



Norwegian University of  
Science and Technology

# Prediction of Wind Power Production Using Beta Transformed Linear Pool

**Andreas Malmgård**

Master of Science in Physics and Mathematics

Submission date: June 2017

Supervisor: Ingelin Steinsland, IMF

Co-supervisor: Joakim Blakstad, Markedskraft ASA

Norwegian University of Science and Technology  
Department of Mathematical Sciences



## **Preface**

This is a master thesis in statistics at NTNU as part of the study program industrial mathematics. The thesis was carried out during the spring semester of 2017 and performed in cooperation with Markedskraft ASA, who provided data and support during the time of writing. The problem description was also carried out in cooperation with Markedskraft ASA.

The thesis assumes the reader has a technological background and is familiar with terms and common methods within statistics.



## **Acknowledgments**

First, I would like to thank Joakim R. Blakstad and Marianne Rypestøl from Markedsraft ASA for being engaged in the planning of this thesis and for being helpful during the project period. I would also like to thank my friends Jens Bruin Ødegaard and Steinar Halsne for helping me read through and edit the thesis. A special thanks goes to my supervisor Professor Ingelin Steinsland for guiding me through this last year of my studies. Our weekly meetings were very helpful and a lot of great ideas have been shared with me and shaped this thesis to a large extent. Finally, I would like to thank Marit Ramsleie Ramstad for always supporting me and for your company at late night study hours. The last few weeks especially, would be a lot harder without you.



## Abstract

In the power trading market, transmission system operators and other actors buy and sell power related to future production. Power production from wind- and solar farms is affected by rapid weather changes, and producers of this power often have to trade in reaction to the changes. Ongoing actual power production data for wind- and solar farms is published by the transmission system operators. These numbers indicate what volumes these actors have to trade to keep up with the weather changes, and are therefore one of the largest cost drivers in the market. Therefore, the ability to forecast power production is highly relevant in the power trading industry.

In this master thesis we do a case study with focus on wind energy, and the main research task is to predict wind power production. We introduce three models named CCPR, UCPR and CPR-LP, where all are based on a new methodology. The methodology starts out with one or two initial forecasts, in the form of cumulative density functions. The CCPR and UCPR use one initial forecast, and the method proceeds by transforming the initial forecast through a beta transformation function, returning a calibrated final forecast. The CPR-LP uses two initial forecasts, where the methodology beta transforms a weighted sum of these. The parameters which define the beta transformation function are modelled as functions of deterministic forecasts related to the wind power production. We divide our test results into groups, based on these deterministic forecasts. UCPR is performing very well compared to the other models for large deterministic forecasts, and CCPR is performing well for small deterministic forecasts. The CPR-LP model on the other hand is preferable when considering all groups as a whole.



## Sammendrag

Systemoperatører og andre aktører i kraftmarkedet kjøper og selger strøm for fremtidig strømproduksjon. Strømproduksjon fra vind- og solenergi blir påvirket av endringer i været, som gjør at produsenter av slik strøm i tillegg må handle på kraftmarkedet for å gjøre opp for disse endringene. Løpende strømproduksjonstall blir publisert av systemoperatørene, og indikerer hvor mye vind- og solprodusentene må handle på kraftmarkedet for å gjøre opp for været. Derfor er disse tallene en av de største prisdriverne i kraftmarkedet. Varsel som estimerer disse strømproduksjonstallene er derfor svært relevant i denne industrien.

I denne masteroppgaven skal vi utføre et case-studie med fokus på vindenergi, der hovedoppgaven er å varsle strømproduksjonen fra vind. Vi introduserer de tre modellene CCPR, UCPR og CPR-LP, som alle er basert på en ny metodologi for å danne probabilistiske varsel. Denne metodologien starter med en eller to initialvarsel i form av kumulative distribusjonsfunksjoner. CCPR og UCPR bruker kun et initialvarsel, og metoden bruker en betafordeling for å transformere dette til et kalibrert endelig varsel. CPR-LP bruker to initialvarsel, hvor betafordelingen transformerer en vektet sum av disse. Parameterne som definerer betafordelingen modelleres som en funksjon av deterministiske varsel, hvor disse inneholder informasjon om strømproduksjonen fra vind. Vi deler testresultatene inn i grupper basert på disse deterministiske varslene. UCPR har best resultater sammenlignet med de andre modellene for høye deterministiske varsel, og CCPR har best resultater for små deterministiske varsel. CPR-LP er den foretrekkende modellen, og har best resultater når vi ser på alle gruppene som en helhet



# Contents

|  |           |
|--|-----------|
| Preface . . . . .  | i         |
| Acknowledgments . . . . .  | iii       |
| Summary . . . . .  | v         |
| Sammendrag . . . . .   | vii       |
| <b>1 Introduction</b>  | <b>1</b>  |
| <b>2 Data</b>  | <b>5</b>  |
| 2.1 Wind Farm Actuals and EC-forecasts . . . . .                   | 5         |
| 2.2 Explorative Analysis . . . . .                                 | 7         |
| <b>3 Background</b>  | <b>12</b> |
| 3.1 Beta Distribution . . . . .                                    | 12        |
| 3.2 Logit Link . . . . .   | 13        |
| 3.3 Beta Transformed Linear Pool (BLP) . . . . .                   | 16        |
| 3.4 Evaluation Methods . . . . .                                   | 17        |
| 3.4.1 RMSE . . . . .   | 18        |
| 3.4.2 PIT-diagram . . . . .  | 18        |
| 3.4.3 CRPS . . . . .   | 20        |
| 3.4.4 Cross Validation . . . . .                                   | 22        |
| <b>4 Methodological Concept</b>                                    | <b>23</b> |
| 4.1 A Toy Example . . . . .  | 23        |
| 4.2 CCPR (Climatology Cumulative Probability Regression) . . . . . | 26        |
| 4.3 UCPR (Uniform Cumulative Probability Regression) . . . . .     | 31        |

|          |  |           |
|----------|--|-----------|
| 4.4      | CPR-LP (Cumulative Probability Regression with Linear Pooling)           | 33        |
| 4.5      | Modelling the Beta Parameters $\mu$ , $\nu$ and the Weight Parameter $w$ | 34        |
| <b>5</b> | <b>Case Study: Models and Evaluation Scheme</b>                          | <b>36</b> |
| 5.1      | Notation: EC- and Persistence Forecast and Hourly Changes                | 36        |
| 5.2      | Model 1: CCPR  | 37        |
| 5.3      | Model 2: UCPR  | 38        |
| 5.4      | Model 3: CPR-LP  | 39        |
| 5.5      | Model 4: ARn-SFR   | 40        |
| 5.6      | Toy Models: Interpretation of Model Coefficients Using Synthetic Data    | 41        |
| 5.6.1    | CCPR coefficients: Synthetic Forecasts                                   | 41        |
| 5.6.2    | UCPR coefficients: Synthetic Forecasts                                   | 43        |
| 5.6.3    | CPR-LP coefficients: Synthetic Forecasts                                 | 45        |
| 5.7      | Inference, Evaluation and Implementation                                 | 47        |
| <b>6</b> | <b>Results</b>   | <b>49</b> |
| 6.1      | Model Coefficients   | 49        |
| 6.1.1    | CCPR coefficients: Actual Forecasts                                      | 49        |
| 6.1.2    | UCPR coefficients: Actual Forecasts                                      | 51        |
| 6.1.3    | CPR-LP coefficients: Actual Forecasts                                    | 51        |
| 6.2      | Model Evaluation   | 54        |
| 6.2.1    | Probabilistic Forecasts  | 54        |
| 6.2.2    | Sharpness  | 56        |
| 6.2.3    | Calibration  | 58        |
| <b>7</b> | <b>Discussion and Conclusion</b>   | <b>63</b> |
| <b>A</b> | <b>Actual Production Sources</b>   | <b>67</b> |
| <b>B</b> | <b>Interpretation of <math>r(\cdot)</math></b>                           | <b>68</b> |

**C Additional Figures** **71**

**Bibliography** **78**



# Chapter 1

## Introduction

Elbas<sup>1</sup> and EPEX<sup>2</sup> are the two largest exchanges for power trading in Northern Europe. They conduct different types of auctions, two of which are called *spot auction* and *intraday auction*. A spot auction happens the day before power delivery, whereas intraday trading is possible down to one hour before delivery. In the time between spot auction and delivery, producers and consumers might fall out of balance. Imbalance means that the final consumption/production from these actors becomes different compared to what they bought and sold the day before. These actors have the opportunity to buy or sell power on the intraday market to correct their balance (NordPool).

Producers of renewable energy like wind- and solarenergy, can not reliably control their production and are very vulnerable to weather changes. A high amount of their revenue might get lost between spot auction and delivery. Intraday is therefore especially important for these producers to restore this revenue. Transmission system operators for each country publish ongoing actual power production data (hereafter referred to as actual production or simply actuals). This includes actuals for wind- and solar production. These numbers are one of the largest cost drivers in the market, because they indicate how much the producers of wind- and solar power have to trade. It is therefore very useful to be able to predict this production.

Until the early 1990s, deterministic forecasts were the most common type of forecast. Weather forecasts, for example, were a deterministic endeavour where one weather prediction was gen-

---

<sup>1</sup>Elbas is a system for intraday auction managed and developed by Nord Pool: <http://www.nordpoolspot.com/TAS/Intraday-market-Elbas/elbas-4/>

<sup>2</sup>EPEX: [http://www.epexspot.com/en/company-info/about\\_epex\\_spot](http://www.epexspot.com/en/company-info/about_epex_spot)

erated (Gneiting and Raftery, 2005). However, all forecasts are associated with some sort of uncertainty and probabilistic forecasts are often preferred over deterministic forecasts to better express this uncertainty. The probabilistic forecast takes the form of a pdf (probability density function) or a cdf (cumulative density function) (Gneiting et al., 2007). The way a probabilistic forecast represent uncertainty allows more nuanced decision making. As a result, probabilistic forecasts have seen increased impact in many applications, such as economics, meteorology and climatology (Gneiting and Ranjan, 2013).

The aim of a probabilistic forecast is to maximise the sharpness subject to calibration. Calibration refers to the statistical consistency between the probabilistic forecast and the observations, whereas sharpness refers to the concentration of the predictive distributions (Gneiting et al., 2007). If we declare an interval or event to have probability  $p$ , the forecast is calibrated if the event happens a proportion  $p$  of the time on average (Raftery et al., 2005). Probability integral transform (PIT) diagrams are commonly used to check for calibration. Clearly, the sharper the forecast the better, as long as it is still calibrated. The width of  $p\%$  prediction intervals measures the sharpness of a probabilistic forecast. The continuous ranked probability score (CRPS) combines both sharpness and calibration to evaluate the performance of a forecast. The score is minimized when the probabilistic forecast is identical to the distribution of the process we want to forecast, and is thus a proper scoring rule (Gneiting et al., 2007).

The dominant approach to probabilistic forecasting in areas such as weather forecasting, has been to use ensembles of deterministic forecasts in which a model is run several times with different initial conditions. A post processing approach could then be used to form a probabilistic forecast from this ensemble. Bayesian model averaging (BMA) is one such post processing technique (Raftery et al., 2005). In recent years, probabilistic forecasts in the form of predictive probability distributions have become more prevalent in various fields, such as economics, finance and meteorology. Research has therefore shifted towards the construction of post processing methods, used to combine several probabilistic forecasts (Bassetti et al., 2015). The prevalent method has been the linear pool, which is a weighted sum of the available probabilistic forecasts. Hora (2004), Ranjan and Gneiting (2010) and Gneiting and Ranjan (2013) have all pointed out shortcomings with the linear pool, all revolving around lack of calibration. Gneiting and Ranjan (2013) introduced the beta transformed linear pool, which composites the linear

pool with a beta transformation to improve calibration. This approach has been generalized further by [Bassetti et al. \(2015\)](#) who introduce a combination of several weighted beta transformed linear pools.

The main research task of this master thesis is to forecast the actual wind power production for several countries in Northern Europe. We introduce a new methodology, inspired by the forecasting method from [Borhaug \(2014\)](#), which in turn is based on the beta transformed linear pool presented by [Gneiting and Ranjan \(2013\)](#). The CCPR (Climatology Cumulative Probability Regression) model from [Borhaug \(2014\)](#) has been reimplemented and used for our case study. This method takes advantage of historical data, i.e. the climatology, and different deterministic forecasts to create a probabilistic forecast. Based on the challenges experienced with the CCPR model for our case study, we introduce two new models named UCPR (Uniform Cumulative Probability Regression) and CPR-LP (Climatology Cumulative Probability Regression with Linear Pooling). These are based on the same methodological concept as the CCPR, but are tailored to increase the forecasting performance when used for our case study. However, the method behind the two new models is presumably applicable to any set of historical data and deterministic forecasts, and opens up an interesting path for further research.

This master thesis is also an extension to [Malmgård \(2016\)](#)<sup>3</sup>, which was a specialisation project with the exact same data and research task. The forecasting models tested by [Malmgård \(2016\)](#) were based on the well known method of linear regression, accompanied by autoregressive models originating from the temporal dependencies. These models had a substantial disadvantage because the probabilistic forecasts did not always match with the nature of the data, i.e. the forecasts did not always make sense physically. The actuals are never below zero, and there is also a maximum limit for actual production. Because the resulting probabilistic forecasts were normal distributions, you could end up with a forecast having considerable density mass for negative production, or above the maximum limit for actual production. The best performing model tested by [Malmgård \(2016\)](#) was the ARn-SFR-model. In addition to not being physically valid, it suffered from consistent underdispersion, i.e. giving too narrow probabilistic forecasts. The main motivation behind our choice of forecasting methodology is thus to obtain physically valid and calibrated forecasts.

---

<sup>3</sup>Please email [andreas@malmgard.com](mailto:andreas@malmgard.com) for the full paper.

The rest of the thesis is organized as follows. Chapter 2 gives an introduction to our case study and a brief explanatory analysis, illustrating some important features of the data. Chapter 3 states the background theory required to develop our forecasting methodology, as well as the methods used to evaluate the forecasting performance. Chapter 4 describes the concept behind our forecasting methodology. First with a toy example to illustrate desired properties, before formally defining the three forecasting models CCPR, UCPR and CPR-LP. Chapter 5 establishes some useful notation and defines the parameter models for each forecasting model used in our case study. An interpretation of the coefficients used to model these parameters is also included. The results of the case study are given in Chapter 6, and Chapter 7 concludes the thesis with a summary and suggestions for further work.

# Chapter 2

## Data

In this chapter we start by introducing the data for our case study in Section 2.1. Section 2.2 includes a brief explanatory analysis, illustrating some important features of the data.

### 2.1 Wind Farm Actuals and EC-forecasts

The observations in our case study are time series of power production from wind plants in megawatts. The resolution is of one value each hour, where this value is the mean production during that hour. Each time series includes the production for one whole bidding area. A bidding area is a geographical area within which market participants are able to exchange power without capacity allocation (Ofgem, 2014). A total of ten bidding areas are included, and a map of these can be seen in Figure 2.1. Further, we also have available time series of forecasts. Both actuals and forecasts are detailed below:

**Actuals** were mentioned in Chapter 1 and are actual wind power production in megawatts. These are in reality estimates provided by the transmission system operators or other external actors. Appendix A includes a list of where actuals for each bidding area are published. The actuals are gathered from these sources and provided as a whole by Markedskraft ASA. An example of the actual production for all areas can be seen in Table 2.1, and plotted in Figure 2.2a.

**Forecasts** (hereafter referred to as EC-forecasts) are forecasts of the actuals given in megawatts. They



Figure 2.1: The map marks the ten different bidding areas considered. Each bidding area has its own color. "RWE/Amp." is divided into two separate geographical areas that comprise a single bidding area.

are calculated based on a wind speed forecast, and provided by Markedskraft ASA two times daily. EC00-forecast starts calculations at midnight, and the result is issued at 08:00. EC12-forecast starts at midday, and the result is issued at 20:00. Because of this delay, Kalman filtering is used to minimize the errors of the EC-forecast at issue time. As an example, when EC00 is available at 08:00, there are already observed actuals from 00:00 to 08:00 which is used in this filtering. The EC-forecast used in this project is a combination of EC00 and EC12. For any given day, the EC-forecast is equal to EC00 from 08:00 to 19:00, and equal to EC12 from 20:00 to 07:00 the next day. An example of the EC-forecast for all areas can be seen in Table 2.2, and plotted in Figure 2.2b.

We also want to introduce the notion of *lead time* of the EC-forecast. This is how far ahead in time we forecast, i.e. the time between the forecast valid time and its issue time. Consider again the EC00-forecast as an example (which is issued at 08:00). The EC00-forecast for 15:00 has a lead time of 7 hours. At the end of Section 2.2, we see how the error

in EC-forecasts increase with increasing lead time.

The length of each provided data series vary, but all of them reach back to at least 4th of August 2015, and last until 6th of June 2016. This time period is used for this project. Table 2.3 displays some fundamental information of both the actuals and EC-forecasts.

## 2.2 Explorative Analysis

We start this section by investigating the correlation between the actual production, EC-forecast and EC-error within each bidding area, where EC-error is defined as actuals minus EC-forecast. Figure 2.3 displays the results. The correlation between actual production and EC-forecast is clear for all areas, whereas the two other combinations does not have a clear common pattern.

The autocorrelation of the actual production is quite high for small lags as shown in Figure 2.4a. The only area standing out is Austria. In contrast to the other bidding areas, Austria is mostly covered by mountains as it is located in the Alps. This creates weather regimes that differ from what is found in the other low land bidding areas, and may explain the low autocorrelations. The autocorrelation curves for the EC-errors in Figure 2.4b are all lower than the curves for actual production as to be expected, but they have a rather large spread among the bidding areas. The EC-errors should ideally have low autocorrelation.

Next we investigate the EC-error sample variance. As mentioned at the start of the chapter, two forecasts are issued each day, at 08:00 and at 20:00. Figure 2.5a displays the relative error variance at each hour of the day. Looking at the mean curve, the variance is clearly increasing in the hours from 08:00 to 16:00, and gets lower towards 20:00. This behaviour of increasing variance with increasing lead time is expected, since an increase in lead time means larger time frame between forecast valid time and its issue time. In addition, Figure 2.5b shows that the relative variance of the actuals have similar shape in the hours between 08:00 and 20:00. This is not a coincidence as larger variance in the actuals makes it harder to forecast, which is another reason for the large EC-error variance in this time frame.

At 20:00, the EC-error variance drops because a new forecast with updated initial conditions gets issued. The increase of variance is not as prominent after 20:00, it is instead rather low during the whole period from 20:00 to 07:00, even though the lead time increases. The explanation

| Date          | AT      | BE     | DK1    | DK2    | DE_ENB | DE_EON  | DE_RWE | DE_VAT  | FR      | NL      |
|---------------|---------|--------|--------|--------|--------|---------|--------|---------|---------|---------|
| 2015-08-05 00 | 503.00  | 555.17 | 374.75 | 30.25  | 36.50  | 2719.00 | 444.50 | 3457.00 | 638.25  | 1040.00 |
| 2015-08-05 01 | 502.75  | 585.75 | 433.50 | 11.25  | 29.50  | 2478.00 | 392.75 | 2689.75 | 589.00  | 949.50  |
| 2015-08-05 02 | 431.00  | 415.08 | 449.25 | 16.50  | 13.75  | 2137.00 | 333.75 | 2149.25 | 639.25  | 792.25  |
| 2015-08-05 03 | 269.75  | 297.09 | 464.00 | 144.75 | 9.25   | 1721.00 | 280.25 | 2072.50 | 742.75  | 702.75  |
| 2015-08-05 04 | 360.25  | 317.14 | 519.00 | 266.75 | 5.00   | 1666.50 | 267.00 | 2009.25 | 789.75  | 637.00  |
| 2015-08-05 05 | 427.75  | 299.65 | 584.50 | 233.50 | 12.50  | 2010.75 | 284.25 | 1646.75 | 935.25  | 605.75  |
| 2015-08-05 06 | 579.00  | 356.79 | 547.75 | 141.25 | 25.75  | 2157.75 | 365.00 | 1399.50 | 1134.25 | 723.00  |
| 2015-08-05 07 | 751.00  | 313.84 | 580.75 | 89.50  | 34.50  | 2047.25 | 477.50 | 1244.50 | 1269.75 | 645.25  |
| 2015-08-05 08 | 727.75  | 242.10 | 472.75 | 92.25  | 36.75  | 1868.25 | 475.75 | 917.25  | 1239.25 | 490.50  |
| 2015-08-05 09 | 914.25  | 201.99 | 474.75 | 49.50  | 26.50  | 1485.50 | 316.25 | 638.25  | 1102.25 | 292.25  |
| 2015-08-05 10 | 1016.50 | 98.45  | 513.50 | 40.50  | 16.50  | 1483.50 | 229.00 | 377.50  | 703.50  | 214.25  |
| 2015-08-05 11 | 288.25  | 67.28  | 514.75 | 43.00  | 13.25  | 1746.75 | 275.25 | 275.50  | 693.00  | 219.00  |
| 2015-08-05 12 | 326.50  | 80.15  | 545.00 | 31.75  | 10.75  | 1861.75 | 277.00 | 247.00  | 920.00  | 154.00  |
| 2015-08-05 13 | 273.50  | 98.76  | 559.25 | 32.75  | 11.50  | 1686.50 | 261.75 | 228.75  | 1050.00 | 114.25  |
| 2015-08-05 14 | 217.00  | 159.29 | 566.50 | 32.00  | 12.50  | 1225.75 | 317.75 | 243.25  | 1044.25 | 110.00  |

Table 2.1: Actual wind production in all ten bidding areas for 15 consecutive hours, starting from 5th of August 2015 at midnight.

| Forecast | Date          | AT     | BE     | DK1    | DK2    | DE_ENB | DE_EON  | DE_RWE | DE_VAT  | FR      | NL     |
|----------|---------------|--------|--------|--------|--------|--------|---------|--------|---------|---------|--------|
| EC12     | 2015-08-05 00 | 303.67 | 358.68 | 445.15 | 205.58 | 13.41  | 2650.55 | 456.73 | 1920.95 | 1171.87 | 486.44 |
| EC12     | 2015-08-05 01 | 385.80 | 313.47 | 503.05 | 192.67 | 11.63  | 2470.65 | 459.81 | 1729.43 | 1325.38 | 478.41 |
| EC12     | 2015-08-05 02 | 574.32 | 286.94 | 522.22 | 178.50 | 14.37  | 2240.40 | 478.57 | 1575.11 | 1413.38 | 457.85 |
| EC12     | 2015-08-05 03 | 706.54 | 295.06 | 520.34 | 163.14 | 15.15  | 2011.80 | 492.78 | 1414.84 | 1433.17 | 462.73 |
| EC12     | 2015-08-05 04 | 690.38 | 275.65 | 501.43 | 139.20 | 12.20  | 1734.76 | 482.47 | 1289.25 | 1378.30 | 476.42 |
| EC12     | 2015-08-05 05 | 620.85 | 265.90 | 462.03 | 114.63 | 5.90   | 1459.22 | 479.51 | 1306.87 | 1216.89 | 521.29 |
| EC12     | 2015-08-05 06 | 583.36 | 225.95 | 567.44 | 111.90 | 8.42   | 1759.44 | 436.08 | 618.98  | 967.52  | 497.46 |
| EC12     | 2015-08-05 07 | 493.79 | 186.40 | 529.23 | 90.55  | 11.26  | 1582.65 | 343.64 | 302.56  | 739.78  | 413.33 |
| EC00     | 2015-08-05 08 | 388.75 | 140.27 | 456.41 | 65.42  | 12.19  | 1345.81 | 303.31 | 190.51  | 696.72  | 313.64 |
| EC00     | 2015-08-05 09 | 322.31 | 121.24 | 443.98 | 47.21  | 13.65  | 1187.00 | 282.55 | 155.64  | 734.51  | 275.51 |
| EC00     | 2015-08-05 10 | 225.48 | 145.71 | 494.98 | 35.17  | 12.93  | 1173.40 | 290.04 | 161.18  | 787.31  | 334.07 |
| EC00     | 2015-08-05 11 | 148.96 | 173.90 | 587.48 | 25.97  | 12.10  | 1099.90 | 307.01 | 195.88  | 865.58  | 379.29 |
| EC00     | 2015-08-05 12 | 98.32  | 189.32 | 641.79 | 18.88  | 11.53  | 870.83  | 311.81 | 182.18  | 972.79  | 440.69 |
| EC00     | 2015-08-05 13 | 71.17  | 185.63 | 597.57 | 14.04  | 10.29  | 698.41  | 307.08 | 153.25  | 1053.80 | 487.62 |
| EC00     | 2015-08-05 14 | 45.68  | 187.61 | 553.04 | 10.36  | 9.39   | 532.28  | 288.33 | 103.87  | 1117.19 | 541.25 |

Table 2.2: EC-forecast of wind production in all ten bidding areas for 15 consecutive hours, starting from 5th of August 2015 at midnight.

