural change that might occur due to changes in the imbalance system [177]. In Table 6-2 recalculated imbalance prices from 2010 – 2014 are compared to MIP. From Table 6-2 it can be seen that in 2013 both MIP and SIP for PAR50 were higher than for other years. However, the difference between MIP and SIP was similar to that seen in 2010, 2012 and 2014. The difference between MIP and SIP was significantly lower in 2011, though it is unclear why this occurs. This indicates that the impact of wind power forecast error

on mean and standard deviation of price per MWh seen in Section 6.3.1 and 6.3.2 are likely to be reflective of the trends which would be seen in 2010, 2012 and 2014.

Table 6-2: Annual average MIP and recalculated SIP for 2010 – 2014 (PAR50)
Source: ELEXON Ltd [161]

Year	Annual average MIP (£/MWh)	Annual average SIP (PAR50) (£/MWh)	$\frac{1}{n}\sum_{t=1}^{n}(SIP_t - MIP_t)$ (£/MWh)
2010*	41.96	43.51	1.55
2011	47.83	47.59	-0.24
2012	45.19	48.68	3.49
2013	50.58	52.94	2.36
2014**	44.13	46.36	2.23

*data available from 15/02/2010 – 31/12/2010
**data available from 01/01/2014 – 17/05/2014

Table 6-3: Annual average MIP and actual SIP for 2015-2017 (PAR50).
Source: ELEXON Ltd [161]

Year	Annual average MIP (£/MWh)	Annual average SIP (PAR50) (£/MWh)	$\frac{1}{n}\sum_{t=1}^{n}(SIP_t - MIP_t)$ (£/MWh)
2015*	39.92	37.87	-2.05
2016	38.87	40.03	1.16
2017**	42.82	43.09	0.27

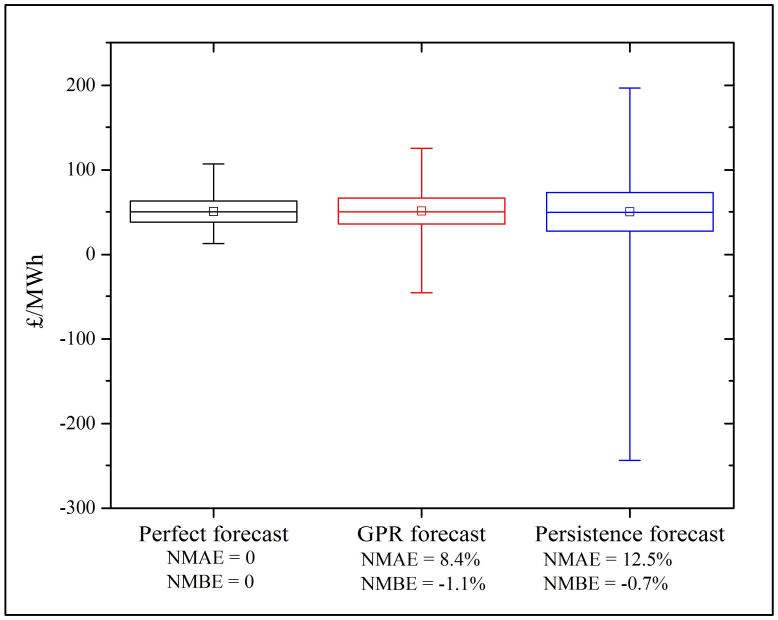
*data available from 05/11/2015 – 31/12/2015
*data available from 01/01/2017 – 31/08/2017

Furthermore, to explore whether the recalculated imbalance prices reflect the difference in MIP and SIP observed once BSC modification was introduced in

November 2015 annual average MIP, SIP and the difference are shown in Table 6-3. So far it is difficult to draw conclusions about the continuing trends in MIP and SIP. In 2016 the difference between MIP and SIP was lower than that seen in the 2010 - 2014 dataset. So far in 2017 there is not much difference between MIP and SIP. However, it is likely that winter prices will change this. This difference will have an impact on the price per MWh and the potential value of using a wind power forecast.

Having considered the overall economic impact of different forecast errors, the value added by the hybrid NWP and GPR model developed in chapter 4 is considered. The hourly income when using the GPR forecast, a perfect forecast and a persistence forecast are compared. Again, these forecasts are for Baillie wind farm, with pricing data from 2013 and a PAR volume of 50 MWh. In Figure 6-10 a box plot of the hourly price per MWh is shown. Both the persistence forecast and the GPR forecast have very small negative bias. As expected, the mean price received per MWh is very similar between all 3 forecasts indicating that over an extended period of time there would be limited difference in total income using either prediction. However, the range in hourly prices is much higher when the persistence forecast is used in contrast to the GPR forecast. The range of prices for the persistence forecast was £440/MWh with a standard deviation of £22.50/MWh whilst for the GPR forecast the range was £170/MWh with a standard deviation of £15.20/MWh. This means that the fluctuation in short term income would be much higher when the persistence forecast is used. In addition to this, the results shown in Figure 6-8 and 6-9 indicate that there may be some benefit in deviating from the predicted output value. Using a more accurate forecast would allow a wind farm operator to choose an appropriate bid volume which manages risk and income with less uncertainty. The next section shows how the deterministic and probabilistic forecasts developed in Chapter 4 and 5 might be used to choose a bid value and the price which would be obtained.

Figure 6-10: Box plot showing price per MWh for a perfect forecast, a hybrid NWP and GPR forecast and a persistence forecast at Baillie wind farm. Box shows the 25, 50 and 75 percentiles, square shows the mean value and the whiskers show the maximum and minimum values.



6.4 Electricity trading with deterministic and probabilistic forecasts

Section 6.3 has demonstrated that with the introduction of BSC modification P305 and the associated changes to the balancing system that the mean price per MWh can be higher when an under prediction is made in comparison to a perfect forecast whilst an over prediction leads to a reduction in mean price per MWh. Furthermore, an under prediction would lower the hourly exposure to deviations in price, creating a steadier income than would be achieved if an over prediction was made. With this in mind, the economic impact of deviating from the most likely prediction of power output is considered.

In this section, how a forecast might be employed to bid in the electricity market is discussed. Of particular interest in this section is how using a probabilistic forecast might differ from the use of a deterministic forecast. Thus far it has been assumed that the most likely power output forecasted is used to bid in the electricity market, and the impact of forecast error is investigated. In Section 2.3.4 the use of wind power forecasts in the derivation of optimal bidding strategies is discussed. Numerous methods exist for choosing the optimal amount to bid in the day ahead electricity markets. Sometimes the optimal bid from an economic perspective may require deviating from the most likely prediction of power output. In this chapter the optimal bidding strategy is not the main concern, as optimal bidding requires forecasting of electricity prices. Instead, the way in which a probabilistic or deterministic forecast could be used is discussed, and the economic impact of doing so. The forecast developed throughout this thesis is used for this purpose. Particularly, the hybrid NWP and GPR power predictions made with a beta likelihood function in Chapter 5 are used. The deterministic forecast is calculated using the most likely value of the predictive distribution calculated in Chapter 5, whilst for the probabilistic forecast the full predictive distribution is used. In Chapter 5 predictions were made for the final 3 months of 2015. Hence price per MWh is calculated for this period.

To begin with, the price per MWh obtained and the standard deviation in price over the year is considered for various deviations from a deterministic forecast. This is done by assuming the bid volume, B_t at trading period *t* is the forecasted amount +/- a percentage of capacity, A, whilst maintaining $0 \leq B_t \leq C$. Hence B_t is calculated using Equation 6-8, and the price per MWh is calculated using Equation 6-9.

$$B_t = \begin{cases} P_{for,t} + \left(\frac{A \times C}{100}\right) & if\ 0 \leq P_{for,t} + \left(\frac{A \times C}{100}\right) \leq C \\ 0 & if\ 0 \geq P_{for,t} + \left(\frac{A \times C}{100}\right) \\ C & if\ P_{for,t} + \left(\frac{A \times C}{100}\right) \geq C \end{cases} \qquad \text{Equation 6-8}$$

where $A \in [-100,100]$

$$R_t = B_t \times TP_t - SIP_t\left(B_t - P_{obs,t}\right) \quad \text{Equation 6-9}$$

In Equation 6-8 $P_{for,t}$ is a deterministic forecast obtained from the method used in Chapter 5 for Baillie wind farm at time *t,* $P_{obs,t}$ is the observed power output, *A* is the percentage under or over bid and *C* is the capacity of the wind farm. In Equation 6-9 TP_t and SIP_t are as used in Equation 6-2.