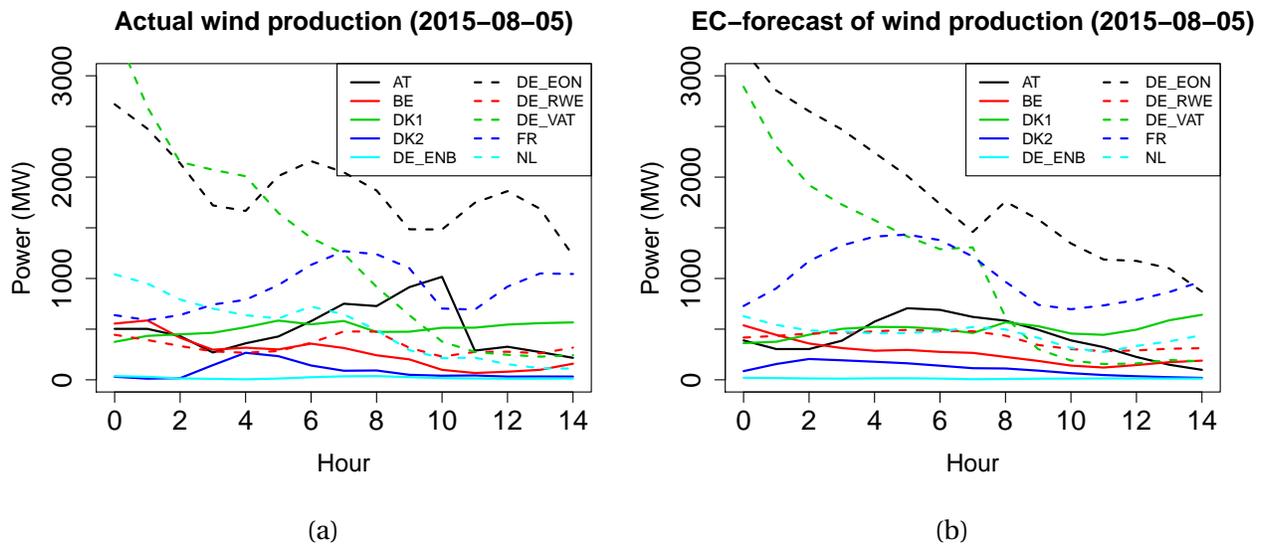


Figure 2.2: The left figure is a plot of the data in Table 2.1 and the right figure is a plot of the data in Table 2.2.

| Area   | StartDates    | Actuals Mean | Forecast Mean | Actual Std | Forecast Std |
|--------|---------------|--------------|---------------|------------|--------------|
| AT     | 2011-01-01 00 | 362.41884    | 342.80874     | 376.53567  | 336.88349    |
| BE     | 2012-12-01 00 | 480.69662    | 685.97492     | 419.51921  | 596.60628    |
| DK1    | 2010-07-08 09 | 978.29551    | 950.81728     | 792.50344  | 788.84747    |
| DK2    | 2010-07-08 09 | 301.93904    | 319.05009     | 261.27683  | 265.07501    |
| DE_ENB | 2010-07-08 09 | 73.74805     | 61.24969      | 97.28731   | 71.42312     |
| DE_EON | 2010-07-08 09 | 2801.89253   | 2834.50144    | 2591.63157 | 2595.89688   |
| DE_RWE | 2010-07-08 09 | 989.52890    | 1010.40328    | 1044.26718 | 940.32963    |
| DE_VAT | 2010-07-08 09 | 3352.37127   | 3163.90813    | 2865.82343 | 2859.58292   |
| FR     | 2014-01-01 00 | 2188.25811   | 2320.98328    | 1604.12345 | 1563.28691   |
| NL     | 2015-08-04 02 | 940.24968    | 920.26733     | 765.78177  | 774.87623    |

Table 2.3: Fundamental information about the data. Including start dates of actuals and EC-forecasts, sample means and sample standard deviations.

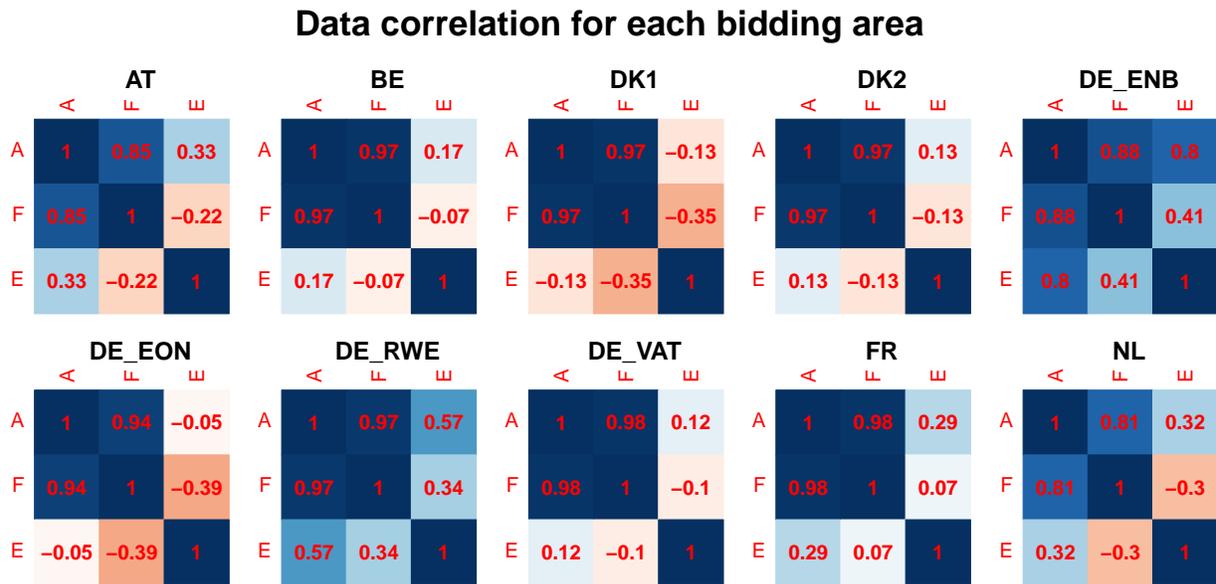


Figure 2.3: Each matrix displays the correlation between actual production(A), EC-forecast(F) and EC-error(E) for each bidding area. The correlation coefficients ranging from  $-1$  to  $1$  are also represented by colors ranging from dark red to dark blue respectively.

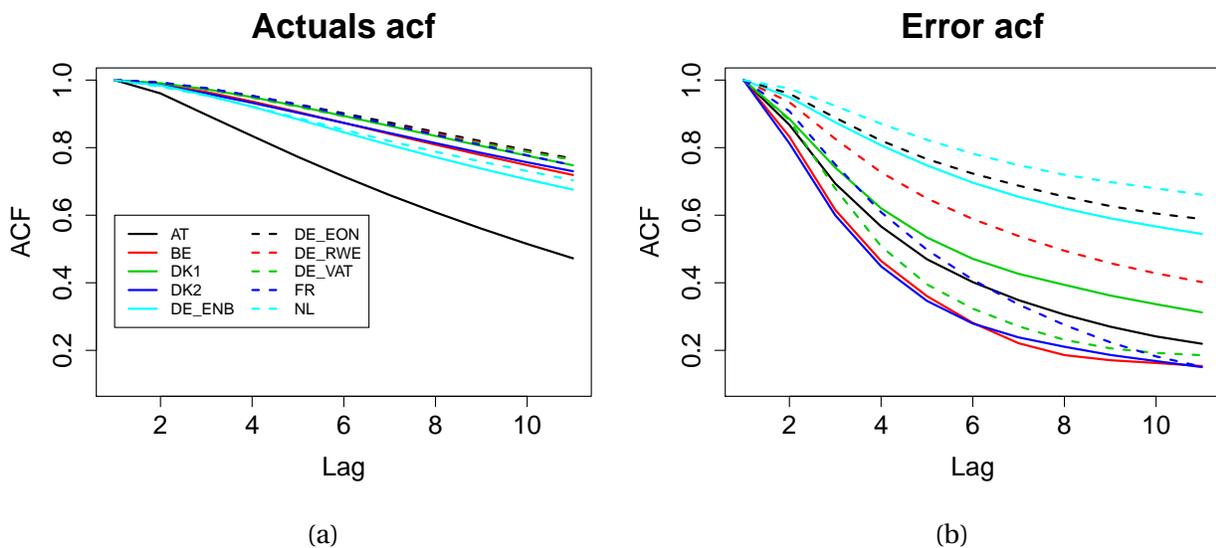


Figure 2.4: Autocorrelation plots for (a) actual production and (b) EC-error for each bidding area. The legend in (a) is also valid for (b).

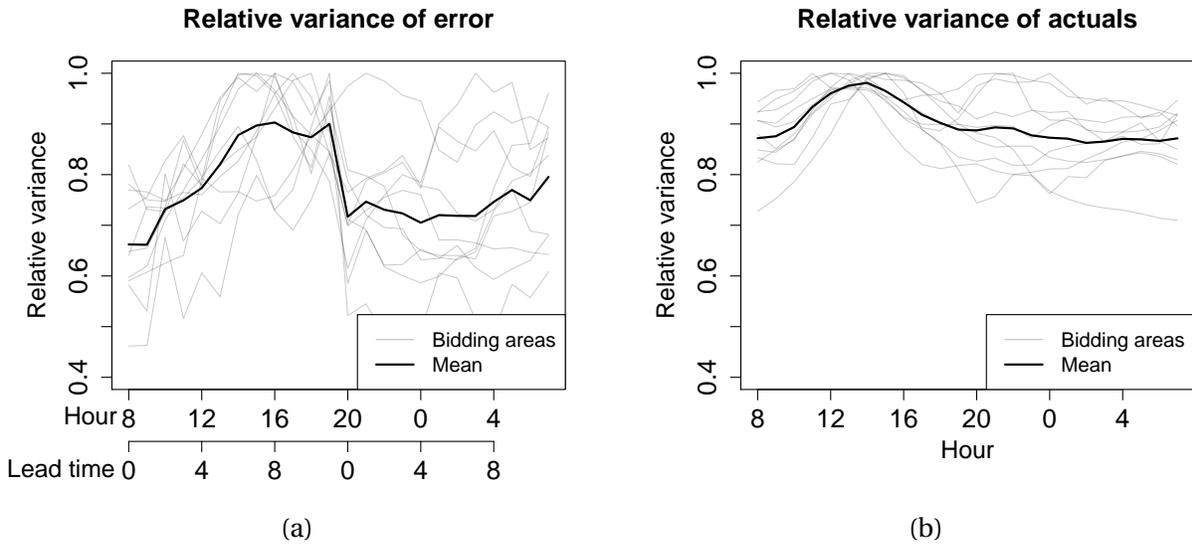


Figure 2.5: Relative variance of (a) the EC-error and (b) actuals at different lead times/hours of the day. The grey curves are the variances for each of the ten bidding areas, while the black curve is the mean of these.

is related to the time varying weather conditions. From Figure 2.5b we see that the actuals have lower variance during night time. This is because the weather is more stable at night with lower wind speeds in general.

As a conclusion, Figure 2.3 showed that the EC-forecast is highly correlated with the actuals, and is going to be an important part of our forecasting models. In addition, the actuals themselves are highly autocorrelated (Figure 2.4a), and are also going to be an important part of our models, especially for shorter lead times. The relative variance of the EC-errors in Figure 2.5a are increasing with increasing lead time. The EC-errors also seem to increase with increasing variance of the actuals, plotted in Figure 2.5b.

# Chapter 3

## Background

In this chapter we introduce the background theory required to develop our forecasting methodology in Chapter 4. We also present the methods used to evaluate the methodology in Chapter 6. Section 3.1 provides a summary of the beta distribution and some of its properties, including several plots to demonstrate the wide variety of shapes the distribution can take. The beta distribution is a key tool in our forecasting methodology where it is used to transform one probabilistic forecast into another. Section 3.2 introduces the logit link function, which maps  $(0,1)$  onto the entire real line. We use the inverse logit link as a tool to model the mean of the beta distribution, which is on the interval  $(0,1)$ . In Section 3.3, we present the beta transformed linear pool, which acts as the foundation of our forecasting models. The methods used to evaluate the sharpness and calibration of our probabilistic forecasts are given in Section 3.4. This includes a brief introduction to the method of cross validation, which is a schematic way of fitting and evaluating a model.

### 3.1 Beta Distribution

A random variable  $Y \in [0,1]$  follows a beta distribution with parameters  $\alpha$  and  $\beta$ , if its pdf  $f(y; \alpha, \beta)$  is

$$f(y; \alpha, \beta) = b_{\alpha, \beta}(y) = \frac{1}{B(\alpha, \beta)} y^{\alpha-1} (1-y)^{\beta-1},$$

where

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)},$$

is the beta function,  $\Gamma(\cdot)$  is the gamma function and  $\alpha, \beta > 0$  (Johnson et al., 1995). We will later denote the beta cdf as  $B_{\alpha, \beta}(y)$ , which should not be confused with the beta function  $B(\alpha, \beta)$ . The expected value and variance are given as

$$E[Y] = \frac{\alpha}{\alpha + \beta} \quad \text{and} \quad \text{Var}(Y) = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}.$$

For modelling purposes, it is convenient to reparameterize the beta distribution in terms of its mean and a precision parameter (Ferrari and Cribari-Neto, 2004). Let  $\mu = \alpha / (\alpha + \beta)$  and  $\nu = \alpha + \beta$ , with  $\mu \in (0, 1)$  and  $\nu > 0$ . It then follows that

$$\alpha = \mu\nu, \quad \beta = \nu(1 - \mu),$$

and

$$E[Y] = \mu, \quad \text{Var}(Y) = \frac{\mu(1 - \mu)}{1 + \nu}.$$

This reparametrization allows us to model the mean  $\mu$  and the precision parameter  $\nu$ .

The beta distribution is well suited for modelling purposes because its density can attain a wide variety of shapes by varying the two parameters  $\mu$  and  $\nu$ . Figure 3.1 shows a few different examples of how the density behaves for different parameter values. We see that the density is symmetric when  $\mu = 0.5$ , and asymmetric when  $\mu \neq 0.5$ . Furthermore, the dispersion increases as  $\nu$  decreases. In particular, the beta distribution can also attain the form of a standard uniform distribution when  $\mu = 0.5$  and  $\nu = 2$ .

## 3.2 Logit Link

Consider a vector of observations  $\mathbf{y} = (y_1, \dots, y_n)$ . This vector is assumed to be a realization of a random variable  $\mathbf{Y} = (Y_1, \dots, Y_n)$ , whose components are independently distributed with means  $\boldsymbol{\mu} = (\mu_1, \dots, \mu_n)$ . Let each observation  $y_i$  be related to a set of  $p$  explanatory variables

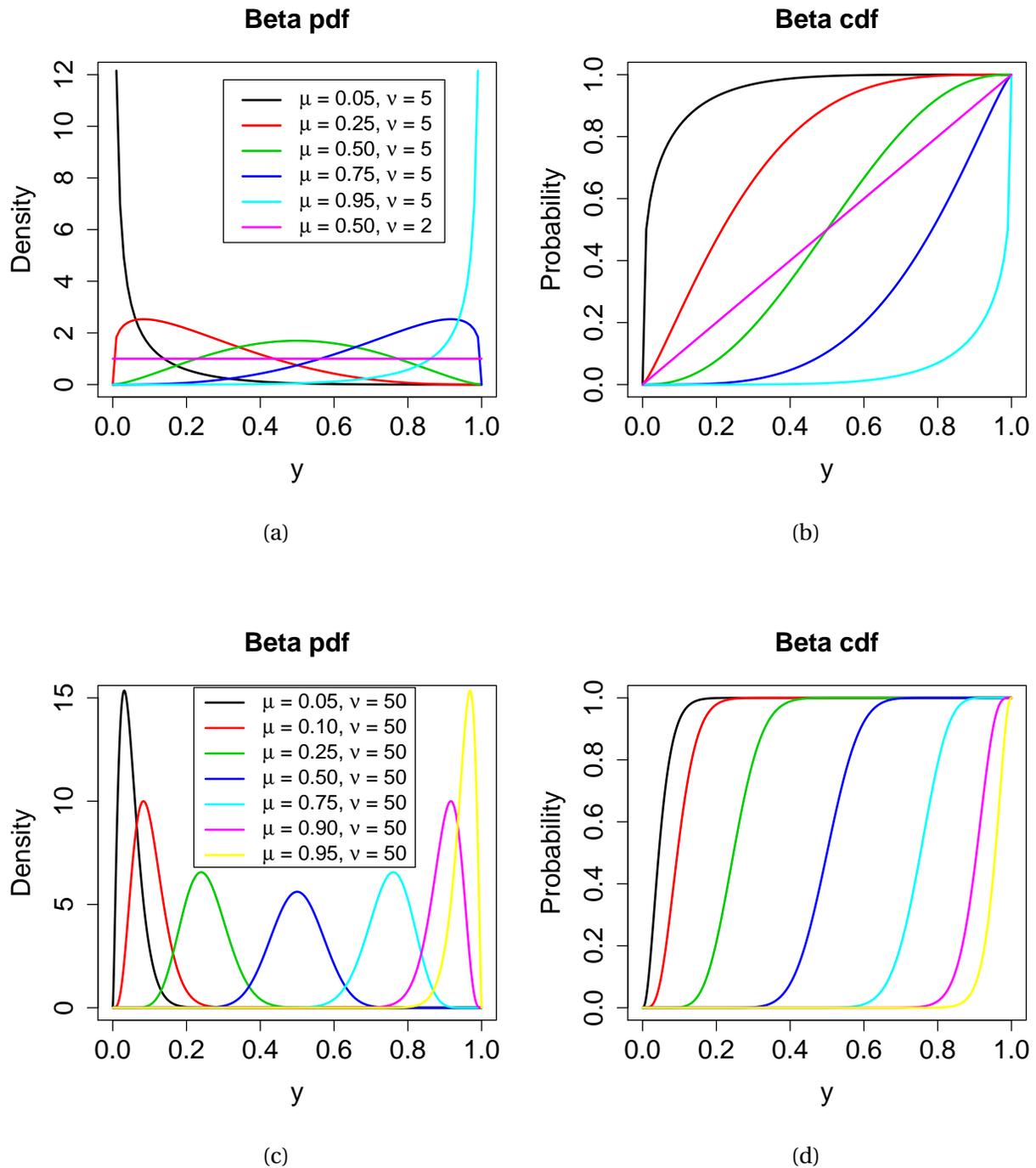


Figure 3.1: Beta pdf and cdf for varying parameter  $\mu$  and fixed parameter  $\nu$ .  $\nu = 5$  in (a)/(b) and  $\nu = 50$  in (c)/(d).

$\mathbf{x}_i = (x_{i1}, \dots, x_{ip})$ . The linear predictor  $\eta_i$  is then defined as

$$\eta_i = \sum_{j=1}^p x_{ij} \beta_j, \quad i = 1, \dots, n,$$

where  $\eta_i$  can take any value on the real line. Least squares estimators of the parameters  $\beta_j$ ,  $j = 1, \dots, p$ , are obtained by fitting the multiple linear regression model

$$y_i = \sum_{j=1}^p x_{ij} \beta_j + \epsilon_i,$$

to the data set  $\{(y_i, x_{i1}, \dots, x_{ip})\}_{i=1}^n$ , where  $\epsilon_i$  is the random error associated with  $y_i$  (Walpole et al., 2012).

For generalized linear models, a link function relates the linear predictor  $\eta_i$  to the mean value  $\mu_i$  of  $Y_i$ . If we write

$$\eta_i = g(\mu_i),$$

then  $g(\cdot)$  is called the link function. In classical linear models the link function is the identity function, and the predictor  $\eta_i = \mu_i$ . However, when  $\mu \in (0, 1)$  as for the beta distribution, we need a link function which maps  $(0, 1)$  onto the real line. The logit function given by

$$\eta_i = \text{logit}(\mu_i) = \log\left(\frac{\mu_i}{1 - \mu_i}\right), \quad (3.1)$$

does exactly this and is continuous and increasing on  $(0, 1)$  (McCullagh and Nelder, 1989). More explicitly, we use the inverse logit function

$$\mu_i = \text{logit}^{-1}(\eta_i) = \frac{1}{1 + \exp(-\eta_i)},$$

to transform the real valued linear predictor  $\eta_i$  into  $\mu_i$ , the mean value of  $Y_i$ . Figure 3.2 shows how the inverse logit function maps the the real axis onto the interval  $(0, 1)$ .

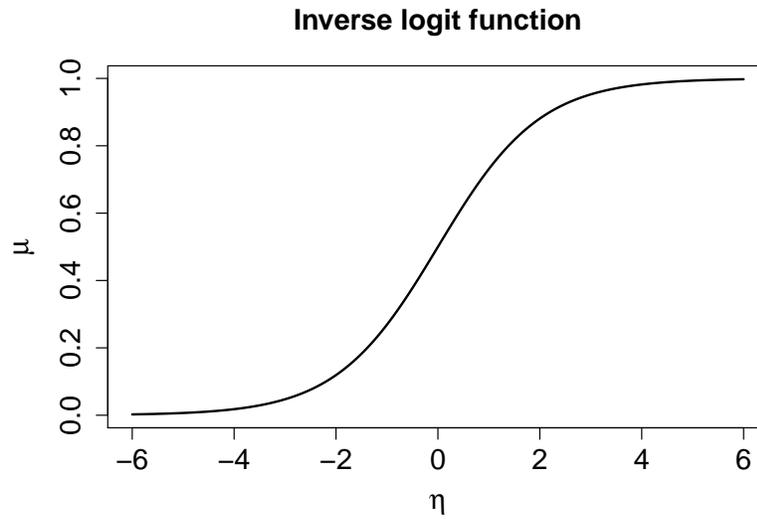


Figure 3.2: The inverse logit link function, connecting  $\mu$  with  $\eta$ .

### 3.3 Beta Transformed Linear Pool (BLP)

A probabilistic forecast can be represented in the form of a cdf  $F(y)$ . In many situations there are several complementary or competing forecasts available from various sources. Let  $\{F_1(y), \dots, F_k(y)\}$ ,  $y \in \mathbb{R}$  be a set of available probabilistic forecasts of the random variable  $Y$ . To aggregate this set of individual predictive distribution functions into a single combined forecast, we have to specify an aggregation method. The prevalent method is the linear pool, which is defined as

$$F(y) = \sum_{i=1}^k w_i F_i(y), \quad (3.2)$$

where the weights  $w_i \geq 0 \forall i$  and  $\sum_{i=1}^k w_i = 1$  (Wallis, 2005). The idea is to weigh each forecast based on their respective performance over a training period.

Despite the success of the linear pool in a large number of applications, Hora (2004) and Ranjan and Gneiting (2010) point at its shortcomings and limitations. They showed for special cases that if each of the individual predictive distributions are calibrated, any nontrivial linear combination is necessarily uncalibrated. A trivial linear pool is the case when one weight is equal to one, and the rest is zero, resulting in one of the original forecasts. Gneiting and Ranjan (2013) generalize these findings and state that dispersion tends to increase under any linear aggregation. As stated in Section 3.4, calibration is a desired and important property for a probabilistic

forecast. Therefore, we want other aggregation methods that outperform the linear pool.