Next, consider how a bid could be obtained from a probabilistic forecast. In the forecasts developed in Chapter 5 a full predictive distribution is derived. A random variable X_t is used to forecast the possible values for the power output at time *t*. A CDF for the power output at time *t* is given by Equation 6-10. In this, $0 \leq x \leq C$ and $p \in [0,1]$.

$$F_{X_t}(x) = P(X_t \leq x) = p \quad \text{Equation 6-10}$$

The inverse CDF for a value *p* gives a value *x* such that the probability that the forecasted value is less than *x* is equal to *p*. The inverse CDF is given by Equation 6-11.

$$F_{X_t}{}^{-1}(p) = \{x \in [0, C] : F_{X_t}(x) = p\} \quad \text{Equation 6-11}$$

Again, for each trading period *t,* a bid B_t must be defined. Rather than obtaining the bid amount at time *t* by increasing or decreasing the forecasted value by a fixed amount as in Equation 6-8, the bid amount is chosen using the inverse likelihood function. If it is expected price per MWh would be increased through underbidding a value to bid at time *t,* B_t would be calculated by evaluating Equation 6-12 for a low value of *p.* Alternatively, if it is expected price per MWh would be increased through overbidding a value to bid at time *t,* B_t would be calculated by evaluating Equation 6-12 for a high value of *p.* Choosing B_t with $p = 1$ is equivalent to bidding full capacity for all trading periods, i.e. choosing $B_t = C, \forall t$. Whilst choosing B_t with $p = 0$ is equivalent to bidding zero for all trading periods, i.e. choosing $B_t = 0, \forall t$.

$$B_t = {F_{X_t}}^{-1}(p) \qquad \text{Equation 6-12}$$

In Figure 6-11 the economic impact of deviating from the deterministic forecast is explored. It can be seen here that the average price obtained per MWh generated is increased by underbidding from the forecasted amount and decreased by overbidding. Furthermore, any deviation from the forecast leads to an increase in the standard deviation of price per MWh. However, there is a larger increase in standard deviation of price per MWh when overbidding than underbidding. This shows, as in Section 6.3.2, that the economic value of using a forecast to make bidding decisions is likely to become more important as the PAR volume decreases.

Figure 6-11: Economic impact of deviation from a deterministic forecast for Baillie wind farm.

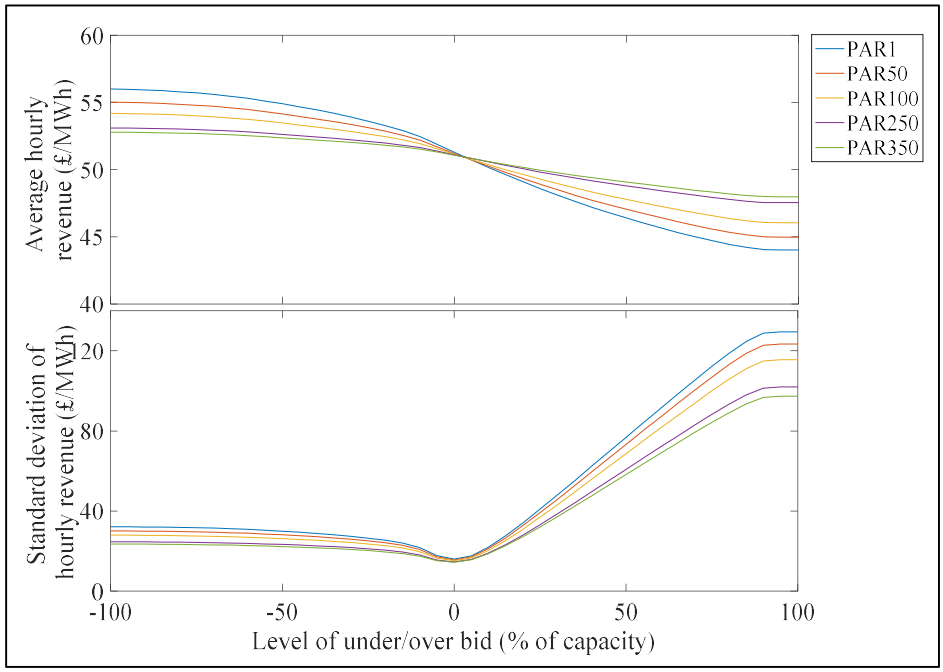
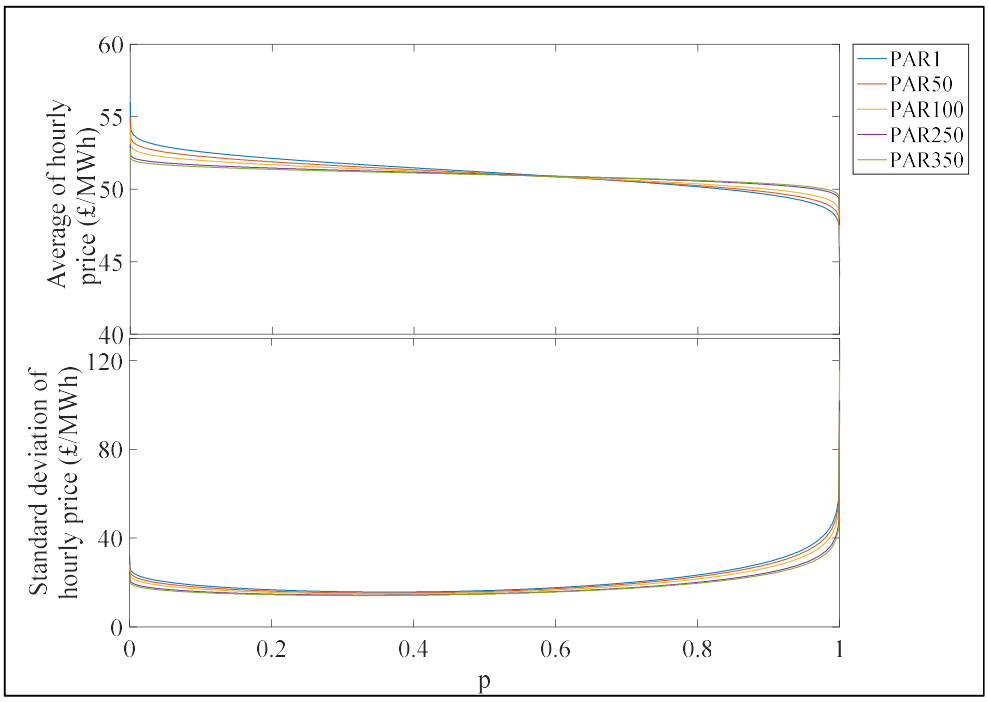


Figure 6-12 explores the economic impact of using a probabilistic forecast with bids chosen using Equation 6-12 for different values of p. In this, it can be seen that there is an increase in mean price per MWh when a low value of p is chosen and a decrease when a high value of p is chosen. However, the impact on standard

deviation of price is smaller. This method of bidding can help limit exposure to changes in MIP and imbalance price using the uncertainty information provided by the probabilistic forecast.

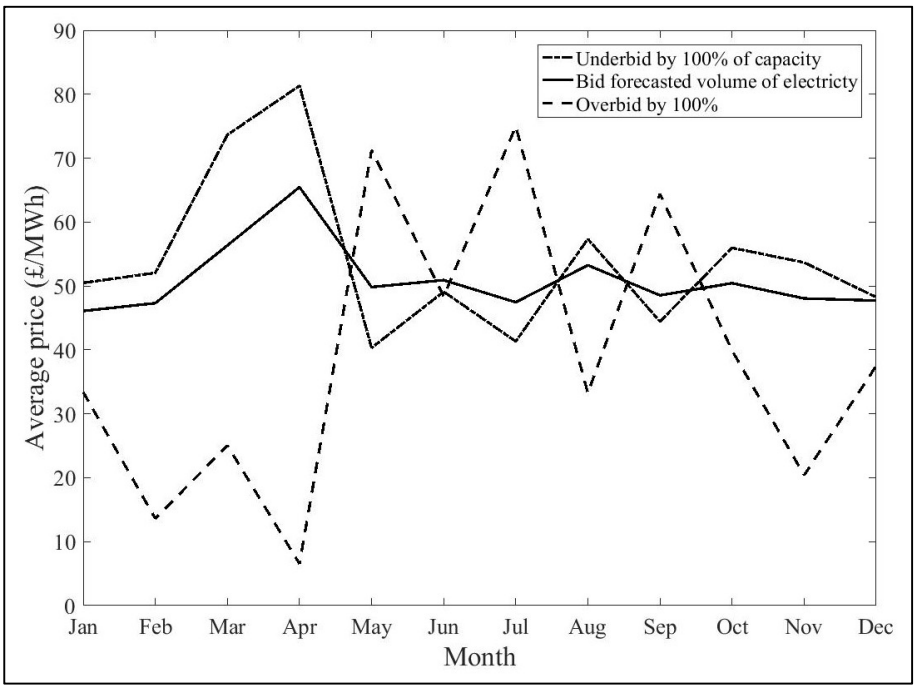
Figure 6-12: Average and standard deviation of hourly price per MWh when a probabilistic forecast is used to choose bid volume.



Figures 6-11 and 6-12 indicate that whilst there may be some increase in the average price per MWh that there is also an increase in the standard deviation in price achieved when a wind power producer under or over bids. This suggests that bidding as close as possible to the actual amount produced can lead to a more stable revenue stream. In order to explore this further, consider the average price received for electricity generated on a monthly basis. Figure 6-13 shows the monthly average price per MWh under extreme underbidding, overbidding and bidding the forecasted volume of electricity. Underbidding leads to an increase in price per MWh over the 12 month period, however, it can be seen in some months it leads to a reduction in price per MWh. In the most extreme case, seen in April in Figure 6-13, the difference between under and overbidding leads to an average difference of

£70/MWh. In addition, even in months where there is less variation between the overbidding and underbidding scenarios the difference between the two is around £20/MWh.

Figure 6-13: Monthly average price per MWh of electricity under different bidding scenarios when the PAR volume is 50MWh.



This means that choosing to underbid based on the annual average price could lead to a significant loss in revenue in some months. For example, in Figure 6-13 in September, choosing to underbid rather than bidding the forecasted volume would lead to a reduction in income of around £175,000 for a wind farm with a 50MW installed capacity. This indicates that forecasting the expected power output from a wind farm can lead to a steadier income for wind farms, and prevent huge fluctuations in income on a monthly basis. This is the case for both deterministic and probabilistic forecasts. In Figure 6-13 an example is shown when the PAR volume is 50 MWh. As the PAR volume decreases further the fluctuations in monthly income will be even larger.

This gives an example of how a deterministic or probabilistic forecast might be used to bid in the electricity market, and shows the potential issues with choosing to bid away from the forecasted volume of electricity. It shows how the additional information provided by the probabilistic forecast could be used to limit exposure to risk. In practice, power output forecasts would be used alongside electricity price forecasts to establish a bidding strategy which optimises price received for electricity generated subject to an acceptable level of risk defined by the wind power producer.

6.5 Conclusions

This chapter has investigated the economic value of a wind power forecast from the perspective of a wind generators participating in the UK electricity market. This has been done by considering the price received per MWh generated. The analysis performed considered the effect of forecast accuracy on price received per MWh and the impact of changing PAR volume for calculating imbalance prices on price received by the generator. Moreover, the hourly price received when the forecasts developed in this thesis were used to define bid volume were considered. Finally, an example of how a deterministic or a probabilistic forecast might be used to bid in the electricity market is given, showing the impact on price of each.

The results shown in Section 6.3 indicate that random error in a wind power forecast has a relatively small impact on mean price per MWh but a larger impact on standard deviation in price per MWh. When the systematic bias in a forecast was zero, a forecast with NMAE of 20% of capacity resulted in less than £0.05 deviation in mean price per MWh in comparison with a perfect forecast. However, the same forecast lead to an increase in standard deviation of up to £21/MWh. This indicates that whilst a reduction in random error in a forecast might not lead to an improvement in mean price per MWh for pricing data from 2013, it can lead to a more stable income stream. It was also shown that systematic bias has a larger impact on mean price than random error. Additionally, it was shown that using the

GPR forecast developed in this thesis to bid in the UK electricity market lead to a reduction in standard deviation of hourly prices of £7.30/MWh in comparison to using a persistence forecast. Because of this, the forecasts developed through this thesis can add value to a wind farm operator participating in the UK electricity market.

Whilst the results in Section 6.3.2 indicate that under predicting wind power output can lead to a higher mean price per MWh than is achieved for a perfect forecast, this doesn't necessarily indicate that a biased forecast is desirable, rather an optimal bidding strategy may require bidding a different volume to the forecasted power output. In addition to this, the results shown in Section 6.3 indicate that as the PAR volume for calculating imbalance prices is reduced the impact of forecast error on price will be amplified. For example, the difference in mean price for a biased forecast is larger for a smaller PAR volume, and the standard deviation in hourly price is higher. This shows that as the PAR volume is reduced a more accurate forecast will be required to maintain a stable source of income. The use of a probabilistic forecast and a deterministic forecast to choose a bid in the electricity market investigated in Section 6.4 indicates that using a probabilistic forecast can limit a generator's exposure to variable prices and decrease the standard deviation in hourly prices.

In general, it can be concluded that a more accurate forecast will limit a generators exposure to the more variable imbalance price and lead to a more stable income. As the PAR volume decreases, imbalance prices are likely to become more variable, and so this may become an important benefit of wind power forecasting. Whether this will lead to a higher mean price depends on the difference between MIP and SIP seen over an extended period of time. It is likely that, as financial support for wind farms is reduced by ROCs being phased out, the price received per MWh will become more important to a wind farms revenue. Therefore, reducing error in a forecast and choosing an appropriate bidding strategy will become increasingly important.

Chapter 7. Conclusions

Wind power has the potential to contribute significantly to the decarbonisation of the electricity supply. Rapid developments have been seen in the wind industry both globally and in the UK in the last decade. This increase in the use of wind and other renewable electricity sources is vital to reducing GHG emissions from the electricity sector. In order to continue this growth, it is important to consider how wind energy can be effectively integrated into the electricity system to maintain a secure and cost effective electricity supply. Wind power forecasting has a part to play in many aspects of integration of wind energy in the UK electricity system, from unit scheduling to meet demand to ensuring a stable income source for wind power producers. Therefore, increasingly accurate wind power forecasts are vital in the electricity sector. In light of this, the main aim of this thesis was to develop and implement a hybrid NWP and GPR model for the prediction of hourly wind speed and power output from 3 – 72 hours ahead. In addition to this, the use of atmospheric stability as an input in the hybrid model, and the value of using a wind power forecast were considered. In this chapter, the overall conclusions drawn from the work performed in this thesis are discussed.

7.1 Research summary and key results

The first object of this thesis was to investigate the use of GPR for the prediction of wind speed and power output, and the second was to explore whether incorporating atmospheric stability into the model could improve model performance. In Chapter 3 the hybrid NWP and GPR model used throughout this thesis was introduced and tested for the prediction of 10 m and hub height wind speeds. The model showed promising results as a method for predicting wind speed, with improvements over an NWP model and the persistence methods. At this stage atmospheric stability was incorporated by subdividing data by atmospheric stability class at the time of the prediction. The hybrid GPR model for the prediction of 10 m wind speed with data

subdivided by atmospheric stability lead to a reduction in MAPE of between 5 and 9% over the Met office NWP. In contrast, the same model without data being subdivided by atmospheric stability class lead to a 2% reduction in MAPE over the met office NWP model. In addition to this, the hybrid NWP and GPR model for the prediction of hub height wind speeds with data subdivided by atmospheric stability class lead to a reduction in MAPE of between 1 and 2% over the same model without subdividing data. Whilst this was not strictly a prediction, it indicated that subdividing data by atmospheric stability class could improve predictions.