One such method investigated by [Gneiting and Ranjan \(2013\)](#) is the beta transformed linear pool (BLP). This is a nonlinear aggregation method defined as

$$F_{\alpha,\beta}(y) = B_{\alpha,\beta}\left(\sum_{i=1}^k w_i F_i(y)\right), \quad (3.3)$$

where  $B_{\alpha,\beta}(\cdot)$  is the beta cdf with  $\alpha > 0$  and  $\beta > 0$ , and  $F_{\alpha,\beta}(y)$  is on the form of a cdf. This method composites on the linear pool (3.2) with a beta transform.

Let  $f_{\alpha,\beta}(y)$  be the corresponding pdf of (3.3), and  $f_i(y)$  the pdf related to  $F_i(y)$ . Using the well known relation  $f_{\alpha,\beta}(y) = \frac{d}{dx} F_{\alpha,\beta}(y)$  and the chain rule of differentiation, the pdf of the beta transformed linear pool is defined as

$$f_{\alpha,\beta}(y) = \left(\sum_{i=1}^k w_i f_i(y)\right) b_{\alpha,\beta}\left(\sum_{i=1}^k w_i F_i(y)\right), \quad (3.4)$$

where  $b_{\alpha,\beta}(y)$  denotes the beta probability density with parameters  $\alpha > 0$  and  $\beta > 0$ .

In practice, the weights  $w_1, \dots, w_k$  and parameters  $\alpha$  and  $\beta$  need to be estimated from training data  $\{(F_{1j}(y), \dots, F_{kj}(y), y_j) : j = 1, \dots, J\}$ , where  $F_{1j}, \dots, F_{kj}$  are probabilistic forecasts of  $y_j$  on the form of cdfs. Let  $f_{1j}, \dots, f_{kj}$  be the corresponding pdfs. Using (3.4) and the training data, we can find the log-likelihood function which is defined as.

$$l(w_1, \dots, w_k, \alpha, \beta | y_1, \dots, y_J) = \sum_{j=1}^J \log\left(\sum_{i=1}^k w_i f_{ij}(y_j)\right) + \sum_{j=1}^J \log\left(b_{\alpha,\beta}\left(\sum_{i=1}^k w_i F_{ij}(y_j)\right)\right). \quad (3.5)$$

We get the maximum likelihood estimators of the parameters  $w_1, \dots, w_k$ ,  $\alpha$  and  $\beta$  by maximizing the likelihood function (3.5) with respect to these parameters.

### 3.4 Evaluation Methods

The purpose of this project is to develop and investigate models for predicting wind power production, and from these be able to provide probabilistic forecasts. The performance of the models is evaluated with respect to both predictive sharpness and calibration, as these are desired forecast properties. As stated by [Gneiting et al. \(2007\)](#), calibration refers to the statistical consis-

tency between the forecast distribution and the observations, whereas sharpness refers to the concentration of the predictive distributions. With both actual values and predictions in hand, we need ways to quantify how well the model did in relation to sharpness and calibration. This topic is addressed in this section.

### 3.4.1 RMSE

The root-mean-square-error (RMSE) is a common way to measure the differences between observed values and predicted values. It represents the sample standard deviation of the differences between predicted values and observed values. Let  $\hat{y}_t$  be the prediction of the observed  $y_t$  at time  $t$ . If  $t = 1, \dots, n$  then

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (y_t - \hat{y}_t)^2}{n}}. \quad (3.6)$$

### 3.4.2 PIT-diagram

The probability integral transform (PIT) relates to the fact that random variables from any given continuous distribution can be transformed into random variables having a standard uniform distribution. [Gneiting et al. \(2007\)](#) describe how this technique can be used to compare a probabilistic forecast with the distribution of the corresponding observations. The PIT thus relates to the calibration of the forecasts.

We assume that the observed outcome  $y_t$  of a process in nature follows the distribution  $G_t$  at times  $t = 1, 2, \dots, n$ , which is the true data generating distribution. As forecasters, we provide probabilistic forecasts in the form of cdfs,  $F_t$  at  $t = 1, 2, \dots, n$ . If  $F_t = G_t$  for all  $t$ , we are perfect forecasters. It is not a straightforward task to compare these two distributions, especially since  $G_t$  is never known, only hypothetical. However, we have the observations  $y_t$  that are samples from  $G_t$ .

Consider now the hypothetical situation where we know  $G_t$ . If we then computed  $p_t = G_t(y_t)$ ,  $p_t$  would be a sample from a standard uniform distribution. Moving back to the real situation where we only know  $F_t$ , we instead do the comparable calculation  $p_t = F_t(y_t)$ . If we are perfect forecasters,  $p_t$  is still a sample from a uniform distribution. However, more often

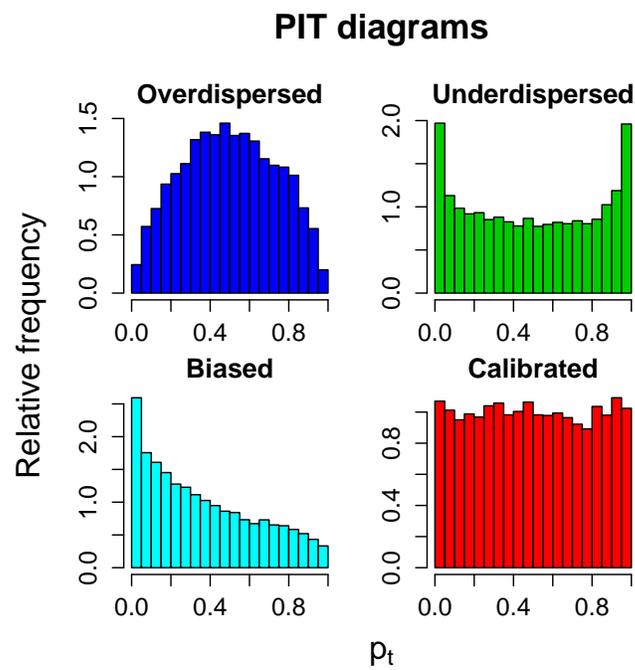
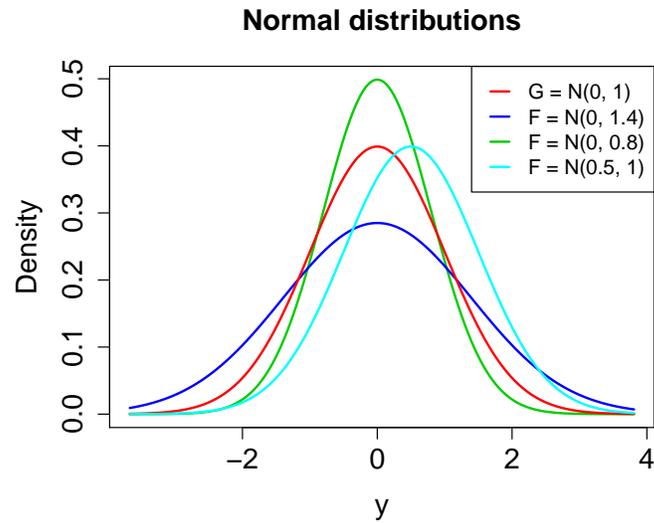


Figure 3.3: (a): The red curve  $G$  is the true data generating distribution, while the other distributions are various probabilistic forecast distributions. (b): PIT-diagrams for the distributions in (a) with respect to observations from  $G$ .

than not,  $F_t \neq G_t$  for all  $t$ , and we want to study how they are different.

A histogram of  $p_t = F_t(y_t)$  for  $t = 1, 2, \dots, n$  often provides hints to the reason for forecast deficiency. Figure 3.3 shows some examples. In the upper plot, we have several suggested probabilistic forecasts  $F$  together with the true data generating distribution  $G$ , plotted in red. The lower plot shows the corresponding PIT-diagrams. We see that the blue curve is overdispersed, which results in a hump-shaped PIT-diagram. The green curve is underdispersed, which results in a U-shaped PIT-diagram. The cyan curve is biased, and results in a triangle-shaped PIT-diagram. If the forecast is properly calibrated we get a uniformly distributed PIT-diagram.

Even though a PIT-diagram can give an indication to the reason for forecast deficiency, uniformity of the PIT-diagram is only a necessary condition for the forecaster to be perfect, not sufficient. Gneiting et al. (2007) reiterate an example from Hamill (2001), where a perfect forecaster having  $F_t = G_t$  for all  $t$ , has three other competing forecasters. The three competitors have slightly different forecasting models compared to the perfect forecaster. Despite this, the PIT-diagrams for all four forecasters are essentially uniform, and cannot be used to distinguish between the ideal forecaster and the competitors.

### 3.4.3 CRPS

The continuous ranked probability score (CRPS) is a verification tool for probabilistic forecast systems, and is a quantity that relates to both forecast sharpness and calibration. It can be defined as follows (adapted from Hersbach, 2000).

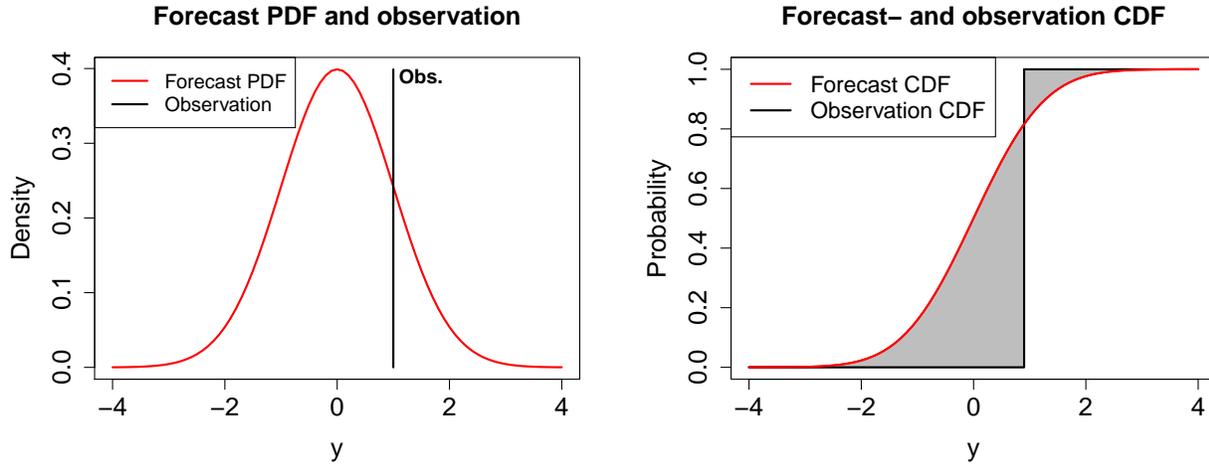
The forecast pdf at time  $t$  is denoted by  $f_t$ , and the corresponding observation by  $y_t$ . Figure 3.4a shows an example where  $f_t$  has a standard normal distribution, and  $y_t = 1$ . The two cdfs for  $f_t$  and  $y_t$  are then respectively given as

$$F_t(y) = \int_{-\infty}^y f_t(\epsilon) d\epsilon \quad \text{and}$$

$$P_t(y) = H(y - y_t),$$

where

$$H(y) = \begin{cases} 0 & \text{for } y < 0 \\ 1 & \text{for } y \geq 0 \end{cases}$$



(a) Example of forecast pdf at time  $t$ , together with the corresponding observation  $y_t$ .

(b) The grey area represents  $CRPS_t$  through the relation in (3.7), such that minimizing the area minimizes  $CRPS_t$ .

Figure 3.4: Visual interpretation of CRPS where the probabilistic forecast is a standard normal distribution.

is the so-called Heaviside function. The CRPS at time  $t$  is then defined as

$$CRPS_t = CRPS(F_t, y_t) = \int_{-\infty}^{\infty} [F_t(y) - P_t(y)]^2 dy. \quad (3.7)$$

A visualization at a given time  $t$  can be seen in Figure 3.4b. The red and black lines represent  $F_t$  and  $P_t$  respectively. The CRPS is negatively oriented, in that a smaller CRPS indicates a better forecast. The area between  $F_t$  and  $P_t$  marked grey is directly linked to  $CRPS_t$ , i.e. minimizing this area also minimizes  $CRPS_t$ . The area (and  $CRPS_t$ ) is minimized when the observation is equal to the median of the probabilistic forecast. In practice, the CRPS is averaged over an interval of time, such that we get one single value for the CRPS:

$$\overline{CRPS} = \frac{\sum_{t=1}^n CRPS_t}{n}.$$

One advantage of the CRPS is that it reduces to absolute error if the forecast is deterministic, i.e. when there is no randomness involved in the forecast. To understand this, consider Figure 3.4b as the forecast cdf becomes a step function, like the observation. Then the area between these two functions is given by the rectangle formed by the two step functions. The

rectangle will have height equal to 1, such that the area directly represents the absolute error. In practice, this makes it possible to directly compare a deterministic forecast with a probabilistic forecast in a consistent manner.

### 3.4.4 Cross Validation

Cross validation is the preferred way to fit a model and compute predictions. Consider a data set  $S$ . Divide the data into two parts: *Training set* and *test set*. The idea is then to fit the model using the training set, and test the model using the test set. Another option when fitting a model would be to use all data available. However, this approach implies that no data is left to do testing, and we are forced to use that same data for testing as we used for fitting. The evaluation of the predictions become too optimistic, because the model is fitted to minimize the error based on this particular data set.

We describe here a common type of cross validation called *k-fold cross validation*. A schematic overview of this method is given as follows:

1. Divide the data set  $S$  into  $k$  groups, or *folds*,  $S_1, \dots, S_k$  of approximately equal size. How the data is divided depends on the nature of the data and model.
2. For  $j = 1, \dots, k$ , do the following:
  - (a) For the  $j$ -th fold (test set), fit the model to the other  $k - 1$  folds (training set).
  - (b) Do testing of the model on the  $j$ -th fold.
3. Combine the test values for all test sets.

In this project,  $k$ -fold cross validation is used, where the data is divided into  $k$  subsequent folds.

# Chapter 4

## Methodological Concept

In this chapter we describe the concept behind our forecasting methodology, and start with a toy example in Section 4.1 to make the idea behind clear. There are several ways to build up a forecasting model following this concept. Section 4.2 introduces the CCPR model presented by Borhaug (2014), which is the basis and inspiration behind the two other forecasting models following the same concept. Based on challenges experienced with the CCPR model, we introduce the UCPR model in Section 4.3 and the CPR-LP model in Section 4.4. All sections in this chapter use the same synthetic data to give examples and illustrations which are used to easier understand the properties of the models.

### 4.1 A Toy Example

Our forecasts are based on explanatory variables, which is available information related to the quantity we want to forecast. We utilize these explanatory variables to construct probabilistic forecasts in the form of a density function, either a pdf or a cdf. In this section we present a toy example to illustrate some desired properties of this forecast. The model behind the example is the CCPR, but the details behind this model is first presented in Section 4.2. In this section we instead focus on how the resulting probabilistic forecast is a function of the explanatory variables. Even though the CCPR model is used to carry out the example, the forecast properties we observe in this section are also valid for the other forecasting models introduced later in this chapter.

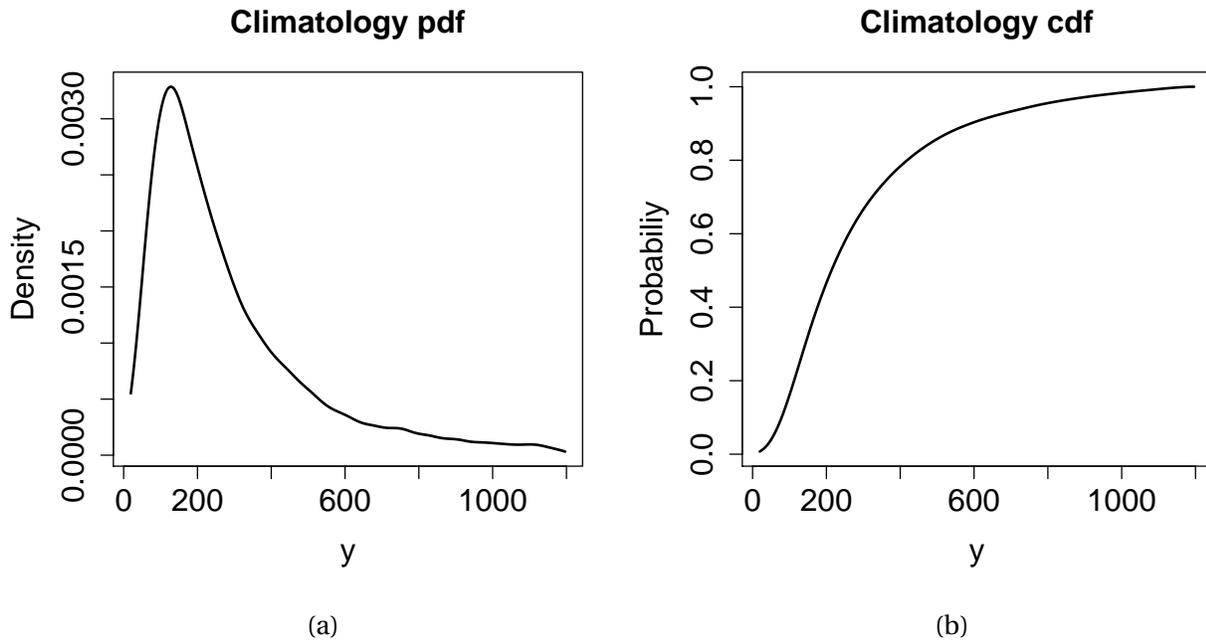


Figure 4.1: Climatology distribution as (a) pdf  $c(y)$  and (b) cdf  $C(y)$ .

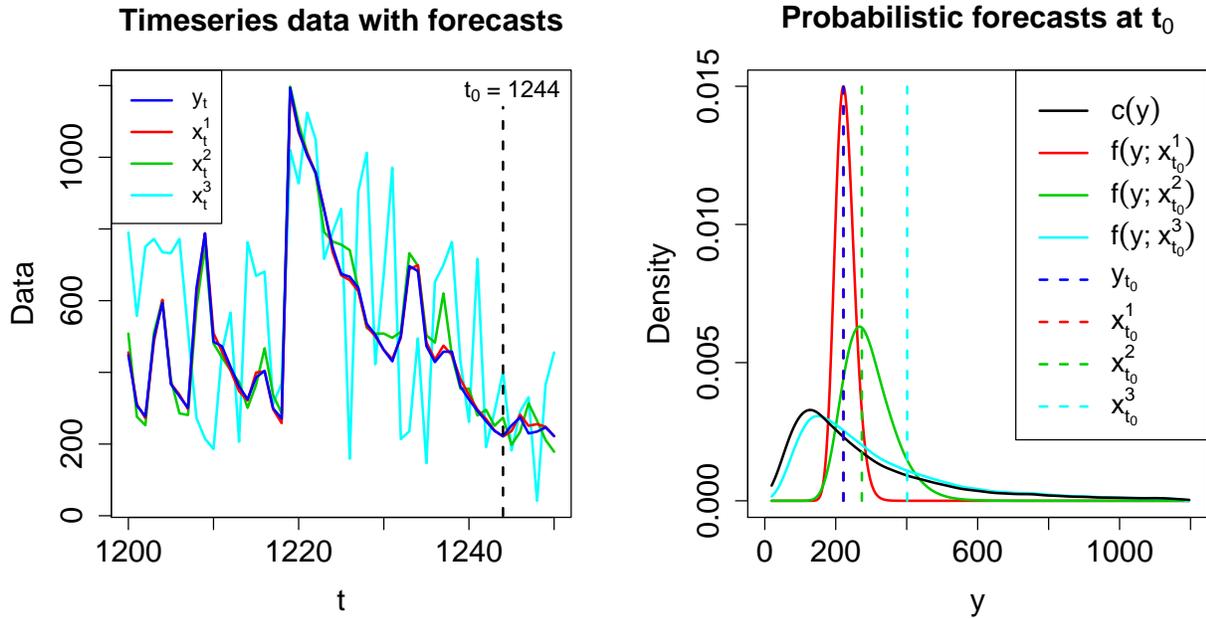
Establishing a probabilistic forecast can be done in numerous ways. Our method starts with an initial forecast which is some prior belief on the forecast density. For the CCPR model, which is the model used in this example, the initial forecast is based on the weather climatology and is denoted by  $c(y)$  (pdf) or  $C(y)$  (cdf). Let the observed time series from the process we want to forecast be denoted by  $\mathbf{y} = (y_1, \dots, y_n)$  where  $n = 10000$ , and  $y_t \in [a, b]$  for  $t = 1, \dots, n$ . The actual wind power production data presented in Section 2.1 has such a restriction. The power production is never below zero, and each area has a maximum production capacity. The climatology pdf  $c(y)$  of the time series  $\mathbf{y}$  is simply a sample distribution function of these observations, and is plotted in Figure 4.1 together with the cdf  $C(y)$ . Like we mentioned in the beginning of this chapter, the observations that form the climatology are synthetic data. However, the data is constructed to make the climatology in Figure 4.1 mimic the climatology of our case data in Chapter 2. All argumentation in this section and the rest of this chapter based on the climatology from the synthetic data is therefore also valid for our case study. This is important to remember throughout the rest of this chapter as some of our forecasting models are developed to perform well when the climatology has the form in Figure 4.1.

If we did not have any other information, the climatology could be used as our probabilistic

forecast for any new observation from the same process. This is not a very sophisticated forecast as it does not change with time, but is the same for all future time points. However, we have access to a deterministic forecast for  $\mathbf{y}$ , denoted by  $\mathbf{x} = (x_1, \dots, x_n)$ . This deterministic forecast can be used to adjust the climatology  $c(y)$ , and results in a new adapted probabilistic forecast  $f(y; x_t)$ . Note the indexing including  $t$ , as we get such an adjusted forecast for each unit of time. Typically, the new forecast pdf  $f(y; x_t)$  is sharper and shifted towards the deterministic forecast  $x_t$ . Section 4.2 explains in detail how this adjustment is done. For now, let us focus on how the shape and sharpness of  $f(y; x_t)$  is dependent on the predictive performance of  $\mathbf{x}$ .

We introduce three different deterministic forecasts for  $\mathbf{y}$  with different predictive performance. These three forecasts are denoted by  $\mathbf{x}^1 = (x_1^1, \dots, x_n^1)$ ,  $\mathbf{x}^2 = (x_1^2, \dots, x_n^2)$  and  $\mathbf{x}^3 = (x_1^3, \dots, x_n^3)$ , and result in the three corresponding probabilistic forecasts  $f(y; x_t^1)$ ,  $f(y; x_t^2)$  and  $f(y; x_t^3)$ . A section of the four data series  $\mathbf{y}$ ,  $\mathbf{x}^1$ ,  $\mathbf{x}^2$  and  $\mathbf{x}^3$  is plotted in Figure 4.2a. As we can see,  $\mathbf{x}^1$  represented by the red line is the most accurate forecast and has a RMSE of 10.1. The second most accurate forecast is  $\mathbf{x}^2$  (green) with a RMSE of 40.7, followed by  $\mathbf{x}^3$  (cyan) with RMSE of 266.6. RMSE is a way to measure the predictive performance of a deterministic forecast, and is defined in (3.6).