This investigation was extended in Chapter 4. In this chapter, the hybrid NWP and GPR model was employed for the prediction of 10 m and hub height wind speeds and wind power for an individual turbine and whole wind farms. In this, the impact of atmospheric stability on predictive performance was tested by using the Obukhov length stability parameter as an input in the model. It was found here that for wind speed prediction the inclusion of the Obukhov length as an input parameter did not lead to improved predictions for 10 m wind speeds but did improve predictions of hub height wind speeds. Using the hybrid NWP and GPR model with the Obukhov length stability parameter as a model input for the prediction of hub height wind speed lead to between 1 and 5% reduction in MAPE over the same model with the Obukhov length stability parameter omitted. The hybrid NWP and GPR model without the stability parameter lead to improved predictions over the NWP model and persistence for both applications. The model was then extended to predict wind power output for both a single turbine and power output for whole wind farms. It was found that including the Obukhov length stability parameter lead to improvements of around 1% in power predictions for a single turbine.

For power predictions at a wind farm the inclusion of the Obukhov length stability parameter did not in general lead to improved prediction. However, some improvement was seen at the 2 offshore sites which were tested. At these sites there was between 0.5 and 1.9% improvement in NMAE when the Obukhov length stability parameter was used as an input parameter in comparison to the simple hybrid NWP and GP model. Overall, the investigation of a hybrid NWP and GPR

model for wind speed and power prediction showed that this method gives results which are comparable to other state of the art forecasting techniques. The inclusion of the Obukhov length stability parameter is beneficial for the prediction of wind power output at offshore wind farms, a result which is important given the current expansion of offshore wind capacity in the UK.

The third objective of this thesis was to explore whether the hybrid NWP and GPR model was appropriate for probabilistic wind power forecasting. In Chapter 5 the model was evaluated in a probabilistic framework for 22 wind farms across the UK. Through this, it is clear that the model generates probabilistic wind power forecasts which are competitive with other state of the art forecasting techniques. In addition to this, it was seen that using a beta likelihood function was more appropriate from a probabilistic perspective than the Gaussian likelihood function. Predicting uncertainty well is an increasingly important aspect of wind power forecasts. It allows wind power generators to define optimal bidding strategies and system operators to define a sufficient level of reserve capacity. The use of the hybrid NWP and GPR model with a beta likelihood function allows predictions to be made with narrower prediction intervals, thus containing less uncertainty. In addition to this, this model is better able to vary the level of uncertainty based on predictive conditions.

The final objective was to consider the economic value of a wind power forecast from a wind farm generators perspective. In Chapter 6, the mean and standard deviation of price received per MWh generated were investigated for forecasts with different levels of systematic bias and random error. In addition to this, Chapter 6 considered the effect of ongoing changes to the imbalance pricing system in the UK electricity market. The results showed that systematic bias had a large effect on both mean price per MWh and standard deviation in price. In addition, random error has a significant impact on standard deviation in price per MWh but less impact on mean price per MWh. This indicated that increasing forecast accuracy has the potential to create a more stable revenue stream for wind farms, leaving them less exposed to fluctuations in day to day price. The hourly price received for electricity generated

when the hybrid NWP and GPR model developed throughout this thesis was used to define the bid volume was also investigated. It was shown that when this method was used to predict wind power output there was no difference in the mean price per MWh obtained over the length of the forecast in comparison to the persistence forecast or a perfect forecast. However, there was a decrease in standard deviation of price per MWh of £7.30. The effect of systematic bias on mean price per MWh also suggests that some benefit may be obtained from bidding a different value to the most likely power output. Many bidding strategies exist which aim to economise on the use of wind power forecasts to improve revenue. This thesis did not focus on developing an optimal strategy, however, did show how a deterministic and a probabilistic forecast might be used to choose a bid volume. The impact on price was discussed for each of these. It was shown that using a probabilistic forecast could help limit a wind farm generators exposure to price fluctuations. Overall, the work in this chapter indicated that the use of a wind power forecast had the ability to limit a wind farms generators exposure to fluctuating electricity prices and charges for deviating from their contracted position. The success of this was dependent on the forecast accuracy. The results shown in this chapter indicate that the forecasts developed throughout this thesis may have some financial value to a wind farm operator participating in the UK electricity market.

Overall, this thesis has demonstrated that a hybrid NWP and GPR model can provide high-quality wind speed and power forecast. Both deterministic and probabilistic forecasts made using this method offered a reduction in error over a simple benchmark model and had the potential to add value for a user.

7.2 Opportunities for further work

The model developed throughout this thesis has proved effective for wind power forecasts. There are various aspects of the model and its applications which would make interesting further research.

In chapters 3 and 4, the hybrid NWP and GPR model was implemented for the prediction of hub height wind speed and power prediction for a single turbine. This was only tested at one site due to the availability of power data and parameters which enabled the calculation of atmospheric stability. It would be beneficial for the future development of the model to test the performance of this model at other sites. The results for the prediction of 10 m wind speeds at 15 MIDAS sites indicated that there was some variation in model performance between sites. The results in Chapter 4 showed that the inclusion of the Obukhov length stability parameter as an input in the model improved the prediction of hub height wind speed and power prediction at a single turbine. It would be beneficial to consider whether this result holds for other sites and whether the site characteristics have any impact on model results.

In chapters 4 and 5, it was shown that for wind farm power prediction the inclusion of the Obukhov length stability parameter as an input had limited impact on prediction accuracy for onshore sites. However, it did improve prediction accuracy for offshore sites where predictions were made using a wind forecast for the nearest onshore location available. This was limited to wind farms where predictions were available for a site within 10 km of the wind farm. It would be beneficial to see whether this result holds for wind farms further offshore, as the majority of UK wind farms are located up to 20 km from the coastline. In addition to this, in the wind farm power predictions made throughout this thesis the wind speed predictions used are for the nearest site. It would be useful to obtain wind speed predictions and observations for offshore sites to attempt to improve predictive accuracy for offshore sites. These extensions could be particularly valuable to both system operators and wind farm operators in the future as the use of offshore wind power in the UK is due to increases significantly.

Additionally, in chapters 4 and 5 the results were shown at 2 sites where only short wind power output datasets were available. This showed that this model was capable of making reasonable wind power predictions for these short datasets. However, it would be useful to consider the length of dataset required to best inform the model.

This would enable the user to create a model which was not too computationally expensive whilst giving accurate predictions.

Finally, chapter 6 showed the value of the forecasts developed throughout this thesis from the perspective of a wind farm operator participating in the UK electricity market and considered the economic impact of improving forecast accuracy. This indicated that there could be some value obtained from increasing forecast accuracy. However, it would be interesting to considering how pricing structures and electricity market mechanisms could be changed to encourage accurate forecasting of electricity generated from non-dispatchable sources. From the perspective of transmission system operators, such as the National Grid in the UK, accurate predictions allow more efficient planning of unit dispatch to meet electricity demand. More efficient unit dispatch can reduce GHG emissions, through reduced reserve requirements, and therefore are important from an environmental perspective. Because of this, it would be useful to consider mechanisms which may increase the economic incentive to increase wind power forecast accuracy.

7.3 Data quality and availability

The availability of high quality wind power output and wind speed forecast data was crucial to the success of the forecasting technique developed throughout this thesis. Furthermore, the availability of data is key to developing efficient, low emission electricity systems. However, obtaining data from the wind industry is not an easy task. Firstly, wind power output data was required. This is generally considered to be commercially sensitive and so only aggregated data was available. This was obtained from Elexon in the form of hourly output from a whole wind farm. It would have been useful in this work to have more detail than this. For example, knowing the number of turbines operational at any one time or individual turbine performance may have allowed wind power forecasts to be improved.

In addition to this, atmospheric stability predictions were only available at a small selection of sites, limiting the site choice and the potential to explore how site

characteristics affected prediction accuracy of wind speed or power output. In particular, the availability of atmospheric stability predictions at more coastal sites would allow a more in-depth study into offshore wind power predictions using the hybrid NWP and GRP model.

On a smaller scale, the data obtained for a single which was used to look at hub height wind speed and power predictions in Chapters 3 and 4 was good quality wind speed and power data available at 10-minute intervals over 3 years. This was very useful for looking at model performance. However, this level of detailed data was only available at one site which limited the potential for further study.

Finally, the lack of publicly available datasets mean that it is difficult for model performance to be compared across different methods seen in literature. This makes it difficult to consider whether new models developed can outperform other models available.

Generally, further availability of high quality detailed wind speed and power data over a selection of sites has the potential to advance wind forecasting techniques and encourage development in the area. Such data is likely to lead to increasingly accurate wind power forecasts and a greater use of such forecast in the wind industry.