Figure 4.2b shows the probabilistic forecasts at time  $t_0 = 1244$ , together with the climatology from Figure 4.1a. The corresponding deterministic forecasts at  $t_0$  are marked in Figure 4.2a. As we can see, the red curve  $f(y; x_{t_0}^1)$  has the most concentrated density. This is in accordance with the fact that  $\mathbf{x}^1$  is the most accurate forecast.  $f(y; x_{t_0}^2)$  is wider because  $\mathbf{x}^2$  is a slightly less accurate deterministic forecast compared to  $\mathbf{x}^1$ , as shown in Figure 4.2a. The third probabilistic forecast  $f(y; x_{t_0}^3)$  is almost identical to the climatology  $c(y)$ . From Figure 4.2a we see how the deterministic forecast  $\mathbf{x}^3$  has no clear correlation with the observations  $\mathbf{y}$ . In this case,  $\mathbf{x}^3$  has next to no information about  $\mathbf{y}$ , and the best probabilistic forecast is therefore the climatology itself. This result marks an important property of the probabilistic forecasts. As the deterministic forecasts get less accurate, the corresponding probabilistic forecasts converge towards the climatology.



(a) Plot of the four time series  $y$ ,  $x^1$ ,  $x^2$  and  $x^3$  in the time interval  $[1200, 1250]$ .

(b) Three probabilistic forecasts  $f(y; x_{t_0}^1)$ ,  $f(y; x_{t_0}^2)$  and  $f(y; x_{t_0}^3)$  of  $y_{t_0}$  at time  $t_0 = 1244$ .  $x_{t_0}^1$ ,  $x_{t_0}^2$  and  $x_{t_0}^3$  are the deterministic forecasts at time  $t_0$  in (a).  $x_{t_0}^1$  is almost not visible, because  $x_{t_0}^1$  and  $y_{t_0}$  are close to being identical.

Figure 4.2

## 4.2 CCPR (Climatology Cumulative Probability Regression)

The previous section focused on how the predictive performance of  $\mathbf{x}$  affects the shape and sharpness of  $f(y; x_t)$ . This is an adjusted probabilistic forecast based on the initial forecast, which was the climatology  $c(y)$  in last section. In this section we introduce the CCPR forecasting model (Borhaug, 2014), which is one way to perform the adjustment. This model is the basis and inspiration behind the development of our other forecasting models which all follow the same methodological concept.

We keep the notation introduced in Section 4.1, but also introduce the notion of a collection of deterministic forecasts at time  $t$ . To repeat, we might have several forecasts  $x^1, x^2, \dots$  of  $y$ . The collection of deterministic forecasts at time  $t$  is then denoted as  $\mathbf{x}_t^{1,2,\dots} = (x_t^1, x_t^2, \dots)$ , where the superscript 1, 2,  $\dots$  denotes the forecasts included.

Our starting point is the initial forecast  $C(y)$ , i.e. the climatology, which is a cdf that trans-

forms observations belonging to an interval  $[a, b]$  into the unit interval  $[0, 1]$ . The climatology cdf in Figure 4.1b illustrates this, where  $a = 0$  and  $b = 1200$ . We want to find a better probabilistic forecast for  $y_t$  than the climatology itself. Given the deterministic forecasts  $\mathbf{x}_t^{1,2,\dots}$ , our approach proceeds by transforming  $C(y)$  into a new cdf  $F_C(y; \mathbf{x}_t^{1,2,\dots})$ . The goal is to make  $F_C(y; \mathbf{x}_t^{1,2,\dots})$  a sharper forecast of  $y_t$  compared to  $C(y)$ , while still calibrated. Note that in Section 4.1, we only used probabilistic forecasts of the form  $F_C(y; x_t^1)$  (or  $f_C(y; x_t^1)$ ), i.e. only a function of one single deterministic forecast. This is only a special case of the more general  $F_C(y; \mathbf{x}_t^{1,2,\dots})$ , which is a function of several deterministic forecasts.

The transformation from  $C(y)$  to  $F_C(y; \mathbf{x}_t^{1,2,\dots})$  is done using the beta cdf  $B_{\alpha,\beta}(\cdot)$ , i.e. a beta transform. This particular choice of transformation function is based on the beta transformed linear pool from [Ranjan and Gneiting \(2010\)](#) and [Gneiting and Ranjan \(2013\)](#). The beta distribution is defined on the interval  $[0, 1]$ , which suits our purpose well as  $C(y)$  has exactly this range. In Section 3.1, we saw that the beta distribution is appropriate for such a transformation, as it can take on a large variety of shapes (Figure 3.1). The  $\alpha$  and  $\beta$  parameters are not necessarily modelled as constants over time, but could both be modelled as a function of the deterministic forecasts  $\mathbf{x}_t^{1,2,\dots}$  ([Borhaug, 2014](#)). This way, the beta distribution is time varying through the deterministic forecasts. Section 4.5 explains how  $\alpha$  and  $\beta$  are modelled.

We can finally link the climatology  $C(y)$  with the adjusted  $F_C(y; \mathbf{x}_t^{1,2,\dots})$  and express the CCPR-model through the following expression.

$$F_C(y; \mathbf{x}_t^{1,2,\dots}) = B_{\mathbf{x}_t^{1,2,\dots}}(C(y)), \quad (4.1)$$

where  $B_{\mathbf{x}_t^{1,2,\dots}}(\cdot)$  denotes the beta cdf determined by the collection of deterministic forecasts. This expression is similar to (3.3) if we set  $k = 1$  and let  $F_1(y) = C(y)$ , i.e. the model in (4.1) is a special case of the beta transformed linear pool (BLP) model from Section 3.3. In the case where  $k = 1$ , the BLP-model can be used to provide calibration and dispersion correction of  $C(y)$  ([Gneiting and Ranjan, 2013](#)).

To make the somewhat abstract transformation in (4.1) more clear, we have illustrated the procedure in the upper two plots of Figure 4.3. We are going to follow the transformation in detail for one point  $y'$ , marked by the dotted lines. The model starts out with the climatology  $C(y)$

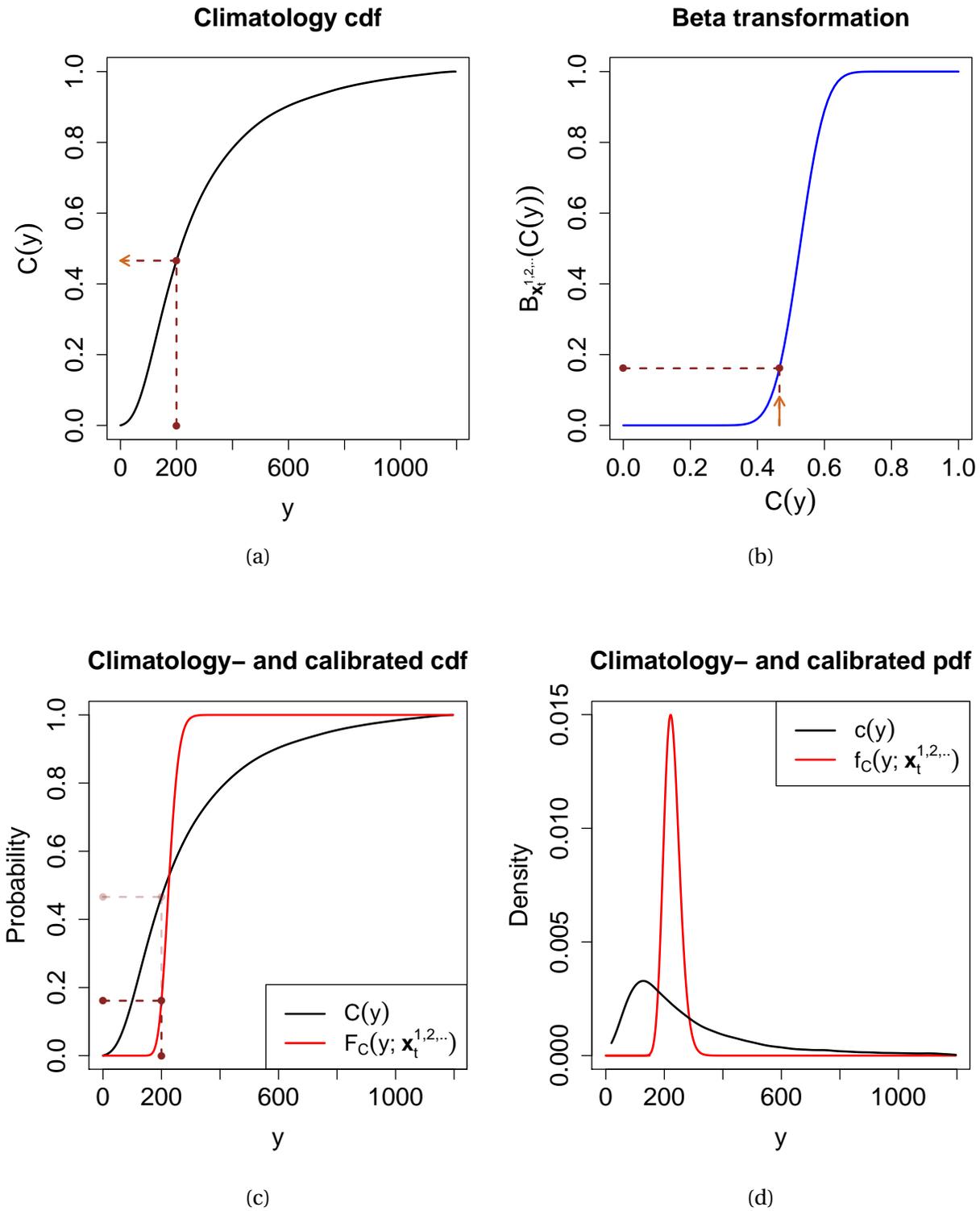


Figure 4.3: Illustration of the transformation procedure through the climatology in (a) and the beta transform in (b), resulting in the final calibrated probabilistic forecast cdf  $F_C(y; \mathbf{x}_t^{1,2,\dots})$  in (c) (and as pdf in (d)).

plotted in Figure 4.3a. Along the x-axis are all possible outcomes of the forecast. Throughout the paper, we refer to the possible outcomes as the domain. In relation to our case study, the domain is represented by the possible values for wind power production. Along the y-axis are the functional values of the climatology  $C(y)$  for the whole domain. We see from Figure 4.3a that  $C(y') = P(y' < 200) \approx 0.47$ , i.e. the climatology tells us that any new production value is less than or equal to 200 with a probability of 0.47. The functional value of  $C(y)$  is a probability, which ranges between 0 and 1.

The beta transformation function  $B_{\mathbf{x}_t^{1,2,\dots}}(C(y))$  is plotted in Figure 4.3b. Along the x-axis are values from the climatology  $C(y)$ , and along the y-axis are functional values of  $B_{\mathbf{x}_t^{1,2,\dots}}(C(y))$  which also range between 0 and 1. The functional value of the beta transformation is in fact also a probability, which means it transforms one probability into another. Following the point  $y'$  from before gives  $B_{\mathbf{x}_t^{1,2,\dots}}(C(y')) = P(C(y') < 0.47) \approx 0.16$ . This result is visualized with the dotted line in Figure 4.3b. The beta transformation  $B_{\mathbf{x}_t^{1,2,\dots}}(C(y))$  thus tells us that any new climatology value  $C(y)$  is less than or equal to 0.47 with a probability of 0.16.

This transformation of  $y'$  is performed for all  $y \in [0, 1200]$ , and we end up with the new probabilistic CCPR forecast  $F_C(y; \mathbf{x}_t^{1,2,\dots})$  plotted in red in Figure 4.3c. The dark dotted line in this plot shows that  $F_C(y'; \mathbf{x}_t^{1,2,\dots}) = B_{\mathbf{x}_t^{1,2,\dots}}(C(y')) \approx 0.16$ . This is a substantial change from the climatology which gave  $C(y') \approx 0.47$  (represented by the transparent dotted line). The new probabilistic CCPR forecast is plotted as a pdf in Figure 4.3d, and compared with the climatology pdf.

The performance of this forecasting model is investigated thoroughly in Chapter 6. However, we end this section by studying the shape of the climatology and how it affects the probabilistic forecasts returned by the CCPR-model. This is relevant as the results are a direct cause for the introduction of another similar forecasting model, namely the UCPR-model in Section 4.3.

Consider the two collections of forecasts  $\mathbf{x}_{t_1}^{1,2,\dots}$  and  $\mathbf{x}_{t_2}^{1,2,\dots}$  (at time  $t_1$  and  $t_2$ ). These two collections give the two corresponding beta transformations  $B_{\mathbf{x}_{t_1}^{1,2,\dots}}(\cdot)$  and  $B_{\mathbf{x}_{t_2}^{1,2,\dots}}(\cdot)$ , plotted in Figure 4.4. To shorten the notation for this example, let the two beta transformations be denoted by  $B_1(\cdot)$  and  $B_2(\cdot)$  respectively. The two beta transformations in Figure 4.4 have similar shape, but  $B_2(\cdot)$  is shifted to the right compared to  $B_1(\cdot)$ . This way, we can compare two similar transformations placed on two different areas of the domain.

Figure 4.5a displays the corresponding CCPR forecast cdfs  $F_{C_1}(y; \mathbf{x}_{t_1}^{1,2,\dots})$  and  $F_{C_2}(y; \mathbf{x}_{t_2}^{1,2,\dots})$ .

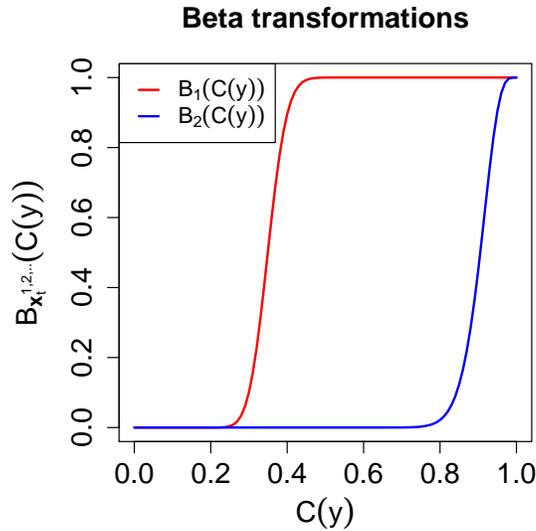


Figure 4.4: The two example beta transformations  $B_{\mathbf{x}_{t_1}^{1,2,\dots}}(\cdot)$  and  $B_{\mathbf{x}_{t_2}^{1,2,\dots}}(\cdot)$  denoted by  $B_1(\cdot)$  and  $B_2(\cdot)$  respectively (to shorten notation).

The corresponding pdfs  $f_{C_1}(y; \mathbf{x}_{t_1}^{1,2,\dots})$  and  $f_{C_2}(y; \mathbf{x}_{t_2}^{1,2,\dots})$  are plotted in Figure 4.5b. This reveals an interesting result. Even though the two beta transformation functions are similar, the resulting CCPR forecasts are very different in shape. This difference is due to the shape of the climatology.

To explain this difference, we look back at the beta transformations in Figure 4.4. The interesting domain of transformation is where the function value is above zero and below one, i.e. the domain where we see the slope of the function. Looking at  $B_1(\cdot)$ , this is approximately when  $C(y) \in (0.2, 0.4)$ . Comparing with Figure 4.5a, the slope of  $C(y)$  is quite steep when  $C(y) \in (0.2, 0.4)$ , which results in a compressed slope for  $F_{C_1}(y; \mathbf{x}_{t_1}^{1,2,\dots})$ . Moving on to  $B_2(\cdot)$  in Figure 4.4, the slope of the function is approximately when  $C(y) \in (0.8, 1.0)$ . Comparing this with Figure 4.5a, the slope of  $C(y)$  is quite gentle when  $C(y) \in (0.8, 1.0)$ , which results in a stretched slope for  $F_{C_2}(y; \mathbf{x}_{t_2}^{1,2,\dots})$ . Comparing with Figure 4.5b, we see how these findings translates to the forecast pdfs: The forecast is sharp when the climatology has high probability density, and not so sharp when the climatology has low probability density. As sharp forecasts are desired, we want to find a way to obtain sharp forecasts on the domain where the climatology has low density, which in this case is for the upper half of the domain.

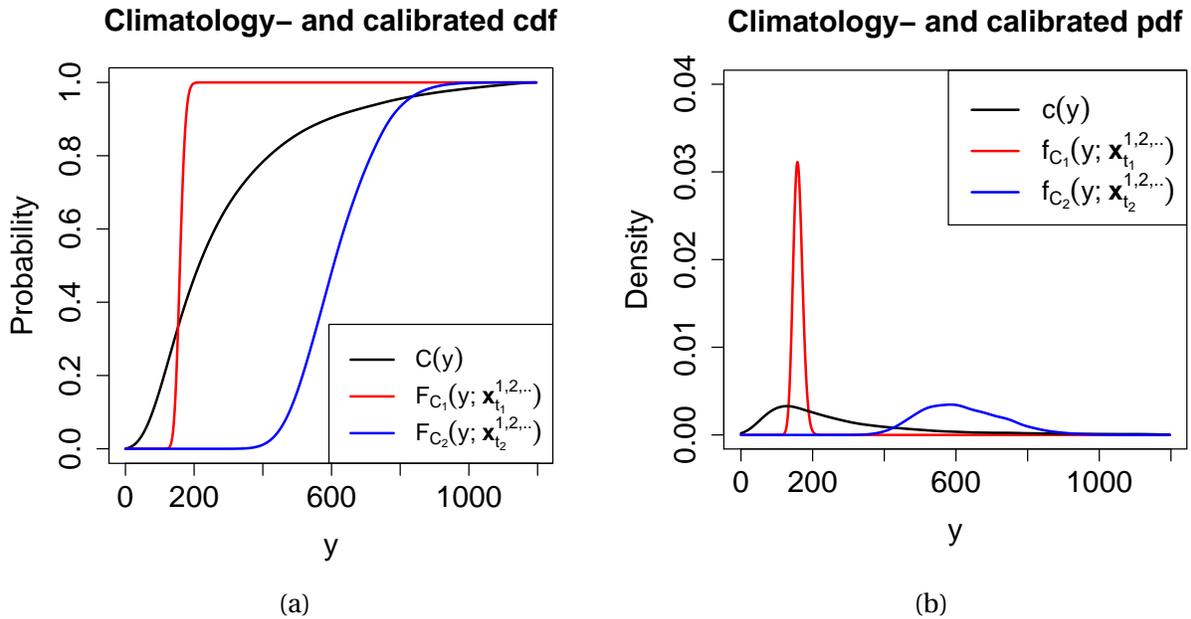


Figure 4.5: Climatology and CCPR forecasts as (a) cdfs and (b) pdfs. The CCPR forecasts are computed using the two beta transformations in Figure 4.4.

### 4.3 UCPR (Uniform Cumulative Probability Regression)

Let  $u(y)$  be the uniform distribution over the domain of possible observations. This uniform distribution has the same density over the whole domain, which means it does not have a diminishing density for the upper half of the domain like the climatology  $c(y)$ . The uniform distribution  $u(y)$  is the initial forecast of the UCPR model. Let  $F_U(y; \mathbf{x}_t^{1,2,\dots})$  be the UCPR forecast at time  $t$ , given the collection of forecasts  $\mathbf{x}_t^{1,2,\dots}$ . The UCPR model is then defined as

$$F_U(y; \mathbf{x}_t^{1,2,\dots}) = B_{\mathbf{x}_t^{1,2,\dots}}(U(y)), \quad (4.2)$$

where  $B_{\mathbf{x}_t^{1,2,\dots}}(\cdot)$  denotes the beta cdf determined by the collection of deterministic forecasts. Substituting the climatology cdf  $C(y)$  in (4.1) by the uniform distribution cdf  $U(y)$ , we end up with the UCPR model in (4.2). The modelling of the parameters that define  $B_{\mathbf{x}_t^{1,2,\dots}}(\cdot)$  are set up in Section 5.3. Like the CCPR-model in (4.1), the UCPR-model in (4.2) is a special case of the beta transformed linear pool in (3.3) where  $k = 1$  and  $F_1(y) = U(y)$ .

Like we did with the CCPR-model in Section 4.2, we want to study the shape of the uniform distribution  $U(y)$  and how it affects the probabilistic forecasts returned by the UCPR-

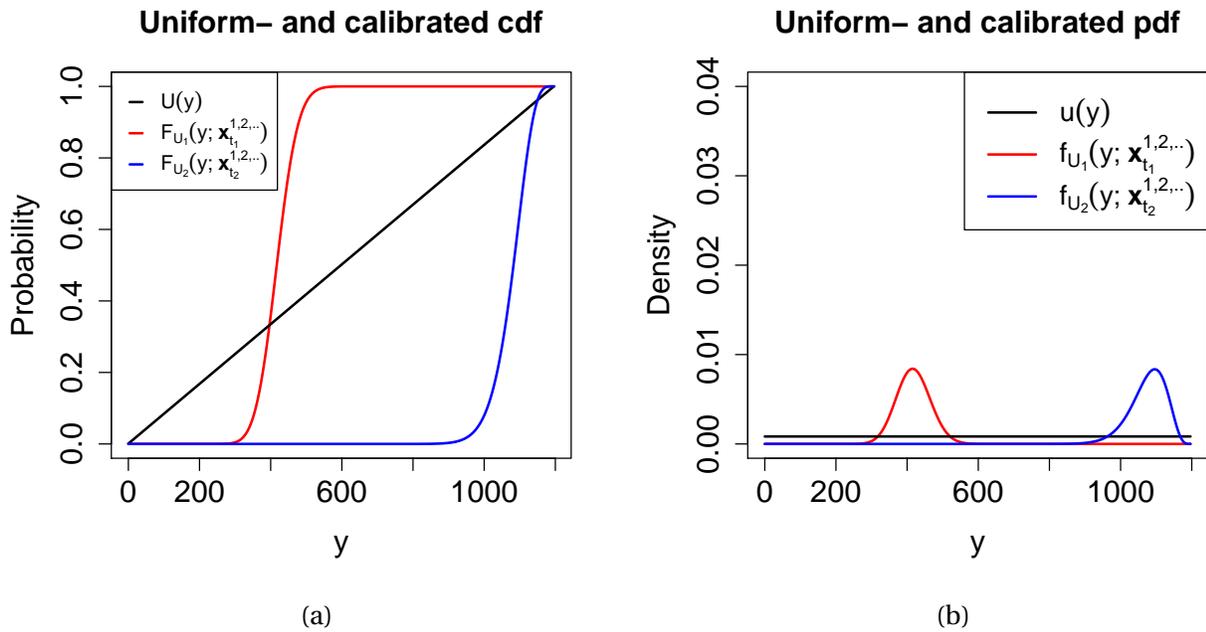


Figure 4.6: Uniform distribution and UCPR forecasts as (a) cdfs and (b) pdfs. The UCPR forecasts are computed using the two beta transformations in Figure 4.4.

model. We still use the two beta transformations  $B_1(\cdot)$  and  $B_2(\cdot)$  from Figure 4.4, which give the UCPR forecasts  $F_{U_1}(y; \mathbf{x}_{t_1}^{1,2,\dots})$  and  $F_{U_2}(y; \mathbf{x}_{t_2}^{1,2,\dots})$  plotted in Figure 4.6a. The corresponding pdfs  $f_{U_1}(y; \mathbf{x}_{t_1}^{1,2,\dots})$  and  $f_{U_2}(y; \mathbf{x}_{t_2}^{1,2,\dots})$  are plotted in Figure 4.6b. The result is quite different from the corresponding CCPR forecasts in Figure 4.5. Because the uniform cdf  $U(y)$  has a constant slope over the whole domain, the two UCPR forecasts  $F_{U_1}(y; \mathbf{x}_{t_1}^{1,2,\dots})$  and  $F_{U_2}(y; \mathbf{x}_{t_2}^{1,2,\dots})$  attain the same shape and slope as the beta transformation itself (Figure 4.4). This translates into the two very similar pdfs in Figure 4.6b, which leads us to the conclusion that UCPR forecasts have similar sharpness and shape on the whole domain for similar beta transformations. Comparing the CCPR and UCPR forecasts in Figures 4.5 and 4.6 respectively, we note that UCPR obtains a sharper forecast for the upper domain, and CCPR obtains a sharper forecast for the lower domain.

## 4.4 CPR-LP (Cumulative Probability Regression with Linear Pooling)

The CPR-LP-model is designed to keep the benefits of the two previous models, while eliminating the disadvantages. Section 4.2 showed how the CCPR was capable of delivering a sharp forecast in the lower area of the domain. Likewise, Section 4.3 showed how the UCPR was capable of delivering a reasonably sharp forecast in the upper area of the domain. These are the properties we want to combine in the CPR-LP model.