7.4 Wider impact

Overall, the aim of reducing emissions from the UK electricity system will not be achieved solely through an increase in renewable electricity capacity. A system-wide approach will need to be employed, making use of a wide range of technologies. For example, reductions in emissions will be achieved through increasing efficiency, reducing demand and increasing the use of renewable electricity. Wind energy alone may not provide the increase in renewable electricity required. However, it will provide a significant contribution, particularly in the UK

where there is an abundant wind resource and a large range of suitable locations for offshore wind farms.

A steady reduction in annual electricity demand since around 2007 and an increase in the use of renewable energy has helped the UK towards its emissions targets so far. However, the UK government's announcement of plans to end the sale of petrol and diesel car and van sales by 2040 is likely to result in an increase in the use of electric vehicles and with it, an increase in electricity demand in the future. In addition to this, increasing the use of variable renewable energy sources further without careful consideration of whole system security could lead to significant increases in the cost of delivering electricity. We have a societal responsibility to ensure that delivering a low carbon electricity system does not lead to an excessive increase in the cost of electricity to consumers, as this is likely to significantly impact upon low-income households. Because of this, continued research focusing on ways to effectively manage renewable energy generation are important.

The forecasting method developed throughout this thesis contributes to the body of literature which has the potential to increase the efficiency of integration of wind energy in the electricity system and reduce the costs associated with this. This has benefits for electricity consumers, generators and the global community who collectively benefit from attempts to reduce GHG emissions. The hybrid NWP and GPR model could be implemented by a number of users as it has shown great potential for the prediction of wind power output. The investigation of forecasting techniques and of input parameters that can increase predictive accuracy continues to be an important area of research. It is unlikely that one method can provide accurate predictions in all scenarios at all sites. For example, as shown in this thesis, the inclusion of the Obukhov length stability parameter is effective in increasing prediction accuracy for power output at offshore sites, but not at onshore sites. Because of this, the work presented throughout this thesis provides an important contribution to the available literature on wind power forecasting.

References

[1] IPCC. 2014, *Impacts, Adaptation, and vulnerability. Summary for Policymakers*.

[2] IPCC. 2014, *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, [Core Writing Team: R.K. Pachauri and L.A. Meyer (eds.)] IPCC, Geneva, Switzerland, pp. 151.

[3] N.H. Stern. 2007, *The economics of climate change: the Stern review*, Cambridge University press, ISBN: 0521700809

[4] United Nations Framework Convention on Climate Change. 2014, *First steps to a safer future: Introducing The United Nations Framework Convention on Climate Change*.

[5] United Nations Framework Convention on Climate Change. *Kyoto protocol*, Accessed: Jan 2016, http://unfccc.int/kyoto_protocol/items/2830.php

[6] United Nations Framework Convention on Climate Change. *Paris Agreement*, Accessed: Oct 2016, http://unfccc.int/paris_agreement/items/9485.php

[7] European Commission. 2010, *Energy 2020: A strategy for competitive, sustainable and secure energy*

[8] European Commission. 2014, *A policy framework for climate and energy in the period from 2020 to 2030*

[9] HM Government. 2008, *Climate Change Act (2008)*.

[10] HM Government. 2016, *Government response to the Committee on Climate Change: Progress on meeting carbon budgets*

[11] DECC. 2011, *UK Renewable Energy Roadmap*.

[12] International Energy Agency (IEA). 2017, *Monthly Electricity Statistics*.

[13] Global Wind Energy Council. 2015, *Global Wind Report Annual Market Update*.

[14] BEIS. 2016, *Energy Trends: December 2016, special feature article - Electricity generation and supply figures for Scotland, Wales, Northern Ireland and England, 2012 to 2015*.

[15] BEIS. 2017, *Energy Trends: March 2017*.

[16] BEIS. 2017, *UK Energy Statistics - 2016 provisional data*.

[17] BEIS. *Renewable Energy Planning Database, extract September 2017*, Accessed: January 2017, https://www.gov.uk/government/collections/renewable-energy-planning-data

[18] BEIS. 2017, *Energy Trends Section 6: Renewables*.

[19] Wind Power Monthly. *Ten of the biggest turbines*, Accessed: Sept 2017, http://www.windpowermonthly.com/10-biggest-turbines

[20] G. Boyle. 2007, *Renewable electricity and the grid: the challenge of variability*, Earthscan, ISBN: 1849772339

[21] Ofgem. 2016, *Electricity System Operator incentives from April 2017*, Accessed: Sept 2017, https://www.ofgem.gov.uk/system/files/docs/2016/08/electricity_system_operator_incentives_from_2017.pdf

[22] Danish Wind Energy Association. 2016, *The Danish Market*.

[23] R. Barthelmie, F. Murray, S. Pryor. 2008, *The economic benefit of short-term forecasting for wind energy in the UK electricity market*, Energy Policy, 36, pp. 1687-1696.

[24] G. Giebel, R. Brownsword, G. Kariniotakis, M. Denhard, C. Draxl. 2011, *The state-of-the-art in short-term prediction of wind power: A literature overview*, Accessed: Jan 2015,

[25] R.B. Stull. 2012, *An introduction to boundary layer meteorology*, Springer Science & Business Media, ISBN: 9400930275

[26] D. Baldocchi. 2012, *Wind and Turbulence (Part 2): Surface Boundary Layer - Theory and Principles*

[27] J.F. Manwell, J.G. McGowan, A.L. Rogers. 2010, *Wind energy explained: theory, design and application*, John Wiley & Sons, ISBN: 0470686286

[28] The Met Office. 2014, *The Coriolis effect*.

[29] E.L. Petersen, N.G. Mortensen, L. Landberg, J. Højstrup, H.P. Frank. 1997, *Wind power meteorology*, Risø National Laboratory, Roskilde, Denmark, Technical Document No. Risø-I-1206 (EN), pp.

[30] R.J. Barthelmie. 1999, *The effects of atmospheric stability on coastal wind climates*, Meteorological Applications, 6, pp. 39-47.

[31] H. Bergström, P.-E. Johansson, A.-S. Smedman. 1988, *A study of wind speed modification and internal boundary-layer heights in a coastal region*, Boundary-Layer Meteorology, 42, pp. 313-335.

[32] J.S. Irwin. 1979, *A theoretical variation of the wind profile power-law exponent as a function of surface roughness and stability*, Atmospheric Environment (1967), 13, pp. 191-194.

[33] D. Spera, T. Richards. 1979, *Modified power law equations for vertical wind profiles*, pp.

[34] S. Emeis, M. Turk. 2007, *Comparison of Logarithmic Wind Profiles and Power Law Wind Profiles and their Applicability for Offshore Wind Profiles*, in: J. Peinke, P. Schaumann, S. Barth (Eds.) Wind Energy: Proceedings of the Euromech Colloquium, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 61-64.

[35] S.M. Weekes, A.S. Tomlin. 2013, *Evaluation of a semi-empirical model for predicting the wind energy resource relevant to small-scale wind turbines*, Renewable Energy, 50, pp. 280-288.

[36] J. Wieringa. 1992, *Updating the Davenport roughness classification*, Journal of Wind Engineering and Industrial Aerodynamics, 41, pp. 357-368.

[37] N. Jelley. 2017, *A Dictionary of Energy Science*, Oxford University Press

[38] J.N. Libii. 2013, *Comparing the calculated coefficients of performance of a class of wind turbines that produce power between 330 kW and 7,500 kW*, World Transactions on Engineering and Technology Education, 11, pp. 36-40.

[39] The Wind Power. *Database of wind turbine power curves*, Accessed: June 2016, www.thewindpower.net

[40] S. Wharton, J. K. Lundquist. 2012, *Atmospheric stability affects wind turbine power collection*, Environmental Research Letters, 7, pp. 014005.

[41] J. Rohatgi, G. Barbezier. 1999, *Wind turbulence and atmospheric stability — Their effect on wind turbine output*, Renewable Energy, 16, pp. 908-911.

[42] J. Sumner, C. Masson. 2006, *Influence of Atmospheric Stability on Wind Turbine Power Performance Curves*, Journal of Solar Energy Engineering, 128, pp. 531-538.

[43] Met Office. *Met Office Numerical Weather Prediction models*, Accessed: June 2015, http://www.metoffice.gov.uk/research/modelling-systems/unified-model/weather-forecasting

[44] S.S. Soman, H. Zareipour, O. Malik, P. Mandal. 2010, *A review of wind power and wind speed forecasting methods with different time horizons*, North American Power Symposium (NAPS).