To aggregate the two forecasts into a single combined forecast we simply use the beta transformed linear pool (Gneiting and Ranjan, 2013) from (3.3). The pool consists of both previous initial forecasts, the climatology  $C(y)$  and the uniform distribution  $U(y)$ . Let  $F_{LP}(y; \mathbf{x}_t^{1,2,\dots})$  be the CPR-LP forecast at time  $t$ , given the collection of forecasts  $\mathbf{x}_t^{1,2,\dots}$ . The model is then defined as

$$F_{LP}(y; \mathbf{x}_t^{1,2,\dots}) = B_{\mathbf{x}_t^{1,2,\dots}}(wC(y) + (1 - w)U(y)), \quad w \in [0, 1], \quad (4.3)$$

where  $B_{\mathbf{x}_t^{1,2,\dots}}(\cdot)$  denotes the beta cdf and  $w$  is the weight parameter. Both  $B_{\mathbf{x}_t^{1,2,\dots}}(\cdot)$  and  $w$  are modelled as a function of the collection of deterministic forecasts. The details behind the modelling of  $w$  and the parameters of  $B_{\mathbf{x}_t^{1,2,\dots}}(\cdot)$  can be found in Section 5.4.

Figure 4.7 illustrates how different values for the weight parameter  $w$  affects the CPR-LP forecast. Like the previous plots in Figure 4.5 and 4.6, the red forecasts are computed using  $B_1(\cdot)$  and the blue forecasts are computed using  $B_2(\cdot)$  (both plotted in Figure 4.4). The red and blue bars to the right indicate the weight parameter used for each plot. We see that the two CPR-LP forecasts are identical to the two CCPR forecasts (Figure 4.5) for  $w = 1$ , and identical to the two UCPR forecasts (Figure 4.6) for  $w = 0$ . Because  $w$  is modelled as a function of the deterministic forecasts, the CPR-LP model is able to weigh the climatology  $C(y)$  when forecasting on the lower part of the domain, and weigh the uniform distribution  $U(y)$  when forecasting on the upper part of the domain. As a result, the model is able to get the sharpest forecasts in both ends of the domain.

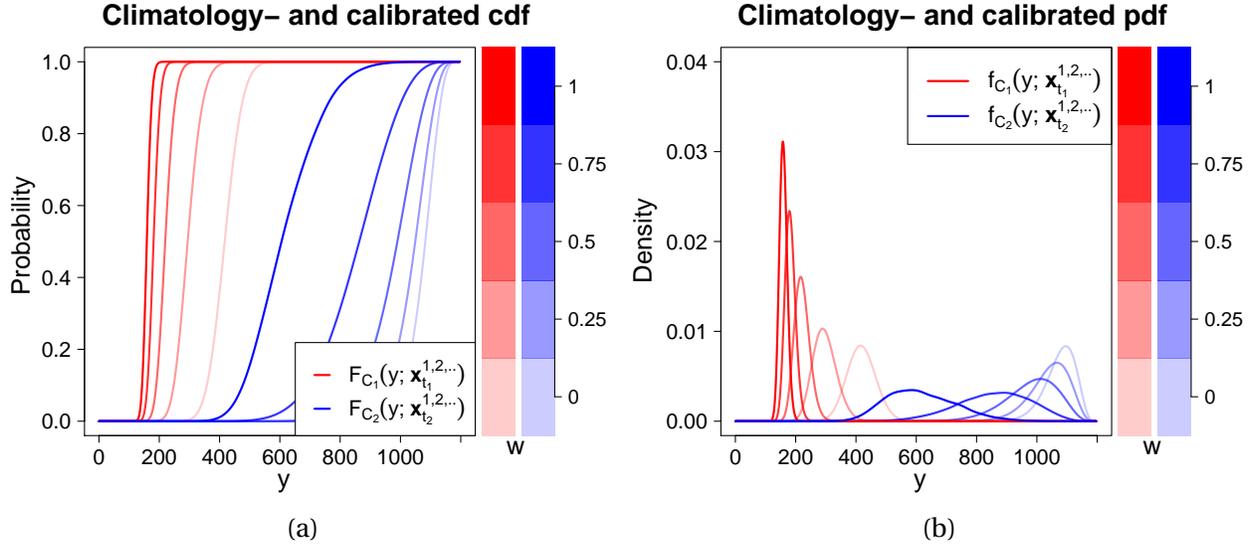


Figure 4.7: CPR-LP forecasts as (a) cdfs and (b) pdfs. The forecasts are computed using the two beta transformations in Figure 4.4.

## 4.5 Modelling the Beta Parameters $\mu$ , $\nu$ and the Weight

### Parameter $w$

The beta cdf  $B_{\mathbf{x}_t^{1,2,\dots}}(\cdot)$  used in the transformation procedure of the forecasting models in this chapter, is parameterized with mean  $\mu$  and precision parameter  $\nu$  (Section 3.1). This allows us to model the mean and the precision of the distribution separately (Ferrari and Cribari-Neto, 2004). To make sure  $\mu \in (0, 1)$ , we model it through the logit function (3.1) (Borhaug, 2014), restated here as

$$\text{logit}(\mu_t) = \log\left(\frac{\mu_t}{1 - \mu_t}\right) = \eta_t, \quad (4.4)$$

where  $\eta_t$  can be written as

$$\eta_t = q(\mathbf{x}_t^{1,2,\dots}),$$

where  $q(\cdot)$  is a function of the deterministic forecasts  $\mathbf{x}_t^{1,2,\dots}$  at time point  $t$ .

The precision parameter  $\nu$  is also modelled as a function of the deterministic forecasts.

$$\nu_t = r(\mathbf{x}_t^{1,2,\dots}).$$

Because  $\nu > 0$ , we need  $r(\mathbf{x}_t^{1,2,\dots}) > 0, \forall \mathbf{x}_t^{1,2,\dots}$ .

Finally, the weight parameter  $w$  is modelled similarly to  $\mu$  because  $w \in [0, 1]$ , i.e.

$$\text{logit}(w_t) = \log\left(\frac{w_t}{1-w_t}\right) = v_t, \quad (4.5)$$

where  $v_t$  can be written as

$$v_t = s(\mathbf{x}_t^{1,2,\dots}),$$

where also  $s(\cdot)$  is a function of the deterministic forecasts  $\mathbf{x}_t^{1,2,\dots}$  at time point  $t$ . Chapter 5 presents the choices of  $q(\cdot)$ ,  $r(\cdot)$  and  $s(\cdot)$  for each model in our case study.

# Chapter 5

## Case Study: Models and Evaluation Scheme

In this chapter we start by establishing some useful notation in Section 5.1. Our case study applies the three models CCPR, UCPR and CPR-LP introduced in Chapter 4, to the data from Chapter 2. Sections 5.2, 5.3 and 5.4 specify how the functions from Section 4.5 are defined for each of these three models respectively. These functions tell how to model the parameters belonging to each model. In addition, we have included a brief explanation of the ARn-SFR-model from Malmgård (2016) in Section 5.5. This model is used as a reference for model performance in Chapter 6. Section 5.6 presents visual interpretations of the model coefficients found in the CCPR, UCPR and CPR-LP models. Section 5.7 describes how the models are fitted and tested.

### 5.1 Notation: EC- and Persistence Forecast and Hourly Changes

As concluded in Section 2.2, the EC-forecast and the temporal dependency between actuals are the important sources of information for our forecasting models. The temporal dependency is critical to introduce a so-called persistence forecast. A persistence forecast relies upon previous observations to forecast the future.

Let  $Y_t$  denote future actual production at time  $t$ . The persistence forecast of  $Y_t$  is simply the last available observation denoted as  $x_{t,l}^P$ , where  $t$  denotes the forecast valid time, and  $l$  is the lead time. In other words, the persistence forecast  $x_{t,l}^P$  is the observation  $y_{t-l}$ . The EC-forecast of  $Y_t$  is denoted as  $x_t^{EC}$ . As we only have one EC-forecast per observation, we do not label it with lead time. Revisit Section 2.1 to see how the EC-forecast is generated. We denote a vector

including both forecasts by  $\mathbf{x}_{t,l}^{EC,P} = (x_t^{EC}, x_{t,l}^P)$ . In this chapter and Chapter 6, we refer to both the EC-forecast and the persistence forecast collectively as the deterministic forecasts.

As we saw in Figure 2.5b, the variance in actual wind power production changes during the day. This is an effect of the different wind conditions during a 24-hour cycle. Typically, the weather is more stable at night with lower wind speeds. Because of this daily cycle, it is natural to fit one climatology for each hour  $h$  of the day. We denote these  $C_h(y)$ , where  $h = 0, \dots, 23$ .

## 5.2 Model 1: CCPR

As we fit one climatology distribution  $C_h(y)$  for each hour  $h$  of the day, we define the CCPR model expression (4.1) as

$$F_C(y; h, \mathbf{x}_{t,l}^{EC,P}) = B_{\mathbf{x}_{t,l}^{EC,P}}(C_h(y)), \quad (5.1)$$

to explicitly indicate that the CCPR forecast does not only vary with the forecast  $y$ , but also with the hour  $h$  of the day. The beta cdf  $B_{\mathbf{x}_{t,l}^{EC,P}}(\cdot)$  in (5.1) is parameterized by  $\mu$  and  $\nu$ , where the mean  $\mu$  is related to the linear predictor  $\eta$  through the logit function in (4.4). The predictor  $\eta$  is modelled as a linear function of the EC-forecast and the persistence forecast, and the coefficients vary with lead time  $l$  and hour  $h$  of the day:

$$\eta_{t,l} = q(\mathbf{x}_{t,l}^{EC,P}) = \gamma_{h,l}^0 + \gamma_{h,l}^{EC} C_h(x_t^{EC}) + \gamma_{h,l}^P C_h(x_{t,l}^P), \quad (5.2)$$

where  $t$  denotes the forecast valid time. By inserting (5.2) into (4.4), we have the complete model for  $\mu_{t,l}$ .

The linear predictor  $q(\cdot)$  is a function of both the EC-forecast and the persistence forecast. This is desirable as the two deterministic forecasts switch between being the most accurate forecast for increasing lead time. Figure 2.4a shows how the temporal dependency between actuals decrease with increasing lead time, making the persistence forecast less desirable for higher lead times. Section 6.1.1 thoroughly investigates this.

Instead of using  $(x_t^{EC}, x_{t,l}^P)$  as covariates, we use  $(C_h(x_t^{EC}), C_h(x_{t,l}^P))$  (Borhaug, 2014). This way,  $\eta_{t,l}$  is affected by changes in  $(C_h(x_t^{EC}), C_h(x_{t,l}^P))$ , and not necessarily by changes in  $(x_t^{EC}, x_{t,l}^P)$ . Consider for example the cumulative climatology in Figure 4.1b, and let  $\eta_1 = q(\mathbf{x}_1^{EC,P})$  and  $\eta_2 =$

$q(\mathbf{x}_2^{EC,P})$ , where  $\mathbf{x}_1^{EC,P} = (800, 800)$  and  $\mathbf{x}_2^{EC,P} = (1000, 1000)$ . If  $\eta_{t,l}$  is defined with  $(C_h(x_t^{EC}), C_h(x_{t,l}^P))$  as covariates,  $\eta_1$  and  $\eta_2$  are almost identical, as the climatology is almost constant from 800 to 1000. However, if  $\eta_{t,l}$  is defined with  $(x_t^{EC}, x_{t,l}^P)$  as covariates,  $\eta_1$  and  $\eta_2$  are more different from each other. This is assuming  $\gamma_{h,l}^{EC}$  and  $\gamma_{h,l}^P$  are significantly different from zero.

The precision parameter  $v$  has a somewhat more complicated model defined by

$$v_{t,l} = r(\mathbf{x}_{t,l}^{EC,P}) = \exp(a_{h,l}^0) + \exp(a_{h,l}^{EC})[C_h(x_t^{EC})(1 - C_h(x_{t,l}^{EC}))] \\ + \exp(a_{h,l}^P)[C_h(x_{t,l}^P)(1 - C_h(x_{t,l}^P))] + \exp(a_{h,l}^D)[1 - |C_h(x_t^{EC}) - C_h(x_{t,l}^P)|], \quad (5.3)$$

where the coefficients vary with lead time  $l$  and hour  $h$  of the day, and  $t$  denotes the forecast valid lead time. The second and third term in (5.3) are only functions of the EC-forecast or the persistence forecast respectively. They model the precision as a symmetric function of the forecasts, with a centred maximum. A more thorough interpretation of these two terms and (5.3) as a whole is done in Appendix B. The fourth term models the precision as an increasing function with increasing absolute difference between the two deterministic forecasts. The intercept is not a function of the deterministic forecasts and is important in the cases where all the other terms are close to zero. For example, when both the EC- and persistence forecast are close to zero and identical.

### 5.3 Model 2: UCPR

With respect to modelling, the UCPR model is very similar to the CCPR model. The only difference between the two is the initial forecast, where the climatology  $C_h(y)$  in CCPR is substituted with the uniform distribution  $U_h(y)$  for the UCPR.

Because we fit the uniform distribution  $U_h(y)$  for each hour  $h$  of the day, the UCPR model expression (4.2) is denoted as

$$F_U(y; h, \mathbf{x}_{t,l}^{EC,P}) = B_{\mathbf{x}_{t,l}^{EC,P}}(U_h(y)), \quad (5.4)$$

to explicitly indicate that the UCPR forecast varies with the hour of the day. The linear predictor

$\eta$  is modelled as

$$\eta_{t,l} = q(\mathbf{x}_{t,l}^{EC,P}) = \gamma_{h,l}^0 + \gamma_{h,l}^{EC} U_h(x_t^{EC}) + \gamma_{h,l}^P U_h(x_{t,l}^P). \quad (5.5)$$

The difference between (5.5) and (5.2) is that the climatology  $C_h(y)$  in (5.2) is substituted with the uniform distribution  $U_h(y)$  in (5.5). In contrast with the CCPR model, there is now no substantial difference between using  $(x_t^{EC}, x_{t,l}^P)$  or  $(U_h(x_t^{EC}), U_h(x_{t,l}^P))$  as covariates. Because  $U_h(y)$  is a linear function, it only scales any deterministic forecast onto the interval  $[0, 1]$ . However, we would like the coefficients of (5.2) and (5.5) to be of similar magnitude, and therefore choose  $(U_h(x_t^{EC}), U_h(x_{t,l}^P))$  as covariates in (5.5).

The precision parameter  $v$  is modelled by

$$v_{t,l} = r(\mathbf{x}_{t,l}^{EC,P}) = \exp(a_{h,l}^0) + \exp(a_{h,l}^{EC}) [U_h(x_t^{EC})(1 - U_h(x_t^{EC}))] \\ + \exp(a_{h,l}^P) [U_h(x_{t,l}^P)(1 - U_h(x_{t,l}^P))] + \exp(a_{h,l}^D) [1 - |U_h(x_t^{EC}) - U_h(x_{t,l}^P)|]. \quad (5.6)$$

This model is similar to the dispersion parameter model for CCPR in (5.3). The difference is that the climatology  $C_h(y)$  in (5.3) is substituted with the uniform distribution  $U_h(y)$  in (5.6).

## 5.4 Model 3: CPR-LP

Also for CPR-LP, we fit the climatology  $C_h(y)$  and the uniform distribution  $U_h(y)$  for each hour  $h$  of the day. Therefore, the CPR-LP model from (4.3) is denoted as

$$F_{LP}(y; h, \mathbf{x}_{t,l}^{EC,P}) = B_{\mathbf{x}_{t,l}^{EC,P}}(wC_h(y) + (1 - w)U_h(y)), \quad w \in [0, 1]. \quad (5.7)$$

The linear predictor  $\eta$  is modelled similarly to the CCPR model, i.e.

$$\eta_{t,l} = q(\mathbf{x}_{t,l}^{EC,P}) = \gamma_{h,l}^0 + \gamma_{h,l}^{EC} C_h(x_t^{EC}) + \gamma_{h,l}^P C_h(x_{t,l}^P). \quad (5.8)$$

The precision parameter  $\nu$  is modelled similarly to the UCPR model, i.e.

$$\begin{aligned} \nu_{t,l} = r(\mathbf{x}_{t,l}^{EC,P}) = & \exp(a_{h,l}^0) + \exp(a_{h,l}^{EC})[U_h(x_t^{EC})(1 - U_h(x_t^{EC}))] \\ & + \exp(a_{h,l}^P)[U_h(x_{t,l}^P)(1 - U_h(x_{t,l}^P))] + \exp(a_{h,l}^D)[1 - |U_h(x_t^{EC}) - U_h(x_{t,l}^P)|]. \end{aligned} \quad (5.9)$$

The weight parameter  $w$  is related to the linear predictor  $\nu$  through the logit function in (4.5).

The predictor  $\nu$  is modelled similarly to  $\eta$  in (5.2). Only the coefficients are different, i.e.

$$\nu_{t,l} = s(\mathbf{x}_{t,l}^{EC,P}) = \omega_{h,l}^0 + \omega_{h,l}^{EC}C_h(x_t^{EC}) + \omega_{h,l}^PC_h(x_{t,l}^P). \quad (5.10)$$

We end this section by commenting the use of either the climatology  $C_h(y)$  or the uniform distribution  $U_h(y)$  in the modelling of the CPR-LP parameters. The CCPR model uses the same initial forecast, the climatology, in both the beta transformation, and to model the beta parameters. The same goes for UCPR, but with the uniform distribution as initial forecast instead of the climatology. Interchanging the initial forecasts has been tried, but the CCPR and UCPR models perform better when we use the same initial forecast to model the parameters as the one used in the beta transformation. However, the CPR-LP model performs a beta transformation of (a weighted sum of) both the climatology and the uniform distribution. Therefore, it is not clear whether or not we should use the climatology or the uniform distribution to model the parameters. With three parameters to model and two initial forecast to choose between, there are several combinations. However, after testing the performance of the CPR-LP model with all possible combinations, the combination presented in this section (with  $C_h(y)$  in (5.8) and (5.10), and  $U_h(y)$  in (5.9)) proved to give the best results.

## 5.5 Model 4: ARn-SFR

Like we mentioned in Chapter 1, the forecasting methodology we propose is in response to the results of Malmgård (2016). The ARn-SFR forecasting model was the best performing model tested by Malmgård (2016), and is used as a reference model with respect to forecasting performance in Chapter 6. In this section we restate the description of the ARn-SFR-model from Malmgård (2016). The notation is changed to follow the notation introduced in Section 5.1.

Let  $\mathbf{y}$  be the vector of all actual production values. These actuals are assumed to be a realization of a random variable  $\mathbf{Y}$  whose components are independent normal variables with means  $\boldsymbol{\mu}$ , and constant variance.

The ARn-SFR-model is defined by modelling  $\mu_t$  through the following linear relationship.

$$\mu_t = \beta_{h,l}^0 + \beta_{h,l}^{EC} x_t^{EC} + \beta_{h,l}^P y_{t-l},$$

where  $x_t^{EC}$  is the EC-forecast at valid time  $t$ , and  $y_{t-l}$  is the last observed actual with lead time  $l$ . The coefficients  $\beta_{h,l}^0$ ,  $\beta_{h,l}^{EC}$  and  $\beta_{h,l}^P$  are unique for each hour  $h$  of the day and lead time  $l$ .

## 5.6 Toy Models: Interpretation of Model Coefficients Using Synthetic Data

In this Section, we are going to take a closer look at the coefficients in the CCPR, UCPR and CPR-LP models. We let both the EC-forecast and the persistence forecast be random synthetic data, without any temporal dependency, i.e. independent and identically distributed random variables from a uniform distribution. We then fit the models using these deterministic forecasts, and plot the coefficients for the various bidding areas and lead times. The purpose is to give more meaning to the somewhat abstract model definitions in this chapter. A similar interpretation of the ARn-SFR model coefficients can be found in [Malmgård \(2016\)](#).

### 5.6.1 CCPR coefficients: Synthetic Forecasts

Figure 5.1 displays the fitted model coefficients. We explained in Section 4.1 that the model will converge towards the climatology when the deterministic forecasts have no information about the observations. The CCPR model returns the climatology when the beta transformation function in (5.1) corresponds to the uniform cdf, i.e. when the mean parameter  $\mu = 0.5$  and the precision parameter  $\nu = 2$  (Section 3.1). Remember that  $\mu$  is related to the linear predictor  $\eta$  through the logit function in (4.4), such that  $\mu = 0.5$  when  $\eta = 0$ .

The coefficients  $\gamma^0$ ,  $\gamma^{EC}$  and  $\gamma^P$  in Figure 5.1 belong to the modelling of  $\eta$  in (5.2). We hereafter refer to these coefficients as the mean coefficients. They are all centred around zero for all

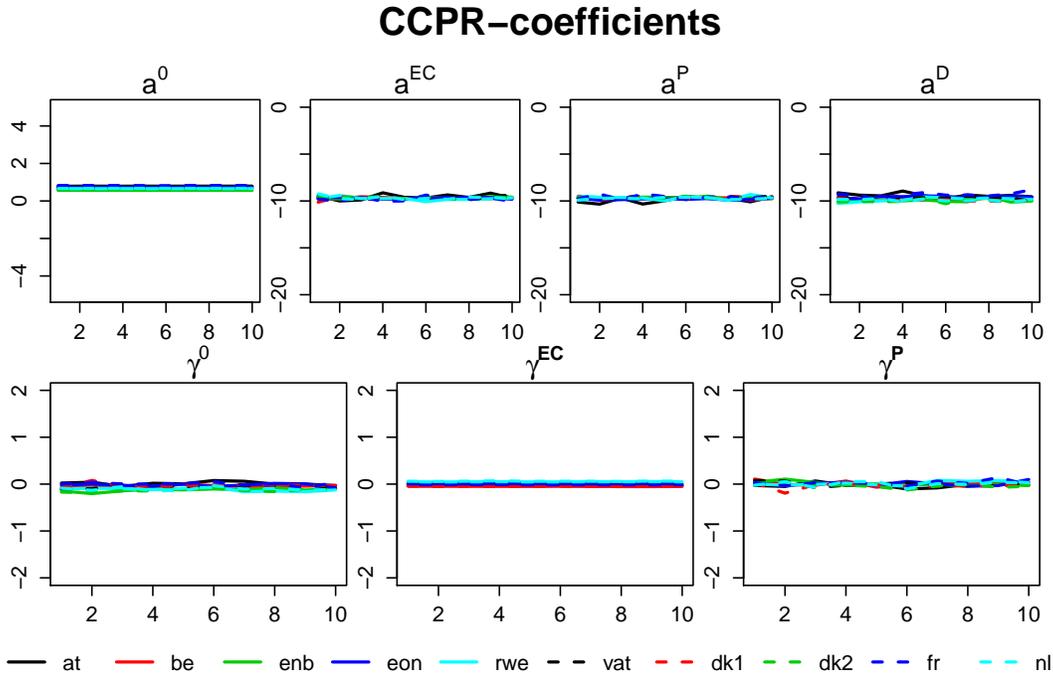


Figure 5.1: The coefficients of the CCPR model for all bidding areas when both the EC-forecast and the persistence forecast are synthetic random time series. Lead times in hours are on the x-axis.

lead times and areas, which give  $\eta = 0$ , i.e.  $\mu = 0.5$ . Consider now the coefficients  $a^0$ ,  $a^{EC}$ ,  $a^P$  and  $a^D$  used to model the precision parameter  $\nu$  in (5.3). We hereafter refer to these coefficients as the precision coefficients.  $a^{EC}$ ,  $a^P$  and  $a^D$  are all close to -10. We work with the exponential of these coefficients in (5.3) and the exponential of a negative number of this magnitude is approximately equal to zero, which means we are only left with  $\exp(a^0)$  in (5.3). The coefficient  $a^0$  is approximately equal to 0.7 for all lead times and areas, such that  $\nu \approx \exp(a^0) \approx 2$ .

These values for  $\mu$  and  $\nu$  make the beta transformation a uniform cdf, and the CCPR forecast in (5.1) is thus the climatology itself as expected. Figure 5.2 illustrates this result for the forecast cdf in the upper row, and for the forecast pdf in the lower row. The first column displays the climatology. The second column displays the beta transformation which is a uniform distribution. The third column displays the CCPR forecast which is similar to the climatology itself because of the uniform beta transformation.

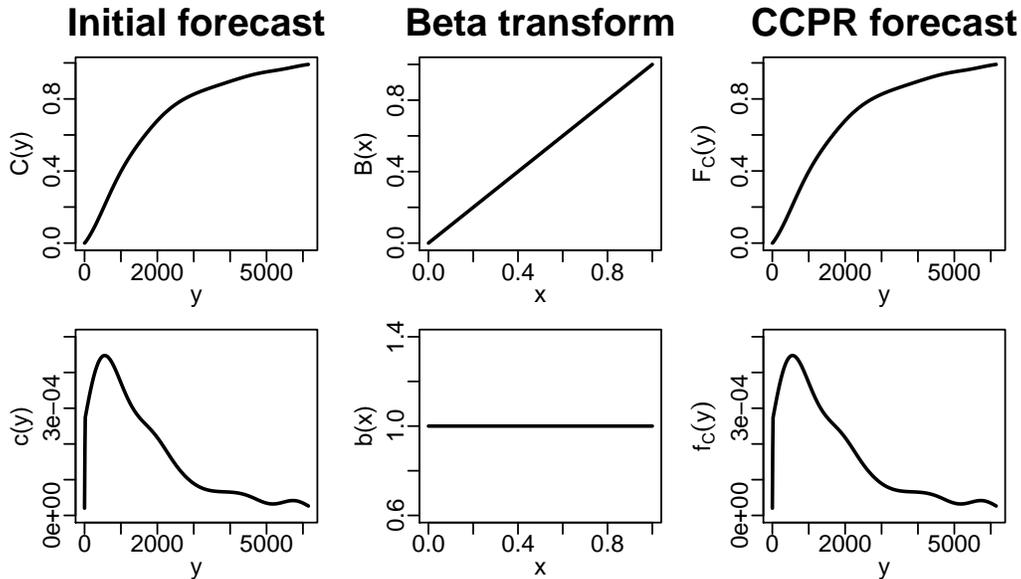


Figure 5.2: Illustration of how the CCPR is similar to the climatology itself when the model coefficients are estimated as in Figure 5.1.

### 5.6.2 UCPR coefficients: Synthetic Forecasts

Figure 5.3 displays the fitted model coefficients. If we compare Figure 5.3 with Figure 5.1, the precision coefficients  $a^0$ ,  $a^{EC}$ ,  $a^P$  and  $a^D$  from (5.3) and (5.6) are almost identical. As a result, the precision parameter  $\nu = 2$ , just like it was for CCPR in Section 5.6.1. Moving over to the mean coefficients  $\gamma^0$ ,  $\gamma^{EC}$  and  $\gamma^P$  from (5.5) however, there is a change compared to CCPR. All coefficients are no longer zero. The coefficient  $\gamma^0$  ranges between -1 and -0.5, varying with bidding area. This results in a mean parameter  $\mu$  ranging between 0.26 and 0.38. Figure 5.4 illustrates how these coefficient values affect the UCPR forecast, as cdf in the upper row, and as pdf in the lower row. The plots show the results for the bidding area RWE in Germany, i.e. when  $\gamma^0 = -1$ .

Section 4.1 and Figure 4.2b explained that our forecasting methodology will converge towards the climatology if there is no information about the observations in the deterministic forecasts. The CCPR model in (5.1) is a beta transformation of the climatology. For this model to return the climatology, the beta transformation has to be a uniform distribution function, which will return the climatology itself like we saw in Figure 5.2. The UCPR in (5.4) on the other hand, is a beta transformation of a uniform distribution. To make the UCPR forecast similar to the climatology, the beta transformation itself has to mimic the shape of the climatology. This is

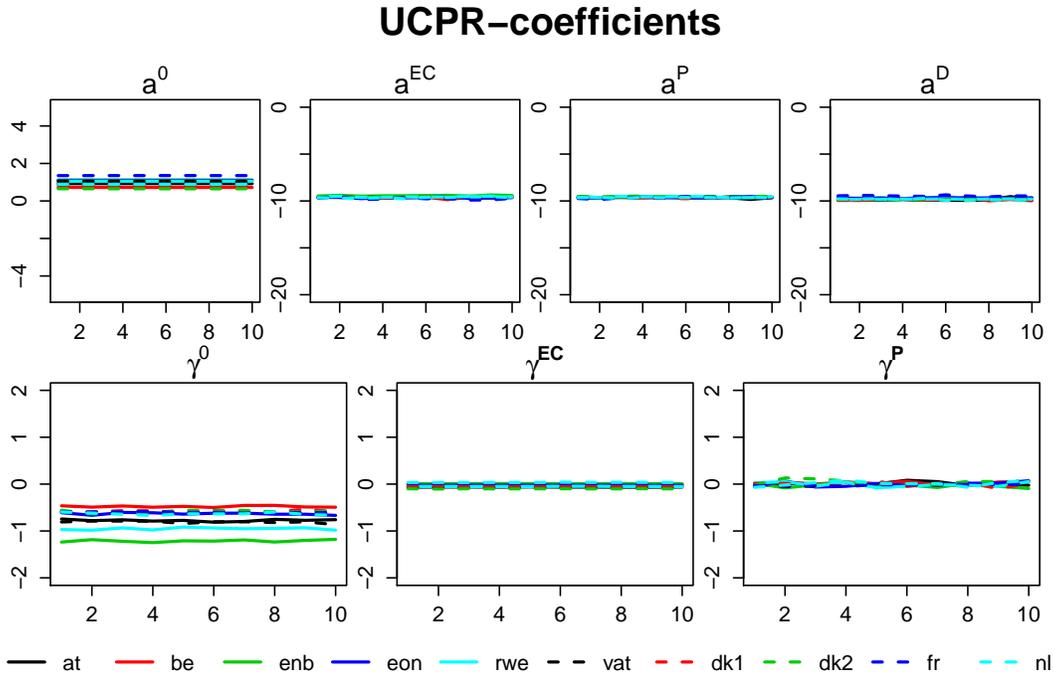


Figure 5.3: The coefficients of the UCPR model for all bidding areas when both the EC-forecast and the persistence forecast are synthetic random time series. Lead times in hours are on the x-axis.

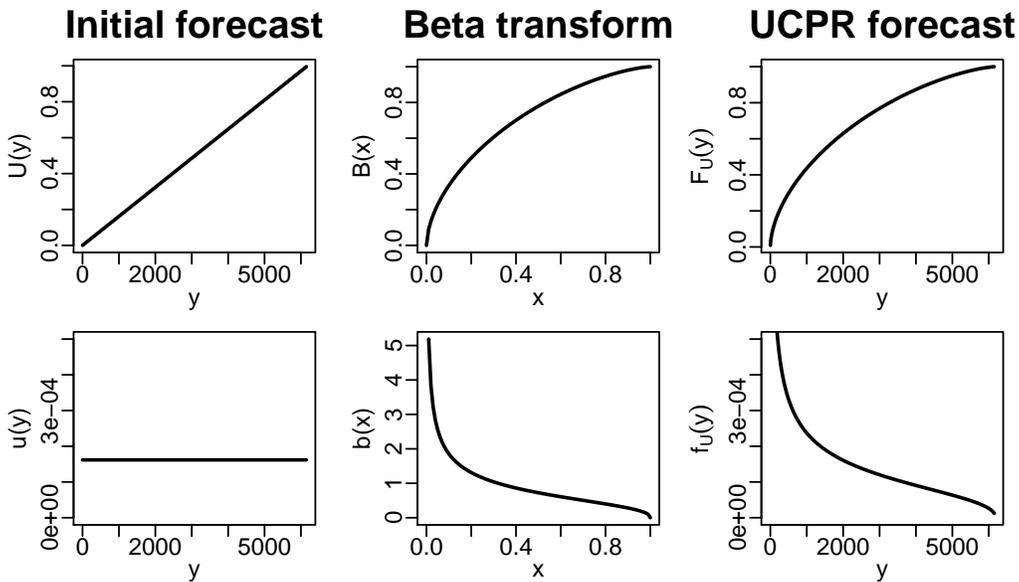


Figure 5.4: Illustration of how the UCPR is similar to the climatology itself when the model coefficients are estimated as in Figure 5.3.

exactly what the beta transformation function in Figure 5.4 does.

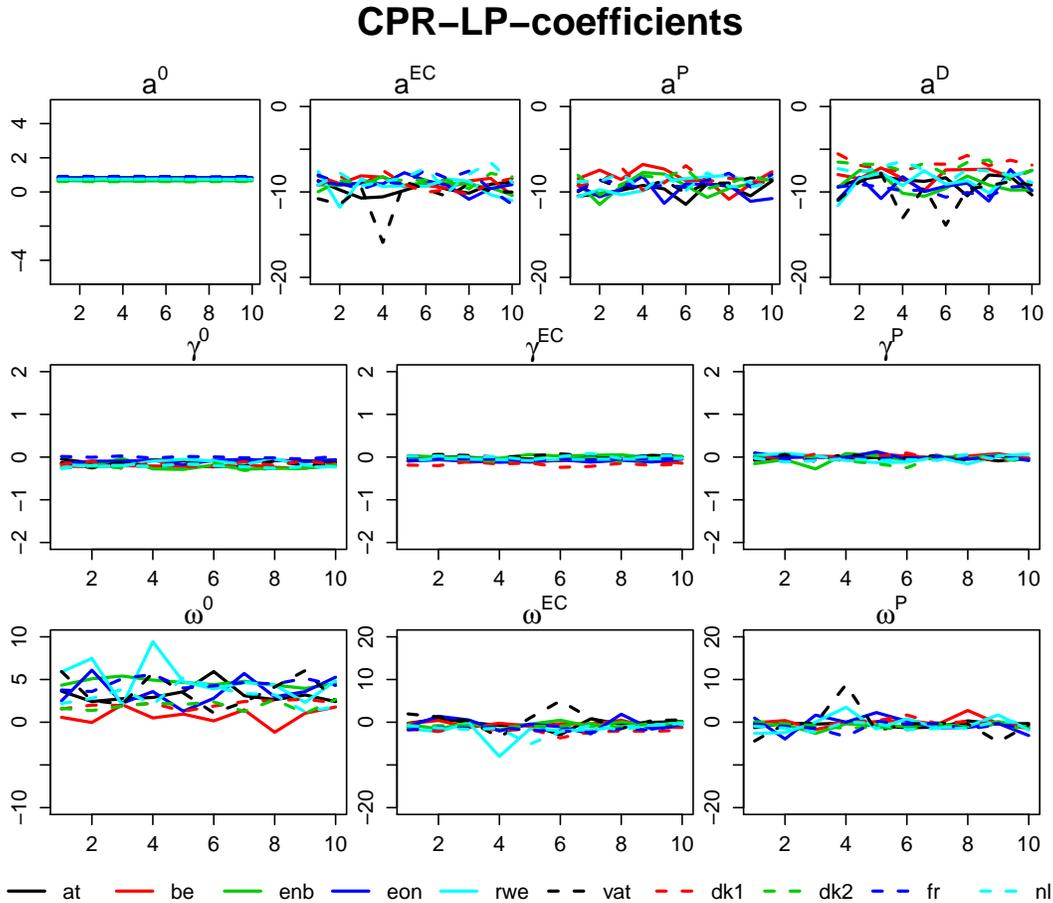


Figure 5.5: The coefficients of the CPR-LP model for all bidding areas when both the EC-forecast and the persistence forecast are synthetic random time series. Lead times in hours are on the x-axis.

### 5.6.3 CPR-LP coefficients: Synthetic Forecasts

The coefficients for the CPR-LP model are plotted in Figure 5.5. There are some clear similarities with the two other models. The precision coefficients  $a^0$ ,  $a^{EC}$ ,  $a^P$  and  $a^D$  found in the modelling of the precision parameter  $\nu$  in (5.9), are just the same as they were for CCPR and UCPR. The result is that  $\nu = 2$  for all lead times and bidding areas.

The mean coefficients  $\gamma^0$ ,  $\gamma^{EC}$  and  $\gamma^P$  in the modelling of the mean parameter  $\mu$  in (5.8), are all close to zero for all areas and lead times. This is similar to CCPR in Figure 5.1, which makes the mean parameter  $\mu = 0.5$ . These values for  $\nu$  and  $\mu$  make the beta transformation in (5.7) correspond to a uniform cdf, as it did in Section 5.6.1 for the CCPR model.

We move over to the coefficients  $\omega^0$ ,  $\omega^{EC}$  and  $\omega^P$ , which are used in (5.10) to model the linear

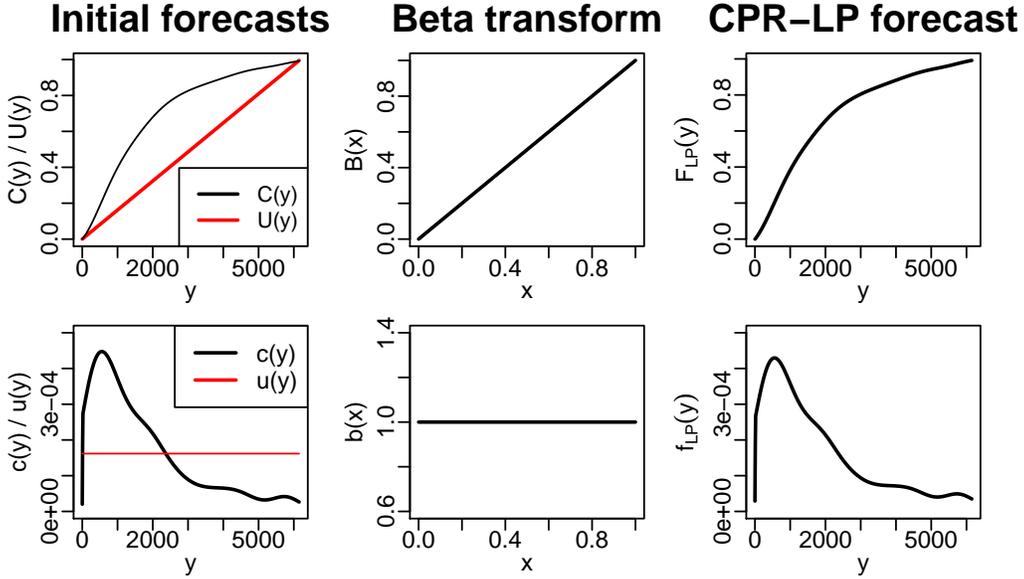


Figure 5.6: Illustration of how the CPR-LP is similar to the climatology itself when the model coefficients are estimated as in Figure 5.5.

predictor  $\nu$ . We hereafter refer to these coefficients as the weight coefficients. The predictor  $\nu$  is related to the weight parameter  $w$  through the logit function in (4.5), which maps the real number  $\nu$  onto the unit interval  $[0, 1]$ . The weight parameter  $w$  is found in the CPR-LP model definition (5.7), and is used to weigh between the climatology  $C(y)$  and the uniform distribution  $U(y)$  inside the beta transformation. If  $w = 1$ , the CPR-LP model is a beta transformation of only the climatology  $C(y)$ , and the model corresponds to the CCPR model in (5.1). If  $w = 0$ , the CPR-LP model is a beta transformation of only the uniform distribution  $U(y)$ , and the model corresponds to the UCPR model in (5.4).

The weight coefficients  $\omega^{EC}$  and  $\omega^P$  in Figure 5.5 are both equal to zero for all lead times and areas. Both the EC-forecast and the persistence forecast are therefore not giving any information to the model, as expected when these forecasts are random as in this example. We are left with the intercept  $\omega^0$  which ranges between approximately 0 and 5, varying with the bidding area. The weight parameter  $w$  is then ranging between 0.5 and 1, which means that the CPR-LP model is mostly similar or equal to the CCPR model. In particular, the model is an even mix between UCPR and CCPR when  $w = 0.5$ , and equal to CCPR when  $w = 1$ .

Figure 5.6 illustrates how these coefficient values affect the CPR-LP forecast, as cdf in the upper row, and as pdf in the lower row. The plots show the results for the bidding area RWE

in Germany, where  $\omega^0$  is close to 5 for all lead times. This gives  $w \approx 1$ , such that the CPR-LP corresponds to a CCPR model. As CCPR in Figure 5.2, the CPR-LP forecast is almost identical to the climatology itself.

## 5.7 Inference, Evaluation and Implementation

To fit and test our forecasting models the software R is used (R Core Team, 2016). R is a free software environment for statistical computing and graphics. Below follows a pseudo code for how the models were tested using the evaluation methods in Section 3.4.

```

for each bidding area do
  |
  | for each lead time l do
  | |
  | | for each hour h do
  | | |
  | | | k-fold cross validation to;
  | | | 1. fit model;
  | | | 2. derive probabilistic forecasts;
  | | | 3. compute CRPS and PIT-value;
  | | end
  | end
end

```

**Algorithm 1:** Algorithm for testing the forecasting model using the evaluation methods in Section 3.4. The data is divided into  $k = 10$  subsequent folds.

The k-fold cross validation is done as described in Section 3.4.4, where the models are fitted to each of the  $k$  training sets. Even though all four models in this chapter follow the scheme in Algorithm 1, the coefficient estimation in the ARn-SFR model is executed differently from the other three models. It simply uses the `lm`-function in R to carry out the linear regression, which returns maximum likelihood estimates of the model coefficients. The estimation procedure for the other three models is self implemented, where all coefficient estimates are maximum likelihood estimates with respect to the log-likelihood function in (3.5). The maximisation procedure is done using the `optim`-function in R. The log-likelihood function is multimodal, i.e. a function with many local maxima. To find the global maxima, we first run `optim` using a small test set of

our data, and put wide search intervals for each coefficient. The coefficient estimates from this optimisation are then used as initial coefficients in the final model fitting.

Let  $F_t(\cdot)$  be any of our probabilistic forecasts at valid time  $t$ , and  $y_t$  the actual production value at valid time  $t$ . The PIT-values are then computed as  $F_t(y_t)$  for all  $t$ , which gives the PIT-diagrams when plotted as a histogram. The CRPS is computed as defined in (3.7), and approximated with a Riemann sum.

# Chapter 6

## Results

In this chapter we present the results when testing the models from Chapter 5 on our case data from Chapter 2. Section 6.1 presents an analysis of the estimated coefficients in the three models CCPR, UCPR and CPR-LP. The ARn-SFR model is excluded from this section, but a corresponding analysis of its coefficients can be found in Malmgård (2016). Section 6.2 presents forecast performance of the models, first by investigating the shape of the probabilistic forecasts, and then with respect to sharpness and calibration.

### 6.1 Model Coefficients

#### 6.1.1 CCPR coefficients: Actual Forecasts

The fitted coefficient estimates for the CCPR model are plotted in Figure 6.1. We start with the mean coefficients  $\gamma^0$ ,  $\gamma^{EC}$  and  $\gamma^P$  from (5.2). First of all we notice the relationship between  $\gamma^{EC}$  and  $\gamma^P$ . As  $\gamma^{EC}$  increases,  $\gamma^P$  decreases accordingly. Based on the decreasing temporal dependency between actuals in Figure 2.4a,  $\gamma^P$  is expected to decrease for increasing lead time. For the shortest lead times, the temporal dependency between actuals is strong enough to make the persistence forecast a better forecast compared to the EC-forecast. However, the EC-forecast quickly becomes the most accurate forecast as the lead time increases, i.e. the corresponding coefficient  $\gamma^{EC}$  increases.

Bidding areas with high values for  $\gamma^{EC}$  tend to have low values for  $\gamma^P$ . This reflects the re-

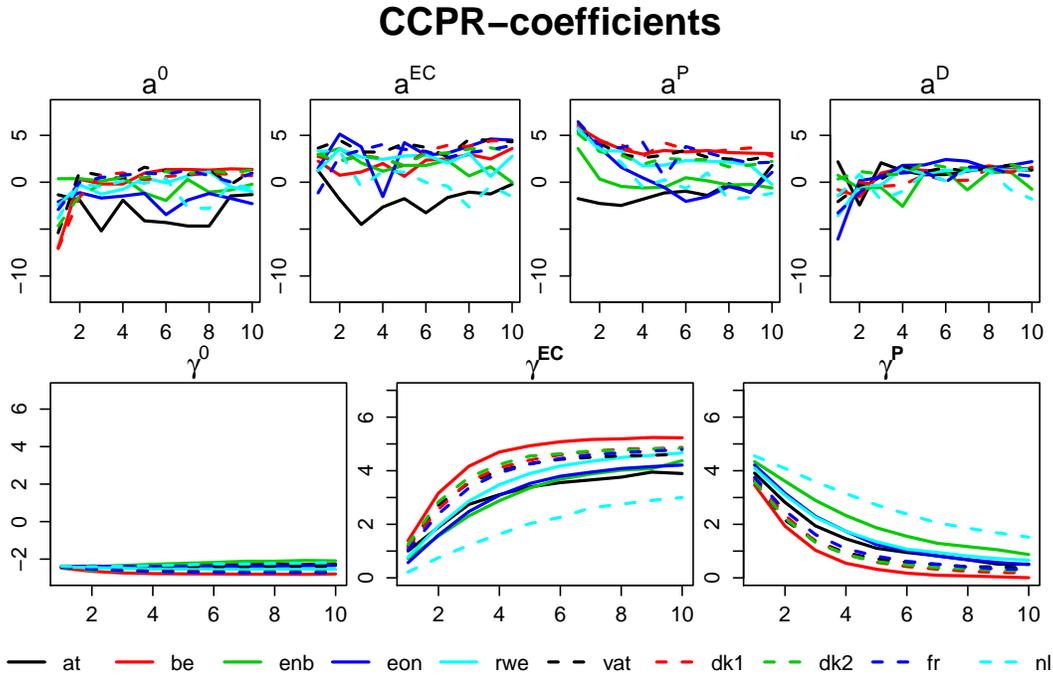


Figure 6.1: The coefficients of the CCPR model for all bidding areas with the EC-forecast and the persistence forecast as defined in Section 5.1. Lead times in hours are on the x-axis.

lation between the predictive performance of the EC-forecast and the persistence forecast. The Netherlands for example, has very low values for  $\gamma^{EC}$  and accordingly high values for  $\gamma^P$ . The relative predictive performance of the EC-forecast compared to the persistence forecast is therefore worse for the Netherlands compared to the other bidding areas. This is in agreement with what we observe in Figure 2.3, where the Netherlands has the poorest correlation between actuals and EC-forecasts.

The remaining coefficients in Figure 6.1,  $a^0$ ,  $a^{EC}$ ,  $a^P$  and  $a^D$ , are the precision coefficients, used to model the precision parameter  $\nu$  in (5.3). The modelling of  $\nu$  multiplies each term with the exponential of the precision coefficients. Therefore, a small change in a precision coefficient might result in a large change in  $\nu$ . The precision coefficients  $a^{EC}$  and  $a^P$  in Figure 6.1 are quite similar. These coefficients are related to the term in (B.2), which is plotted in Figure B.2. Because both  $a^{EC}$  and  $a^P$  are of substantial magnitude, both the EC-forecast and the persistence forecast are important to this part of the precision parameter model. However,  $a^P$  decreases for the first few lead times, which suggests that the persistence forecast is most important for shorter lead times. This is similar to what we have already seen for  $\gamma^P$ , which is also related to the persistence

forecast.

The two coefficients  $a^0$  and  $a^D$  are of smaller magnitude compared to  $a^{EC}$  and  $a^P$ .  $a^0$  is the intercept of the precision model in (5.3), and the term related to  $a^D$  models the precision as a function of the absolute difference between the EC-forecast and the persistence forecast. Because  $a^D$  is non-negative, but centred around zero, the absolute difference between the two deterministic forecasts is somewhat important in the modelling of the precision parameter. However, it is far from the magnitude of  $a^{EC}$  and  $a^P$ , especially when we take into account the exponential of these coefficients, which is used in the precision modelling (5.3).