[45] M. Nodet. 2012, *Introduction to data assimilation*.

[46] Met Office. *Data assimilation - optimising the use of observations in data assimilation*, Accessed: March 2017, https://www.metoffice.gov.uk/research/weather/data-assimilation/use-of-observations

[47] P.L. Houtekamer, H.L. Mitchell. 1998, *Data Assimilation Using an Ensemble Kalman Filter Technique*, Monthly Weather Review, 126, pp. 796-811.

[48] Met Office. *Data assimilation methods*, Accessed: April 2017, https://www.metoffice.gov.uk/research/weather/data-assimilation

[49] Met Office. *UK observations network*, Accessed: Jan 2015, https://www.metoffice.gov.uk/guide/weather/observations-guide/uk-observations-network

[50] L. Buisson, W. Thuiller, N. Casajus, S. Lek, G. Grenouillet. 2010, *Uncertainty in ensemble forecasting of species distribution*, Global Change Biology, 16, pp. 1145-1157.

[51] L. Yu, S. Wang, K.K. Lai. 2005, *A novel nonlinear ensemble forecasting model incorporating GLAR and ANN for foreign exchange rates*, Computers & Operations Research, 32, pp. 2523-2541.

[52] F.A. Eckel, C.F. Mass. 2005, *Aspects of Effective Mesoscale, Short-Range Ensemble Forecasting*, Weather and Forecasting, 20, pp. 328-350.

[53] Z. Toth, Y. Zhu, T. Marchok. 2001, *The Use of Ensembles to Identify Forecasts with Small and Large Uncertainty*, Weather and Forecasting, 16, pp. 463-477.

[54] Y. Zhu, Z. Toth, R. Wobus, D. Richardson, K. Mylne. 2002, *The Economic Value Of Ensemble-Based Weather Forecasts*, Bulletin of the American Meteorological Society, 83, pp. 73-83.

[55] R.N. Tubbs, S.P. Ballard, D. Simonin, Z. Li, J.-F. Caron, G. Kelly, L. Hawkness-Smith, H. Buttery, C. Piccolo, B. Golding, C. Pierce, N. Gaussiat. 2012, *Nowcasting for a British summer*.

[56] J.L. Torres, A. García, M. De Blas, A. De Francisco. 2005, *Forecast of hourly average wind speed with ARMA models in Navarre (Spain)*, Solar Energy, 79, pp. 65-77.

[57] R.G. Kavasseri, K. Seetharaman. 2009, *Day-ahead wind speed forecasting using f-ARIMA models*, Renewable Energy, 34, pp. 1388-1393.

[58] S.A. Kalogirou. 2001, *Artificial neural networks in renewable energy systems applications: a review*, Renewable and Sustainable Energy Reviews, 5, pp. 373-401.

[59] A. More, M.C. Deo. 2003, *Forecasting wind with neural networks*, Marine Structures, 16, pp. 35-49.

[60] G. Li, J. Shi. 2010, *On comparing three artificial neural networks for wind speed forecasting*, Applied Energy, 87, pp. 2313-2320.

[61] G. Li, J. Shi, J. Zhou. 2011, *Bayesian adaptive combination of short-term wind speed forecasts from neural network models*, Renewable Energy, 36, pp. 352-359.

[62] I.J. Ramirez-Rosado, L.A. Fernandez-Jimenez, C. Monteiro, J. Sousa, R. Bessa. 2009, *Comparison of two new short-term wind-power forecasting systems*, Renewable Energy, 34, pp. 1848-1854.

[63] E. Cadenas, W. Rivera. 2009, *Short term wind speed forecasting in La Venta, Oaxaca, México, using artificial neural networks*, Renewable Energy, 34, pp. 274-278.

[64] J. Catalão, H. Pousinho, V. Mendes. 2011, *Short-term wind power forecasting in Portugal by neural networks and wavelet transform*, Renewable Energy, 36, pp. 1245-1251.

[65] I.G. Damousis, M.C. Alexiadis, J.B. Theocharis, P.S. Dokopoulos. 2004, *A fuzzy model for wind speed prediction and power generation in wind parks using spatial correlation*, IEEE Transactions on Energy Conversion, 19, pp. 352-361.

[66] H. Liu, H.-Q. Tian, C. Chen, Y.-f. Li. 2010, *A hybrid statistical method to predict wind speed and wind power*, Renewable Energy, 35, pp. 1857-1861.

[67] K.A. Larson, K. Westrick. 2006, *Short-term wind forecasting using off-site observations*, Wind Energy, 9, pp. 55-62.

[68] I.T. Jolliffe, D.B. Stephenson. 2012, *Forecast verification: a practitioner's guide in atmospheric science*, John Wiley & Sons, ISBN: 1119961076

[69] C.J. Willmott, K. Matsuura. 2005, *Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance*, Climate Research, 30, pp. 79.

[70] T. Chai, R.R. Draxler. 2014, *Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature*, Geosci. Model Dev., 7, pp. 1247-1250.

[71] N. Chen, Z. Qian, X. Meng. 2013, *Multistep wind speed forecasting based on wavelet and Gaussian processes*, Mathematical Problems in Engineering, 2013, pp.

[72] P. Louka, G. Galanis, N. Siebert, G. Kariniotakis, P. Katsafados, I. Pytharoulis, G. Kallos. 2008, *Improvements in wind speed forecasts for wind power prediction purposes using Kalman filtering*, Journal of Wind Engineering and Industrial Aerodynamics, 96, pp. 2348-2362.

[73] N. Chen, Z. Qian, I.T. Nabney, X. Meng. 2014, *Wind power forecasts using Gaussian processes and numerical weather prediction*, Power Systems, IEEE Transactions on, 29, pp. 656-665.

[74] F. Shu, J.R. Liao, R. Yokoyama, C. Luonan, L. Wei-Jen. 2009, *Forecasting the Wind Generation Using a Two-Stage Network Based on Meteorological Information*, Energy Conversion, IEEE Transactions on, 24, pp. 474-482.

[75] P. Pinson, H.A. Nielsen, J.K. Møller, H. Madsen, G.N. Kariniotakis. 2007, *Non-parametric probabilistic forecasts of wind power: required properties and evaluation*, Wind Energy, 10, pp. 497-516.

[76] T. Gneiting, K. Larson, K. Westrick, M.G. Genton, E. Aldrich. 2006, *Calibrated Probabilistic Forecasting at the Stateline Wind Energy Center*, Journal of the American Statistical Association, 101, pp. 968-979.

[77] T. Gneiting, A.E. Raftery. 2007, *Strictly proper scoring rules, prediction, and estimation*, Journal of the American Statistical Association, 102, pp. 359-378.

[78] S. Lerch, A. Jordan, F. Krüger. 2016, *scoringRules-A software package for probabilistic model evaluation*, EGU General Assembly Conference Abstracts.

[79] J.B. Bremnes. 2004, *Probabilistic wind power forecasts using local quantile regression*, Wind Energy, 7, pp. 47-54.

[80] A.U. Haque, M.H. Nehrir, P. Mandal. 2014, *A hybrid intelligent model for deterministic and quantile regression approach for probabilistic wind power forecasting*, IEEE Transactions on Power Systems, 29, pp. 1663-1672.

[81] G. Sideratos, N.D. Hatziargyriou. 2012, *Probabilistic wind power forecasting using radial basis function neural networks*, IEEE Transactions on Power Systems, 27, pp. 1788-1796.

[82] J.W. Messner, A. Zeileis, J. Broecker, G.J. Mayr. 2014, *Probabilistic wind power forecasts with an inverse power curve transformation and censored regression*, Wind Energy, 17, pp. 1753-1766.

[83] H.A. Nielsen, T.S. Nielsen, H. Madsen, G. Giebel, J. Badger, L. Landberg, K. Sattler, L. Voulund, J. Tofting. 2006, *From wind ensembles to probabilistic*

information about future wind power production--results from an actual application, International Conference on Probabilistic Methods Applied to Power Systems.

[84] H.A. Nielsen, H. Madsen, T.S. Nielsen. 2006, *Using quantile regression to extend an existing wind power forecasting system with probabilistic forecasts*, Wind Energy, 9, pp. 95-108.