As a final comment to the coefficients of the CCPR model in Figure 6.1, we want to point out Austria's values for the precision coefficients. All precision coefficients, except for  $a^D$ , are small for Austria compared to the other bidding areas. We know from Section 2.2 and Figures 2.3 and 2.4a that Austria is a difficult area to forecast. This fact is reflected through the small precision coefficients of Austria, which lead to small precision, i.e. large variance.

### 6.1.2 UCPR coefficients: Actual Forecasts

The fitted coefficient estimates are plotted in Figure 6.2. We start with the mean coefficients  $\gamma^0$ ,  $\gamma^{EC}$  and  $\gamma^P$  from (5.5). As for the CCPR model in Figure 6.1,  $\gamma^{EC}$  increases while  $\gamma^P$  decreases for increasing lead time. With the exception of EON in Germany, the relation between the two coefficients for each bidding area is still the same. That is, areas with high values for  $\gamma^{EC}$  tend to have small values for  $\gamma^P$ .

The precision coefficients (from Equation (5.6)) of UCPR in Figure 6.2 are very similar to the precision coefficients of the CCPR model in Figure 6.1. Austria still has lower values for  $a^0$ ,  $a^{EC}$  and  $a^P$  compared to the other bidding areas, and we still observe the decrease in  $a^P$  for the first shorter lead times.

### 6.1.3 CPR-LP coefficients: Actual Forecasts

Figure 6.3 displays the fitted CPR-LP coefficient estimates. We start with the mean coefficients  $\gamma^0$ ,  $\gamma^{EC}$  and  $\gamma^P$ , which are used in the modelling of the mean parameter  $\mu$  through the linear predictor  $\eta$  in (5.8). These coefficients are very similar to what we saw for both CCPR and UCPR

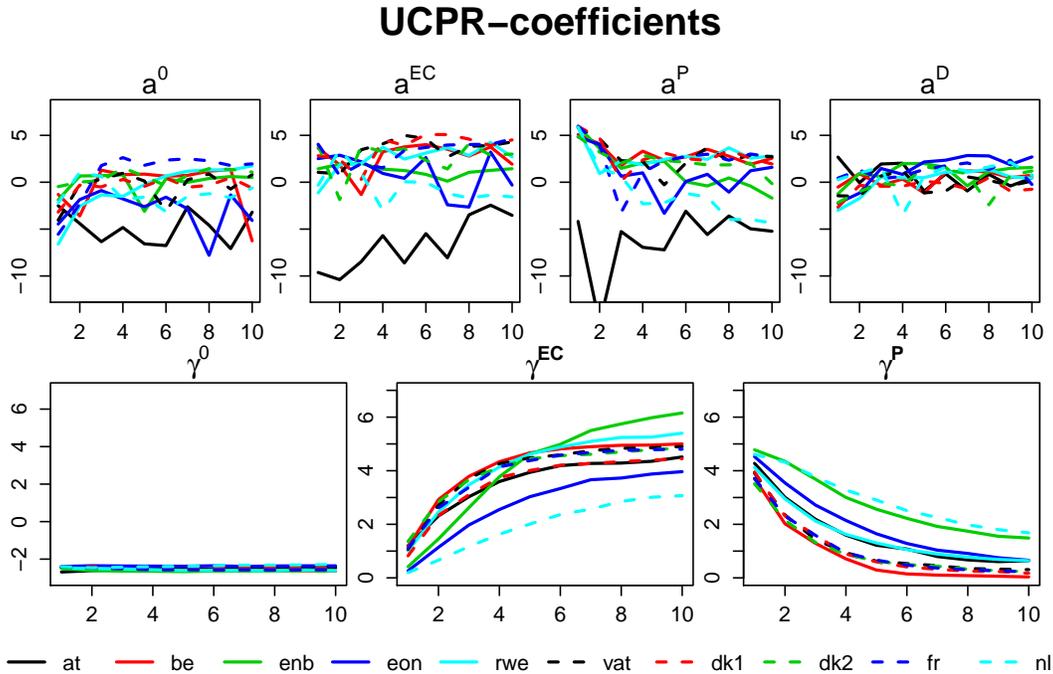


Figure 6.2: The coefficients of the UCPR model for all bidding areas with the EC-forecast and the persistence forecast as defined in Section 5.1. Lead times in hours are on the x-axis.

in Figures 6.1 and 6.2 respectively. The coefficient  $\gamma^{EC}$  is increasing while  $\gamma^P$  is decreasing for increasing lead time. In addition, areas with high values for  $\gamma^{EC}$  tend to have low values for  $\gamma^P$ . We refer to Section 6.1.1 for a detailed evaluation of these observations, as it also applies to this model.

We continue with the precision coefficients  $a^0$ ,  $a^{EC}$ ,  $a^P$  and  $a^D$  from (5.9), used to model the precision parameter. These coefficients are different from what we saw for CCPR and UCPR in Figures 6.1 and 6.2, respectively. The coefficient  $a^{EC}$  is below -8 at all lead times for all areas, which is much lower than it was for CCPR and UCPR. The term belonging to this coefficient in Equation (5.9) is therefore neglected in the modelling of the precision parameter (because we use the exponential of the coefficient in the model). The other 3 precision coefficients  $a^0$ ,  $a^P$  and  $a^D$  are therefore increased compared to CCPR and UCPR, to maintain the level of precision. In particular,  $a^D$  has increased compared CCPR and UCPR. This suggests that low absolute difference between EC-forecast and persistence forecast is more important for the CPR-LP forecast precision. In addition,  $a^P$  decreases as lead times increase, which suggests that the persistence forecast is most important for shorter lead times.

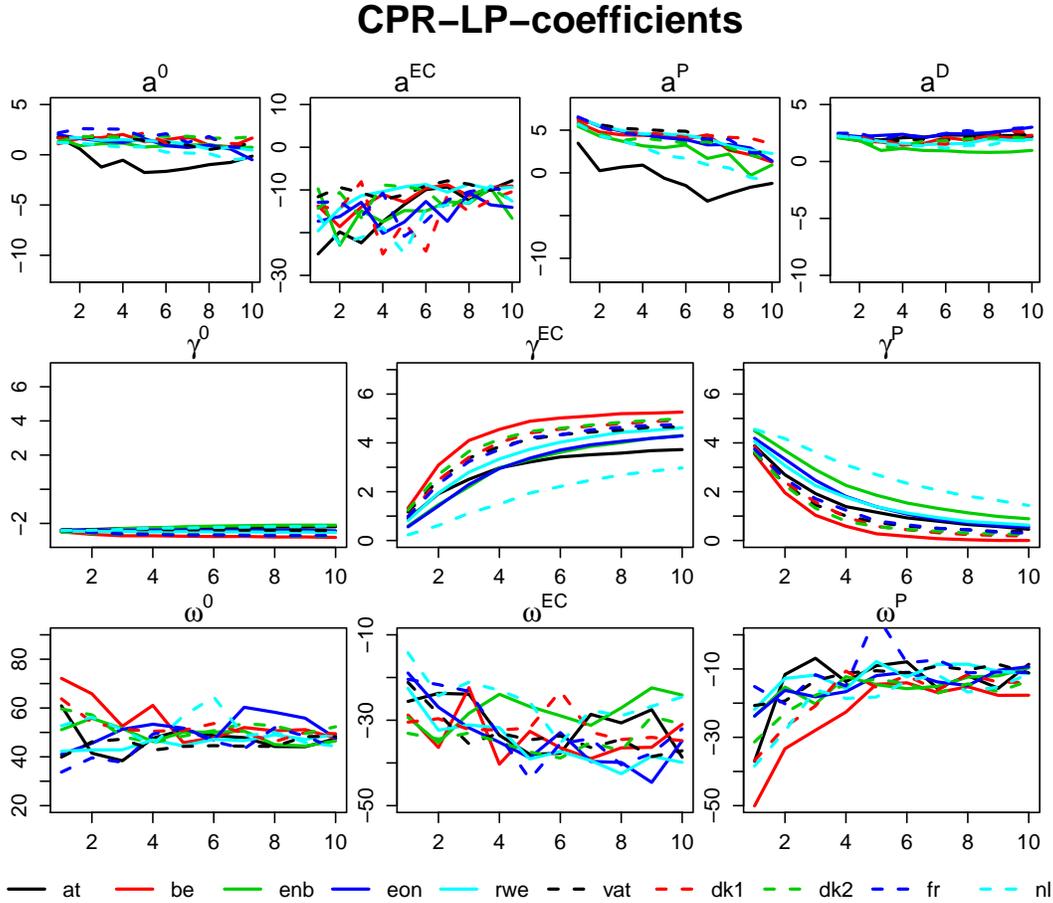


Figure 6.3: The coefficients of the CPR-LP model for all bidding areas with the EC-forecast and the persistence forecast as defined in Section 5.1. Lead times in hours are on the x-axis.

We move over to the weight coefficients  $\omega^0$ ,  $\omega^{EC}$  and  $\omega^P$  which are used in (5.10) to model the linear predictor  $\nu$ . This predictor is related to the weight parameter  $w$  through the logit function in (4.5). The intercept  $\omega^0$  is always positive, but with various magnitudes depending on the bidding area, centred around 50. This implies that the weight parameter is close to 1 when both the EC-forecast and the persistence forecast are zero, and makes the CPR-LP model in (5.7) correspond to a CCPR model.

The weight coefficients  $\omega^{EC}$  and  $\omega^P$  on the other hand, are negative, and makes the weight parameter  $w$  decrease for increasing lead time. When the weight parameter decreases and approaches zero, the CPR-LP model in (5.7) becomes more similar to the UCPR model.  $\omega^P$  is more negative than  $\omega^{EC}$  for the shortest lead times, which suggests that the persistence forecast is more important compared to the EC-forecast when it comes to lowering the weight parame-

ter for the shortest lead times. This is in accordance with all other coefficients related to the persistence forecast, which all suggest that this forecast is less accurate for longer lead times.

To summarize, the patterns in the estimated weight coefficients behave as expected. Because  $\omega^0$  is positive, and  $\omega^{EC}$  and  $\omega^P$  are negative, the weight parameter  $w$  starts at 1, and decreases towards zero for increasing EC-forecast and persistence forecast. Hence, the CPR-LP model in (5.7) starts like a CCPR model for low forecasts, and becomes more like a UCPR model for higher forecasts. This form is exactly what we intended when we constructed the model in Section 4.4.

## 6.2 Model Evaluation

The forecasting methodology we introduce in this thesis, is a response to weaknesses of the forecasting methods presented by Malmgård (2016). The best performing forecasting model from Malmgård (2016) was the ARn-SFR-model, which is described in Section 5.5. The results for this model are included in this section, and the model is now considered as a reference forecast for the other three models. When testing the forecasting models, we see similar results for all bidding areas. Therefore, most of the results are only described for one area, RWE in Germany.

### 6.2.1 Probabilistic Forecasts

Deterministic forecasts near zero will typically lead to probabilistic forecasts also centred near zero. This property is shared by all of the four forecasting models. However, the shape of the four different probabilistic forecasts might differ. One model might have the sharpest forecast for deterministic forecasts near zero, while still having the most dispersive forecast for deterministic forecasts near the maximum production capacity.

Figure 6.4 displays all four probabilistic forecasts together with the corresponding deterministic forecasts and actual value, for three different forecast valid times (one per column) and four different lead times (one per row). The valid times are referred to by their date, and are picked deliberately to plot three forecasts covering three different areas over the domain. Even though we only see forecasts at three individual valid times, they represent the general differences between each model well.

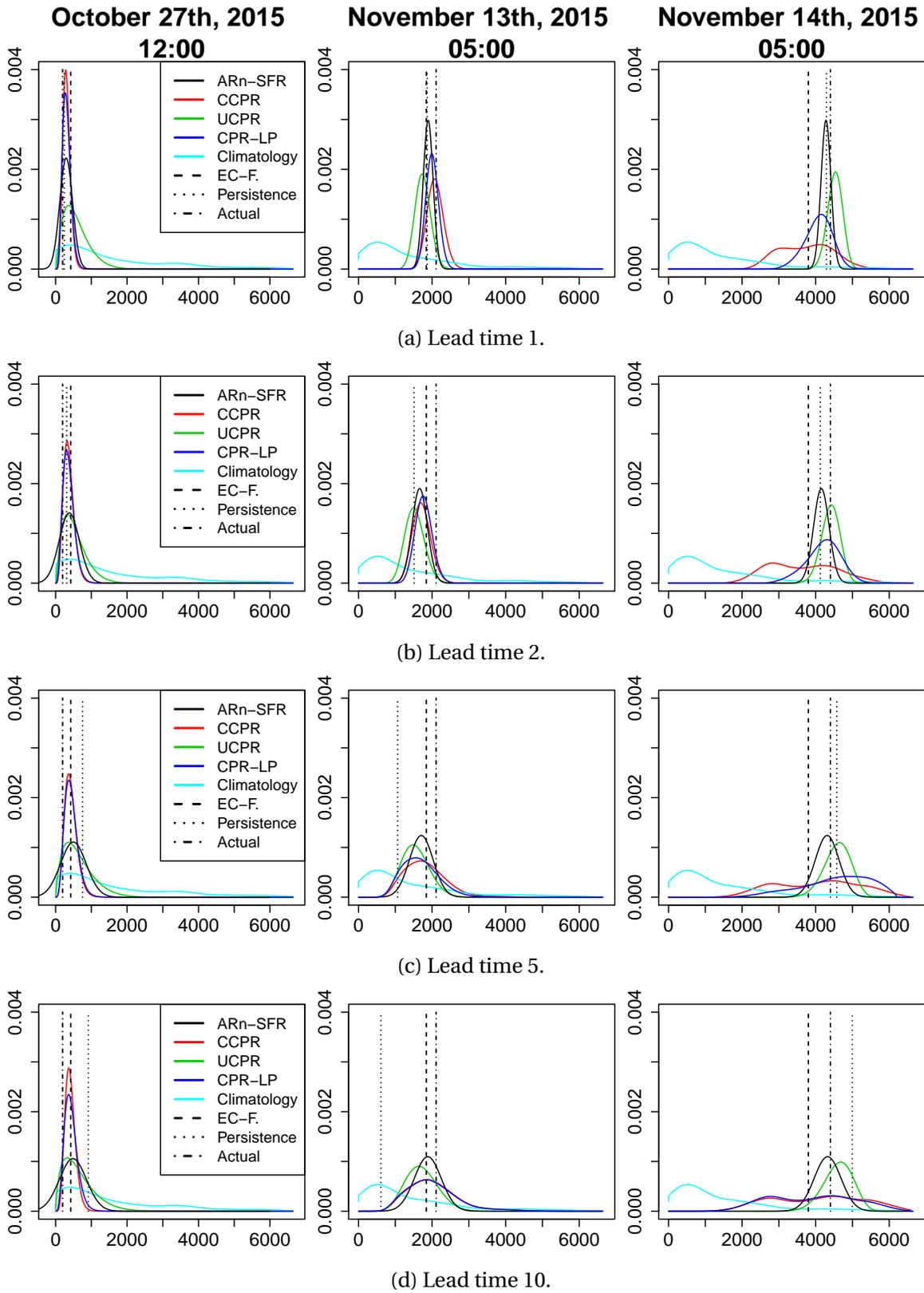


Figure 6.4: Probabilistic forecasts from all four models for RWE in Germany at three different forecast valid times and four different lead times: Lead time 1 (a), lead time 2 (b), lead time 5 (c) and lead time 10 (d). In addition to the probabilistic forecasts, each plot includes the corresponding deterministic forecasts and actual value. Density on the y-axis and MW on the x-axis.

The probabilistic forecasts from CCPR and CPR-LP on October 27th are quite similar. This is expected, as the CPR-LP is weighed towards the CCPR for small deterministic forecasts. Both are also sharper compared to the other two forecasts at this valid time, because they are based on the climatology  $C(y)$ , and not the uniform distribution  $U(y)$  like the UCPR is. Moving on to November 13th, the CCPR and CPR-LP forecasts have decreased sharpness and are quite similar to the other two models. The exception is at lead time 1, where ARn-SFR is the sharpest forecast for both November 13th and November 14th. The CCPR and CPR-LP forecasts, on the other hand, are rather dispersive on November 14th for all lead times. The sharpness of the UCPR forecast is comparable to ARn-SFR at this valid time.

With increasing lead time we see a clear tendency of more dispersive forecasts, for all models and valid times. The change is most prominent for the shortest lead times. Compared to the transition from lead time 1 to lead time 2, the transition from lead time 5 to lead time 10 sees almost no notable changes to the shape of the forecasts. Increasing the lead time also makes the CPR-LP forecast more similar to the CCPR forecasts, even for larger deterministic forecasts. The CPR-LP model is designed to weigh the UCPR forecast for large deterministic forecasts, but this is seemingly most desirable for short lead times.

Before ending this section, we want to take a closer look at the ARn-SFR forecast on October 27th for all lead times. We mentioned in Chapter 1 how the ARn-SFR forecasts are not always physically valid, i.e. below zero production or above maximum production. This problem is visualized in Figure 6.4 where the ARn-SFR forecast has some of its density below zero.

### 6.2.2 Sharpness

The CRPS for all four models is plotted in Figure 6.5. The corresponding plots for the other nine bidding areas are found in Appendix C, Figure C.1. Of the four models, the CCPR forecast has the highest CRPS values for all lead times, followed by the UCPR forecast. The best performing model with respect to CRPS is the ARn-SFR model, where the CPR-LP is the second best, only slightly worse.

We know from Chapter 4 that a model might perform well when forecasting low values, and worse for high values, or vice versa. We want to investigate how this is reflected through the CRPS. Figure 6.6 is similar to Figure 6.5, only now we have divided the data into three groups,

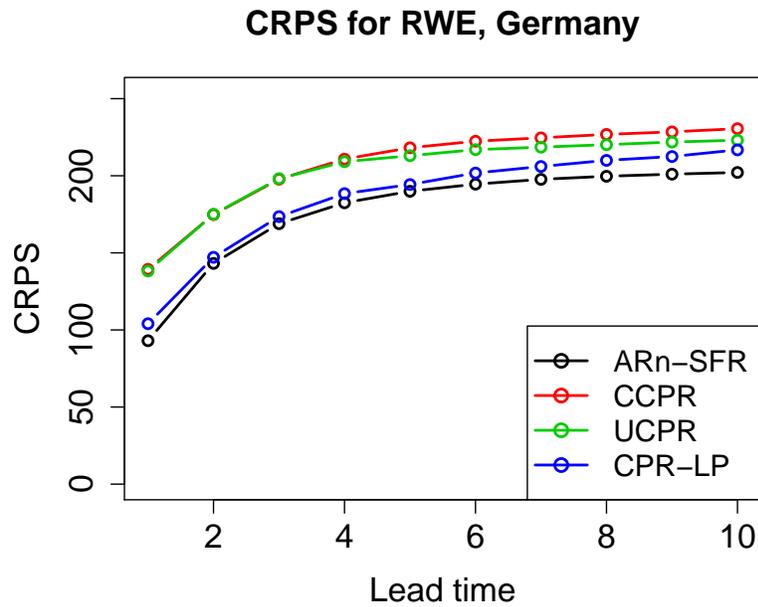


Figure 6.5: CRPS for the bidding area RWE in Germany.

with respect to the deterministic EC-forecasts. The first plot includes the lowest 60% of all EC-forecasts, and is called group low. The third plot includes the highest 10% of all EC-forecasts, and is called group high. The remaining 30% of the data is plotted in the middle, and is called group intermediate. Note that the lower 60% of the EC-forecasts in group low only spans between 0 MW and 1290 MW, where the upper 10% of the EC-forecasts in group high spans between 3289 MW and 6459 MW, i.e. a much wider domain. This clearly shows we are forecasting low production values most of the time.

We know from Section 4.3 that the UCPR model struggles to attain sharp probabilistic forecasts for small deterministic forecasts. This is why the CRPS of the UCPR is much larger compared to the other forecasts in group low. The CCPR, CPR-LP and ARn-SFR forecasts have very similar performance in group low. The CRPS has increased for all models in group intermediate. The UCPR is no longer an outlier, but is more similar in performance with the other models. Figures C.4, C.5 and C.6 show that when we consider all bidding areas at the same time, the ARn-SFR is the overall best performing forecast with respect to CRPS in group intermediate. The order of the remaining three forecasts varies with lead time and between areas. Group high clearly illustrates the weakness of the CCPR model, as it attains much higher CRPS values than

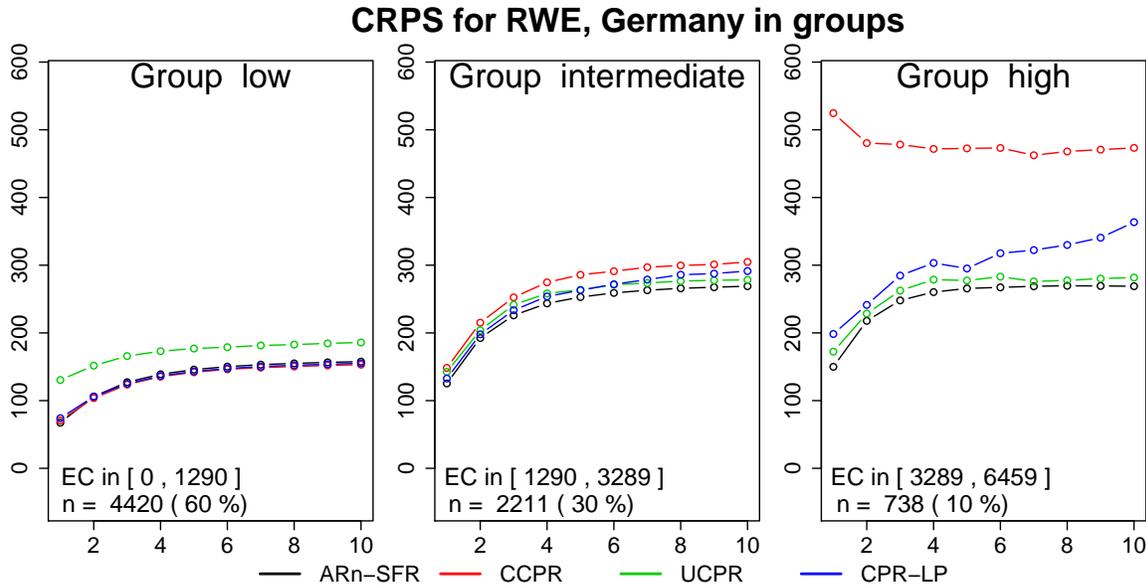


Figure 6.6: CRPS for the bidding area RWE in Germany divided into three groups with respect to different values of EC-forecasts. CRPS on the y-axis and lead time on the x-axis. The minimum and maximum EC-forecast in MW are shown in the brackets, and  $n$  is the number of data points belonging to this group.

the other models. This is the result of the difficulties illustrated in Section 4.2, where the model fails to attain sharp forecasts in the upper part of the domain.

Further, we want to investigate the width of the 95% prediction intervals. We continue to use the same division of data based on different EC-forecasts. A box plot of the prediction interval widths for the different models and lead times is given in Figure 6.7.

As we observed in Figure 6.4, all models perform worse as lead times increase. That is, all models have increasing 95% interval widths. Moving from group low up to group high, the ARn-SFR and UCPR models are barely changing. This is in contrast to CCPR and CPR-LP, which have increasing 95% interval widths. However, the widths of CCPR and CPR-LP are shorter compared to the two other models for group low.

### 6.2.3 Calibration

The PIT-diagrams for each model are plotted for lead times 1, 2, 5 and 10 in Figure 6.8. We still use RWE in Germany as an example representing the results for all bidding areas. The PIT-diagrams for the remaining nine areas are found in Appendix C, Figures C.2 and C.3.