[85] A. Carpinone, R. Langella, A. Testa, M. Giorgio. 2010, *Very short-term probabilistic wind power forecasting based on Markov chain models*, IEEE 11th International Conference on Probabilistic Methods Applied to Power Systems (PMAPS).

[86] R.J. Bessa, V. Miranda, A. Botterud, Z. Zhou, J. Wang. 2012, *Time-adaptive quantile-copula for wind power probabilistic forecasting*, Renewable Energy, 40, pp. 29-39.

[87] J. Juban, L. Fugon, G. Kariniotakis. 2007, *Probabilistic short-term wind power forecasting based on kernel density estimators*, European Wind Energy Conference and exhibition, EWEC.

[88] Y. Zhang, J. Wang, X. Luo. 2015, *Probabilistic wind power forecasting based on logarithmic transformation and boundary kernel*, Energy Conversion and Management, 96, pp. 440-451.

[89] P. Kou, D. Liang, F. Gao, L. Gao. 2014, *Probabilistic wind power forecasting with online model selection and warped Gaussian process*, Energy Conversion and Management, 84, pp. 649-663.

[90] J. Tastu, P. Pinson, P.-J. Trombe, H. Madsen. 2014, *Probabilistic forecasts of wind power generation accounting for geographically dispersed information*, IEEE Transactions on Smart Grid, 5, pp. 480-489.

[91] P. Kou, F. Gao, X. Guan. 2013, *Sparse online warped Gaussian process for wind power probabilistic forecasting*, Applied Energy, 108, pp. 410-428.

[92] P. Pinson, G. Papaefthymiou, B. Klockl, J. Verboomen. 2009, *Dynamic sizing of energy storage for hedging wind power forecast uncertainty*, 2009 IEEE Power & Energy Society General Meeting.

[93] J. Garcia-Gonzalez, R.M.R.d.l. Muela, L.M. Santos, A.M. Gonzalez. 2008, *Stochastic Joint Optimization of Wind Generation and Pumped-Storage Units in an Electricity Market*, IEEE Transactions on Power Systems, 23, pp. 460-468.

[94] N.P. Padhy. 2004, *Unit commitment-a bibliographical survey*, IEEE Transactions on Power Systems, 19, pp. 1196-1205.

[95] B. Saravanan, S. Das, S. Sikri, D. Kothari. 2013, *A solution to the unit commitment problem--a review*, Frontiers in Energy, 7, pp. 223.

[96] B.J. Kirby. 2003, *Spinning reserve from responsive loads*.

[97] P. Denholm, M. Hand. 2011, *Grid flexibility and storage required to achieve very high penetration of variable renewable electricity*, Energy Policy, 39, pp. 1817-1830.

[98] P.D. Lund, J. Lindgren, J. Mikkola, J. Salpakari. 2015, *Review of energy system flexibility measures to enable high levels of variable renewable electricity*, Renewable and Sustainable Energy Reviews, 45, pp. 785-807.

[99] H. Kondziella, T. Bruckner. 2016, *Flexibility requirements of renewable energy based electricity systems–a review of research results and methodologies*, Renewable and Sustainable Energy Reviews, 53, pp. 10-22.

[100] N. Kumar, P. Besuner, S. Lefton, D. Agan, D. Hilleman. 2012, *Power plant cycling costs*, National Renewable Energy Laboratory (NREL), Golden, CO.

[101] First Hydro Company. *Dinorwig Power station*, Accessed: May 2016, http://www.fhc.co.uk

[102] Ofgem. 2015, *GB electricity market*.

[103] DECC. 2014, *Technical Assessment of the Operation of Coal & Gas Fired Plants*.

[104] A.J. Pimm. 2016, *The Effect of Sharpened Imbalance Prices on the Value of Electricity Storage*, Offshore energy and storage symposium, Malta.

[105] Ofgem. 2015, *Balancing and Settlement Code (BSC) P305: Electricity Balancing Significant Code Review Developments.*

[106] ELEXON Ltd. 2017, *Market Index Definition Statement for Market Index Data Provider(s) (version 8.0)*.

[107] Ofgem. 2017, *Wholesale energy markets in 2016*.

[108] DECC. 2013, *Energy Act Policy*.

[109] DECC. 2013, *Electricity Market Reform: Capacity Market: Detailed Design Proposals*.

[110] BEIS. 2016, *Policy paper: Contracts for difference*.

[111] Ofgem. 2016, *About the RO*.

[112] D. Bunn, T. Yusupov. 2015, *The progressive inefficiency of replacing renewable obligation certificates with contracts-for-differences in the UK electricity market*, Energy Policy, 82, pp. 298-309.

[113] DECC. 2015, *2010 to 2015 government policy: low carbon technologies*.

[114] Ofgem. 2016, *RO closure*.

[115] BEIS. 2015, *Contracts for difference (CFD) allocation round one outcome*.

[116] BEIS. 2017, *Contracts for Difference second allocation round results*.

[117] KPMG. 2016, *Second CfD Allocation Round*, Accessed: April 2017, https://home.kpmg.com/content/dam/kpmg/uk/pdf/2016/11/second-cfd-allocation-round.pdf

[118] The National Grid. 2015, *Provisional Auction Results: T-4 Capacity market auction for 2019/20*

[119] DECC. 2015, *Press release: The first ever Capacity Market auction official results have been released today*.

[120] The National Grid. 2016, *Provisional Auction Results T-4 Capacity market auction 2020/2021*.

[121] Fusion media Ltd. *Carbon Emissions Futures*, Accessed: October 2016, https://uk.investing.com/commodities/carbon-emissions-historical-data

[122] Sandbag. *The UK carbon floor price*, Accessed: August 2015, https://sandbag.org.uk/site_media/pdfs/reports/Sandbag_Carbon_Floor_Price_2013_final.pdf

[123] UK Parliament. 2013, *Briefing paper: Carbon price floor*.

[124] David Hirst. 2018, *Briefing paper (No. 05927): Carbon price floor (CPF) and the price support mechanism.*

[125] BEIS. 2017, *Energy trends: December 2017*.

[126] M.L. Ahlstrom, R.M. Zavadil. 2005, *The Role of Wind Forecasting in Grid Operations & Reliability*, IEEE/PES Transmission & Distribution Conference & Exposition: Asia and Pacific, Dalian, China.

[127] R. Doherty, M. O'Malley. 2005, *A new approach to quantify reserve demand in systems with significant installed wind capacity*, Power Systems, IEEE Transactions on, 20, pp. 587-595.

[128] C. Lowery, M. O'Malley. 2012, *Impact of wind forecast error statistics upon unit commitment*, IEEE Transactions on Sustainable Energy, 3, pp. 760-768.

[129] M.A. Matos, R.J. Bessa. 2011, *Setting the operating reserve using probabilistic wind power forecasts*, IEEE Transactions on Power Systems, 26, pp. 594-603.

[130] H. Bludszuweit, J.A. Domínguez-Navarro. 2011, *A probabilistic method for energy storage sizing based on wind power forecast uncertainty*, IEEE Transactions on Power Systems, 26, pp. 1651-1658.

[131] A. Botterud, Z. Zhou, J. Wang, J. Sumaili, H. Keko, J. Mendes, R.J. Bessa, V. Miranda. 2013, *Demand dispatch and probabilistic wind power forecasting in unit commitment and economic dispatch: A case study of Illinois*, IEEE Transactions on Sustainable Energy, 4, pp. 250-261.

[132] Z. Zhou, A. Botterud, J. Wang, R. Bessa, H. Keko, J. Sumaili, V. Miranda. 2013, *Application of probabilistic wind power forecasting in electricity markets*, Wind Energy, 16, pp. 321-338.

[133] M. Zugno, T. Jónsson, P. Pinson. 2013, *Trading wind energy on the basis of probabilistic forecasts both of wind generation and of market quantities*, Wind Energy, 16, pp. 909-926.

[134] A. Botterud, Z. Zhou, J. Wang, R.J. Bessa, H. Keko, J. Sumaili, V. Miranda. 2012, *Wind power trading under uncertainty in LMP markets*, IEEE Transactions on power systems, 27, pp. 894-903.

[135] P. Pinson, C. Chevallier, G.N. Kariniotakis. 2007, *Trading wind generation from short-term probabilistic forecasts of wind power*, IEEE Transactions on Power Systems, 22, pp. 1148-1156.

[136] V. Robu, R. Kota, G. Chalkiadakis, A. Rogers, N.R. Jennings. 2012, *Cooperative virtual power plant formation using scoring rules*, Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 3.

[137] C.E. Rasmussen, C.K.I. Williams. 2006, *Gaussian Processes for Machine Learning*, The MIT press, ISBN: 0-262-18253-X

[138] J. Melo. 2012, *Gaussian processes for regression: a tutorial*, Accessed: Oct 2017,

[139] T. Chen, J. Morris, E. Martin. 2007, *Gaussian process regression for multivariate spectroscopic calibration*, Chemometrics and Intelligent Laboratory Systems, 87, pp. 59-71.

[140] D. Nguyen-Tuong, J.R. Peters, M. Seeger. 2009, *Local Gaussian process regression for real time online model learning*, Advances in Neural Information Processing Systems.

[141] H. He, W.C. Siu. 2011, *Single image super-resolution using Gaussian process regression*, IEEE Conference: Computer Vision and Pattern Recognition (CVPR).

[142] S.M. Weekes. 2014, *Small-scale wind energy: methods for wind resource assessment*, University of Leeds, Doctoral Degree.

[143] C. Zhang, H. Wei, X. Zhao, T. Liu, K. Zhang. 2016, *A Gaussian process regression based hybrid approach for short-term wind speed prediction*, Energy Conversion and Management, 126, pp. 1084-1092.

[144] J. Hu, J. Wang. 2015, *Short-term wind speed prediction using empirical wavelet transform and Gaussian process regression*, Energy, 93, pp. 1456-1466.

[145] D. Duvenaud. 2014, *Automatic model construction with Gaussian processes*, University of Cambridge, Doctoral Dissertation.

[146] C.E. Rasmussen, H. Nickisch. 2010, *Gaussian processes for machine learning (GPML) toolbox*, The Journal of Machine Learning Research, 11, pp. 3011-3015.

[147] H. Bijl. 2016, *Gaussian process regression techniques with applications to wind turbines*, Delft University of Technology, Doctoral degree.

[148] J. Nocedal, S.J. Wright. 2006, *Conjugate gradient methods*, Numerical optimization, pp. 101-134.

[149] Z. Chen, B. Wang. 2016, *How priors of initial hyperparameters affect Gaussian process regression models*, arXiv preprint arXiv:1605.07906, pp.

[150] Met Office. *Research news: The UKV model - Kilometre-scale forecasting over the UK with the Unified Model*, Accessed: March 2015, http://www.metoffice.gov.uk/research/news/ukv

[151] Met Office. *3 hourly weather forecast and observational data - UK locations*, Accessed: September 2015, http://data.gov.uk

[152] Met Office. *Met Office Integrated Data Archive (MIDAS) Land and Marine Surface Stations Data (1853-current)*, Accessed: March 2015, http://catalogue.ceda.ac.uk/

[153] I. Sandu, A. Beljaars, P. Bechtold, T. Mauritsen, G. Balsamo. 2013, *Why is it so difficult to represent stably stratified conditions in numerical weather prediction (NWP) models?*, Journal of Advances in Modeling Earth Systems, 5, pp. 117-133.

[154] A. Holtslag, G. Svensson, P. Baas, S. Basu, B. Beare, A. Beljaars, F. Bosveld, J. Cuxart, J. Lindvall, G. Steeneveld. 2013, *Stable atmospheric boundary layers and diurnal cycles: challenges for weather and climate models*, Bulletin of the American Meteorological Society, 94, pp. 1691-1706.

[155] U. Focken, M. Lange, K. Mönnich, H.-P. Waldl, H.G. Beyer, A. Luig. 2002, *Short-term prediction of the aggregated power output of wind farms—a statistical analysis of the reduction of the prediction error by spatial smoothing effects*, Journal of Wind Engineering and Industrial Aerodynamics, 90, pp. 231-246.

[156] A. Peña, A.N. Hahmann. 2012, *Atmospheric stability and turbulence fluxes at Horns Rev—an intercomparison of sonic, bulk and WRF model data*, Wind Energy, 15, pp. 717-731.

[157] L. Jensen. 2007, *Array efficiency at Horns Rev and the effect of atmospheric stability*, Dong Energy, Fredericia, Denmark, pp.

[158] K. Ashrafi, G.A. Hoshyaripour. 2010, *A model to determine atmospheric stability and its correlation with CO concentration*, International Journal of Civil and Environmental Engineering, 2, pp.

[159] D.B. Turner. 1964, *A diffusion model for an urban area*, Journal of Applied Meteorology, 3, pp. 83-91.

[160] J. Hu, J. Wang. 2015, *Short-term wind speed prediction using empirical wavelet transform and Gaussian process regression*, Energy, 93, pp. 1456-1466.

[161] ELEXON Ltd. *Balancing Mechanism Reporting System (BRMS)*, Accessed: 2017, http://www.bmreports.com.

[162] T. Foken. 2006, *50 Years of the Monin–Obukhov Similarity Theory*, Boundary-Layer Meteorology, 119, pp. 431-447.

[163] S.-E. Gryning, E. Batchvarova, B. Brümmer, H. Jørgensen, S. Larsen. 2007, *On the extension of the wind profile over homogeneous terrain beyond the surface boundary layer*, Boundary-Layer Meteorology, 124, pp. 251-268.

[164] D. Golder. 1972, *Relations among stability parameters in the surface layer*, Boundary-Layer Meteorology, 3, pp. 47-58.

[165] R.J. Barthelmie, J. Badger, S.C. Pryor, C.B. Hasager, M.B. Christiansen, B.H. Jørgensen. 2007, *Offshore Coastal Wind Speed Gradients: Issues for the Design and Development of Large Offshore Windfarms*, Wind Engineering, 31, pp. 369-382.

[166] J. Dowell, G. Hawker, K. Bell, S. Gill. 2016, *A Review of probabilistic methods for defining reserve requirements*, Power and Energy Society General Meeting (PESGM), 2016.

[167] R.J. Bessa, M.A. Matos. 2010, *Comparison of probabilistic and deterministic approaches for setting operating reserve in systems with high penetration of wind power*, pp.

[168] B.S. Jensen, J.B. Nielsen, J. Larsen. 2013, *Bounded Gaussian process regression*, IEEE International Workshop on Machine Learning for Signal Processing (MLSP).

[169] E. Snelson, C.E. Rasmussen, Z. Ghahramani. 2004, *Warped Gaussian processes*, Advances in neural information processing systems, 16, pp. 337-344.

[170] C.E. Rasmussen, H. Nickisch. 2016, *The GPML Toolbox version 4.0*, Accessed: Nov 2016, http://www.gaussianprocess.org/gpml/code/matlab/doc/manual.pdf

[171] A. Jordan. *Closed form expressions for the continuous ranked probability score*, Accessed: Mar 2017, https://github.com/FK83/scoringRules/blob/master/crps.pdf

[172] A. Botterud, J. Wang, R. Bessa, H. Keko, V. Miranda. 2010, *Risk management and optimal bidding for a wind power producer*, Power and Energy Society General Meeting, 2010 IEEE.

[173] E. Ibanez, M. Milligan. 2012, *A probabilistic approach to quantifying the contribution of variable generation and transmission to system reliability*, National Renewable Energy Lab Conference Paper. NREL/CP-5500-56219.

[174] DECC. 2013, *EMR consultation. Annex C: Reliability standard methodology*.

[175] ELEXON Ltd. *Loss of load probability calculation statement*, Accessed: April 2016, https://www.elexon.co.uk/wp-content/uploads/2015/08/Loss-of-Load-Probability-Calculation-Statement-V0-1-for-Consultation.pdf

[176] R.J. Bessa, M.A. Matos, I.C. Costa, L. Bremermann, I.G. Franchin, R. Pestana, N. Machado, H.P. Waldl, C. Wichmann. 2012, *Reserve Setting and Steady-State Security Assessment Using Wind Power Uncertainty Forecast: A Case Study*, IEEE Transactions on Sustainable Energy, 3, pp. 827-836.

[177] ELEXON Ltd. *P305 Historic Analysis*, Accessed: April 2016, https://www.elexon.co.uk/wp-content/uploads/2014/05/P305_AC_X-Historic-Analysis-v1.0.pdf