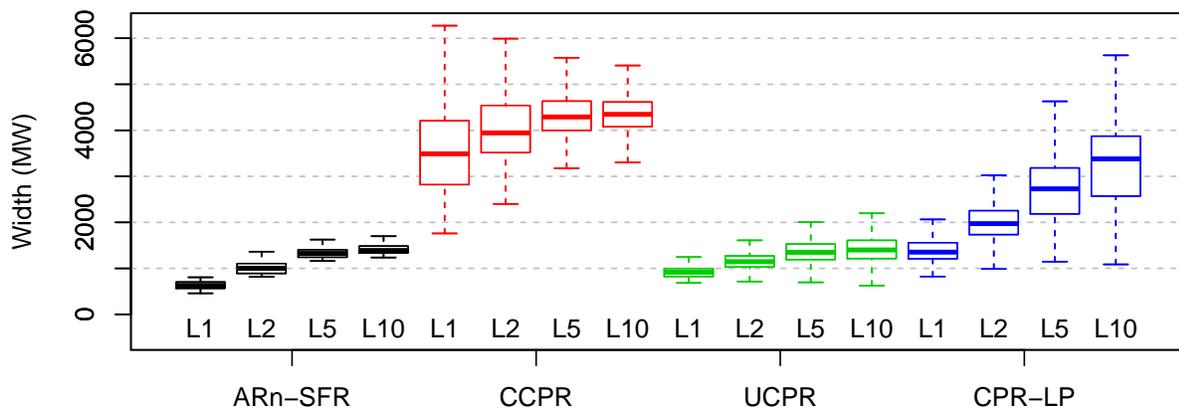
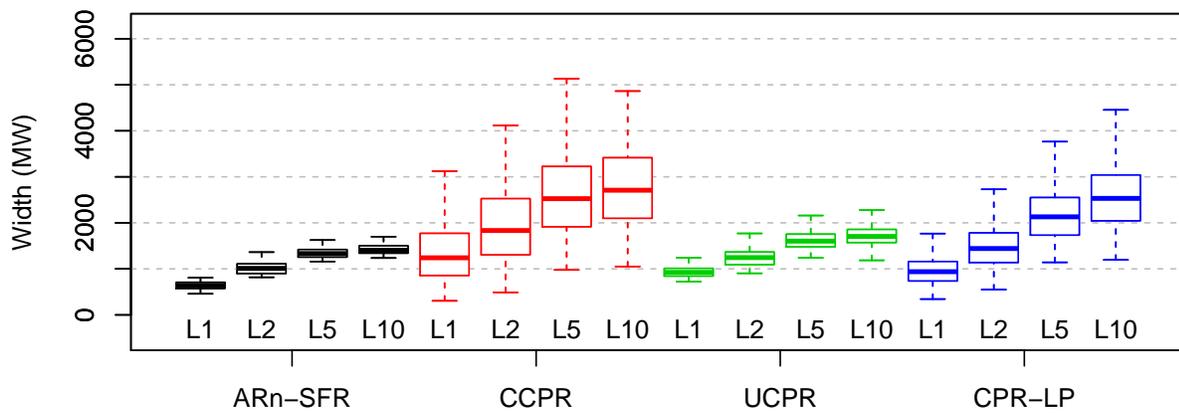
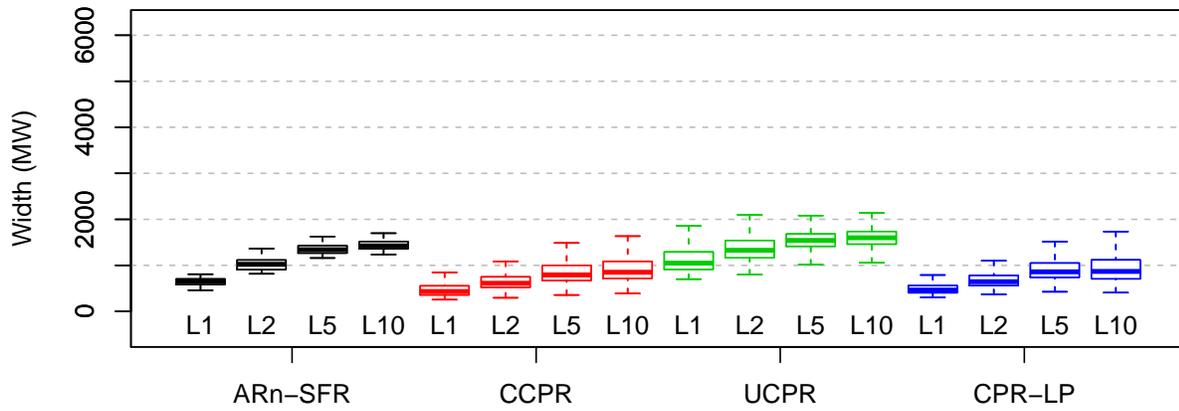


Figure 6.7: Box plots of the 95% prediction interval widths for RWE in Germany.

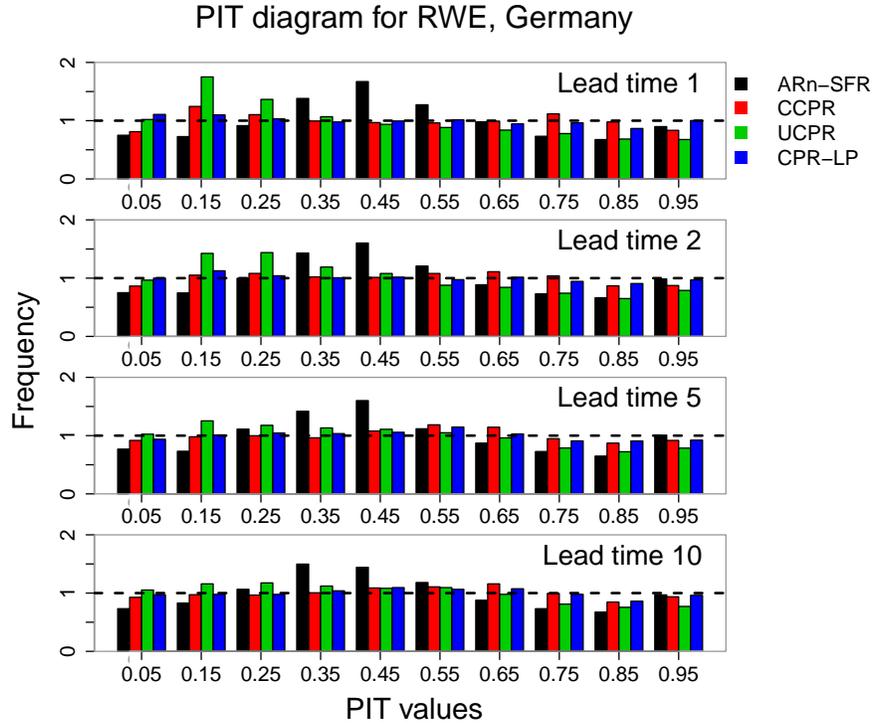


Figure 6.8: PIT-diagram for the bidding area RWE in Germany.

Comparing the PIT-diagram of CPR-LP with the one of ARn-SFR in Figure 6.8, we see a great improvement in calibration. The diagram for CPR-LP is close to a uniform distribution for all lead times. The same can be said about the CCPR model which is quite similar to CPR-LP. The ARn-SFR on the other hand shows tendencies to being overdispersed, but the frequencies are quite satisfactory for higher PIT-values. The diagram of the UCPR is also overdispersed, but with a bias, especially for lower lead times.

Figure 6.9 also plots the RWE, Germany PIT-diagrams for each model at lead times 1, 2, 5 and 10, but the data is now divided into the three groups we also investigated in Section 6.2.2. Starting with group low, we notice that ARn-SFR and UCPR are even more overdispersed for all lead times, compared to the diagrams in Figure 6.8. The 27th of October in Figure 6.4 displays that these two forecasts are wider compared to the other two forecasts for group low. According to the PIT-diagrams, these probabilistic forecasts are in fact too wide. The CPR-LP is arguably the best performing model for group low. However, it does have a slight U-shape for lead time 10, which is an indication of underdispersion.

Moving on to group intermediate we note that most models seem more calibrated compared

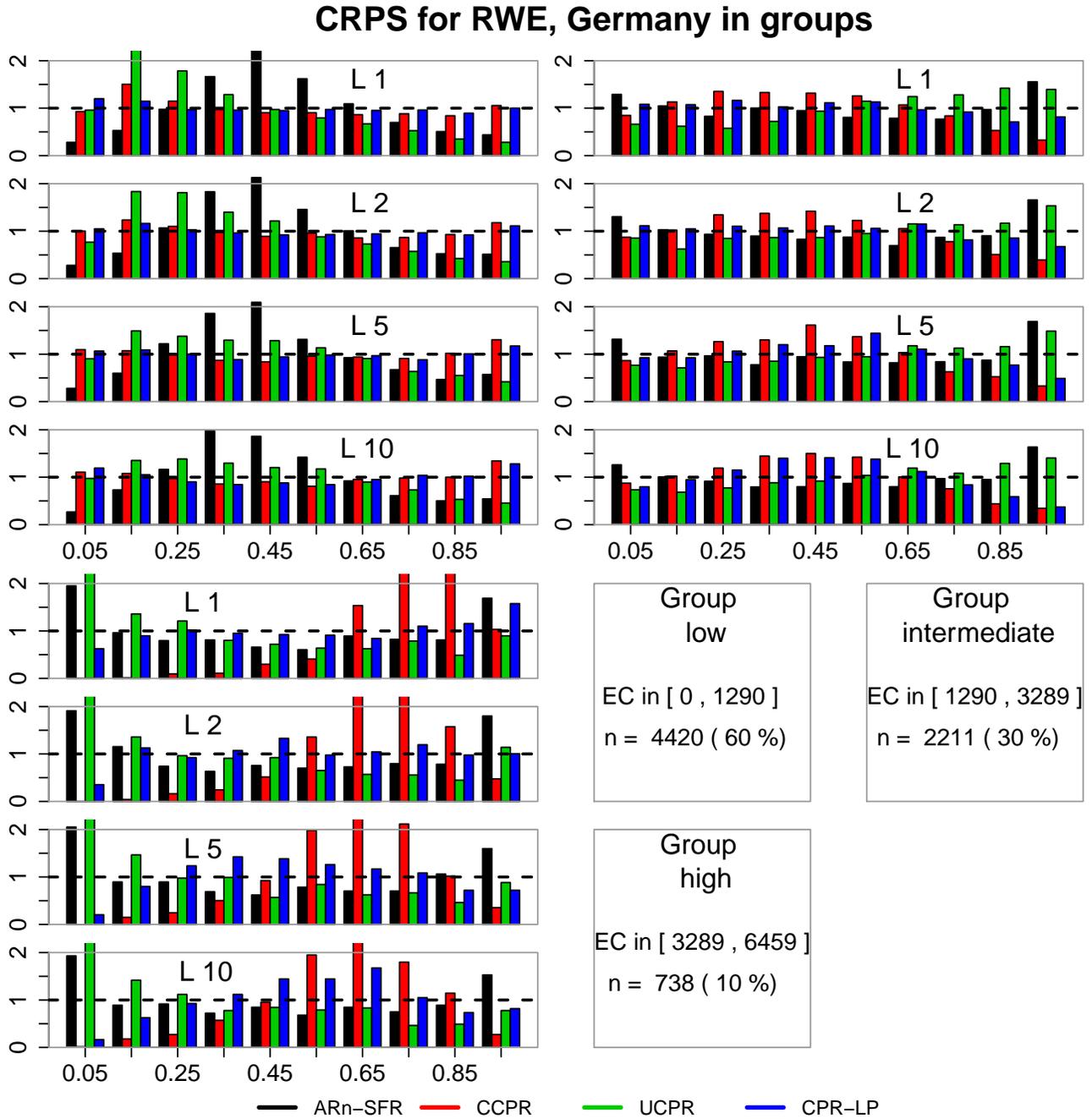


Figure 6.9: PIT-diagrams for the bidding area RWE in Germany divided into three groups with respect to different values of EC-forecasts. Frequency on the y-axis and PIT-values on the x-axis. The lower right corner displays information about the three groups plotted, including the EC interval limits in MW, and the number of data points in each interval.

to group low. The UCPR diagram is now biased, with too high frequencies to the right, whereas the CCPR model on the other hand is overdispersed. The CPR-LP model is performing quite well for all lead times, but is also somewhat overdispersed, especially for lead time 10. The ARn-SFR forecast is slightly underdispersed for all lead times. This shape carries over to group high, where the underdispersion is even more prominent. Figure 6.4 displays that the ARn-SFR model attains the sharpest forecasts of all models on November 13th and November 14th, especially for lead time 1. However, based on the PIT-diagrams, these forecasts are too sharp on average.

Considering all PIT-diagrams for group high in Figure 6.9 as a whole reveals that all models struggle to forecast high production values. The PIT-diagram of CCPR is clearly biased with too high frequencies to the right, especially for the first two lead times. The UCPR is biased in the opposite direction, but not as substantial as CCPR. Yet again, the CPR-LP is arguably the most calibrated diagram, even though it is far from uniformly distributed.

# Chapter 7

## Discussion and Conclusion

The main research task of this thesis is to predict wind power production in the bidding areas depicted in Figure 2.1, based on the production data itself, and the forecast referred to as the EC-forecast. The predictions should be given as probabilistic forecasts, where the goal is to maximise their sharpness subject to calibration. [Malmgård \(2016\)](#) already developed probabilistic forecasts with the exact same research task in mind, thus using the same case data. The forecasting models tested by [Malmgård \(2016\)](#) were based on the well known method of linear regression, accompanied by autoregressive models originating from the temporal dependencies in the wind power production. However, these forecasts had problems related to calibration, and were not always physically valid. That is, the models could forecast impossible production values, like negative production. In this thesis, we introduce a new forecasting methodology, which is a response to these weaknesses. The best performing model tested by [Malmgård \(2016\)](#) was the ARn-SFR model, which is used as a reference for model performance.

The new forecasting methodology is inspired by the CCPR method from [Borhaug \(2014\)](#), which in turn is based on the beta transformed linear pool presented by [Gneiting and Ranjan \(2013\)](#). The methodology starts out with an initial probabilistic forecast, in the form of a cdf. In the case of the CCPR, this forecast is the climatology, i.e. a sample cdf of wind power production history. We proceed by transforming the initial forecast using a beta cdf. That is, the beta transform takes a cdf as input, and returns a new cdf, which is our probabilistic forecast. The beta transform is parameterized with mean parameter  $\mu$  and precision parameter  $\nu$ . These parameters, and thus the shape of the beta transform, are modelled as a function of the EC-forecast and

previous observations, i.e. a persistence forecast.

Chapter 4 includes a more thorough introduction to the CCPR model, as well as the new UCPR and CPR-LP models. The two latter models are developed as a response to the CCPR model performance for the case study. As we saw in Section 4.2, the CCPR is unable to attain sharp forecasts when the deterministic forecasts are large, i.e. when forecasting on the upper area of the domain. The UCPR model introduced in Section 4.3, tries to solve this problem by substituting the climatology in the CCPR model by a uniform distribution (over the domain of possible production values). However, this sacrifices the sharp forecasts obtained by the CCPR for small deterministic forecasts, i.e. when forecasting on the lower part of the domain. The solution is the CPR-LP model introduced in Section 4.4, which is designed to keep the benefits of the two other models, while still eliminating the disadvantages. This is done by beta transforming a weighted sum of the climatology and the uniform distribution, where the weight parameter is a function of the deterministic forecasts. This gives the CPR-LP model the ability to mimic the CCPR model for small deterministic forecasts, and to mimic the UCPR model for large deterministic forecasts.

The models are fitted and tested using k-fold cross validation as explained in Section 3.4.4. The fitting sequence estimates the model coefficients presented in Chapter 5 for each area, lead time, and hour of the day. Section 6.1 investigates the behaviour of these coefficients and reveals that all coefficients related to the persistence forecast tend to decrease with increasing lead time. This behaviour is in agreement with the observation in Figure 2.4a, where the actual wind power production has decreasing temporal dependency for increasing lead time. The testing sequence computes the CRPS and PIT-values as defined in Section 3.4.3 and 3.4.2, respectively. The CRPS is a quantity that relates to both forecast sharpness and calibration, and the CRPS results for all models are presented in Section 6.2.2. Figure 6.5 reveals that the CPR-LP performs better than both the CCPR and the UCPR with respect to CRPS, but is slightly worse compared to ARn-SFR. However, if we include the CRPS results for all bidding areas in our analysis, plotted in Figure C.1, the CPR-LP is in fact the best performing model for certain bidding areas.

The PIT-diagram is used to test the forecasting models for calibration, and the results are analysed in Section 6.2.3. The PIT-diagram of the ARn-SFR model in Figure 6.8 is overdispersed, and is arguably the least calibrated of all models. This is one of the major problems with this

forecasting model, and an important reason why we have chosen the new methodology in this thesis. The PIT-diagrams for the CPR-LP and CCPR models on the other hand, are close to uniform distributions for all lead times. Another problem with the ARn-SFR model is that it is not always physically valid, i.e. it may forecast production values which are not physically possible to obtain. The new methodology behind the CCPR, UCPR and CPR-LP forecasts omits this weakness.

To conclude, we argue that the main research task of predicting the wind power production in various bidding areas in Northern Europe is fulfilled. We have developed a new forecasting methodology, where the best performing forecast is arguably the CPR-LP forecast. It is well calibrated, physically valid, and competing with the ARn-SFR with respect to CRPS performance. In fact, it is actually better than ARn-SFR with respect to CRPS for certain bidding areas. However, there are still alternatives and extensions to the CPR-LP model that should be investigated further.

Recall the definition of the term *initial forecasts*, which is a common term for the climatology and the uniform distribution, i.e. they are both initial forecasts. The CCPR model uses the same initial forecast, the climatology, in both the beta transformation, and to model the beta parameters. The same goes for UCPR, but with the uniform distribution as initial forecast instead of the climatology. This choice leads to best forecasting performance. The CPR-LP on the other hand is a beta transformation of the climatology and the uniform distribution summed together, and it is not clear whether or not we should use the climatology or the uniform distribution to model the beta parameters. An interesting alternative is to instead introduce a beta mixture ([Bassetti et al., 2015](#)). We would then get a weighted sum of two separate beta transformations, where one transforms the climatology and the other transforms the uniform distribution. This leads to two sets of beta parameters, where each set could be modelled using the initial forecast corresponding to the initial forecast in the beta transformation.

Section 4.2 explained why the CCPR model is unable to attain sharp probabilistic forecasts for the upper part of the domain, i.e. for large deterministic forecasts. In short, the reason behind this is the shape of the climatology which has a diminishing density in this area. The UCPR-LP model solves the problem by introducing a combination of both the climatology and the uniform distribution. It beta transforms a weighted sum of these two initial forecasts. How-

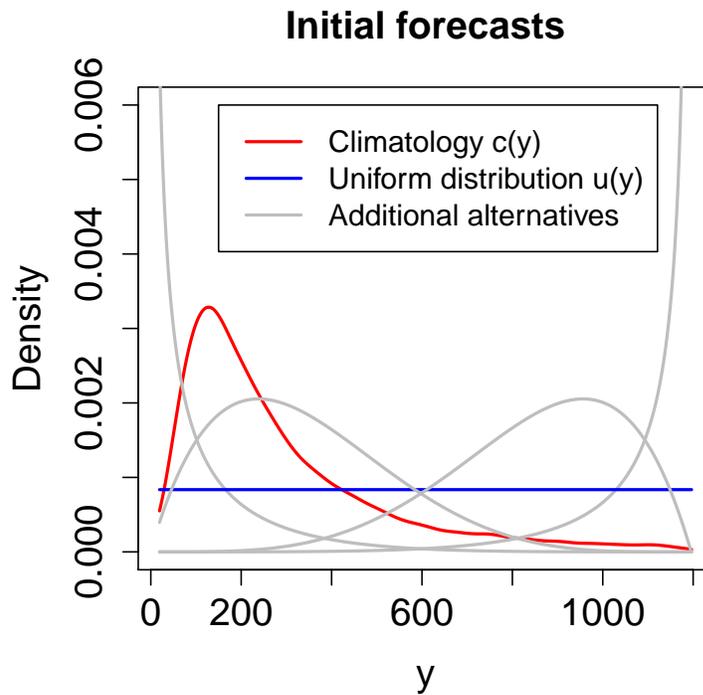


Figure 7.1: Example of a climatology  $c(y)$  and a uniform distribution  $u(y)$  together with other alternative initial forecasts.

ever, the method is not limited to including only two initial forecast. An extended version of the CPR-LP model could include a weighted sum of several initial forecasts in the beta transform. Figure 7.1 illustrates an example where the climatology and the uniform distribution are plotted together with additional initial forecasts. Further research should investigate how to choose the number of initial forecasts and their shape, depending on the climatology of the problem in question.

# Appendix A

## Actual Production Sources

**Austria** APG: <https://www.apg.at/en/market/Markttransparenz/generation/Erzeugung%20pro%20Typ>

**Belgium** Elia: <http://www.elia.be/en/grid-data/power-generation/wind-power>

**Denmark** NordPool: <http://www.nordpoolspot.com/Market-data1/Power-system-data/Production1/Wind-Power/ALL/Hourly1/?view=table>

**Germany** EEX: <https://www.eex-transparency.com/homepage/power/germany/production/usage/solar-wind-power-production/solar-wind-power-production-table>

**France** EEX: [http://clients.rte-france.com/lang/an/clients\\_distributeurs/vie/prod/realisation\\_production.jsp](http://clients.rte-france.com/lang/an/clients_distributeurs/vie/prod/realisation_production.jsp)

**Netherlands** EEX: <https://transparency.entsoe.eu/generation/r2/actualGenerationPerProductionType/show>

# Appendix B

## Interpretation of $r(\cdot)$

The function  $r(\cdot)$  is used to model the precision parameter  $\nu$  of the CCPR, UCPR and CPR-LP models, and is given as

$$\begin{aligned} \nu_{t,l} = r(\mathbf{x}_{t,l}^{EC,P}) = & \exp(a_{h,l}^0) + \exp(a_{h,l}^{EC})[G_h(x_t^{EC})(1 - G_h(x_t^{EC}))] \\ & + \exp(a_{h,l}^P)[G_h(x_{t,l}^P)(1 - G_h(x_{t,l}^P))] + \exp(a_{h,l}^D)[1 - |G_h(x_t^{EC}) - G_h(x_{t,l}^P)|], \quad (\text{B.1}) \end{aligned}$$

where  $G_h(\cdot)$  is either the climatology  $C_h(\cdot)$  or the uniform distribution  $U_h(\cdot)$ . Let us first convince ourselves that  $\nu_{t,l} = r(\mathbf{x}_{t,l}^{EC,P}) > 0, \forall \mathbf{x}_{t,l}^{EC,P}$  as this is required for  $\nu_{t,l}$ . The expression (B.1) consists of four terms in total. All exponentials are always positive, meaning the the first term is positive. Because  $G_h(\cdot)$  is a cdf, the range of the function is between 0 and 1. The second and third term of (B.1) is thus an exponential multiplied by two functional values between 0 and 1, which gives a non-negative value. The fourth term subtracts a functional value between -1 and 1 from 1, giving a value between 0 and 2. Multiplying this with an exponential value, gives a non-negative value. With all terms considered, it is clear that  $\nu_{t,l} = r(\mathbf{x}_{t,l}^{EC,P}) > 0, \forall \mathbf{x}_{t,l}^{EC,P}$ .

The design of (B.1) is a reaction to the performance of earlier builds of our models. One such build modelled the precision parameter as a constant. Figure B.1 displays the CRPS of this particular CPR-LP build for RWE in Germany at lead time 1, both as a function of the EC-forecast and the persistence forecast. Note how the CRPS reaches lower values at the edges, but is considerably higher in the intermediate domain. This observation is a result of the CPR-LP forecast being sharper at the edges, compared to the intermediate domain. The variance of the

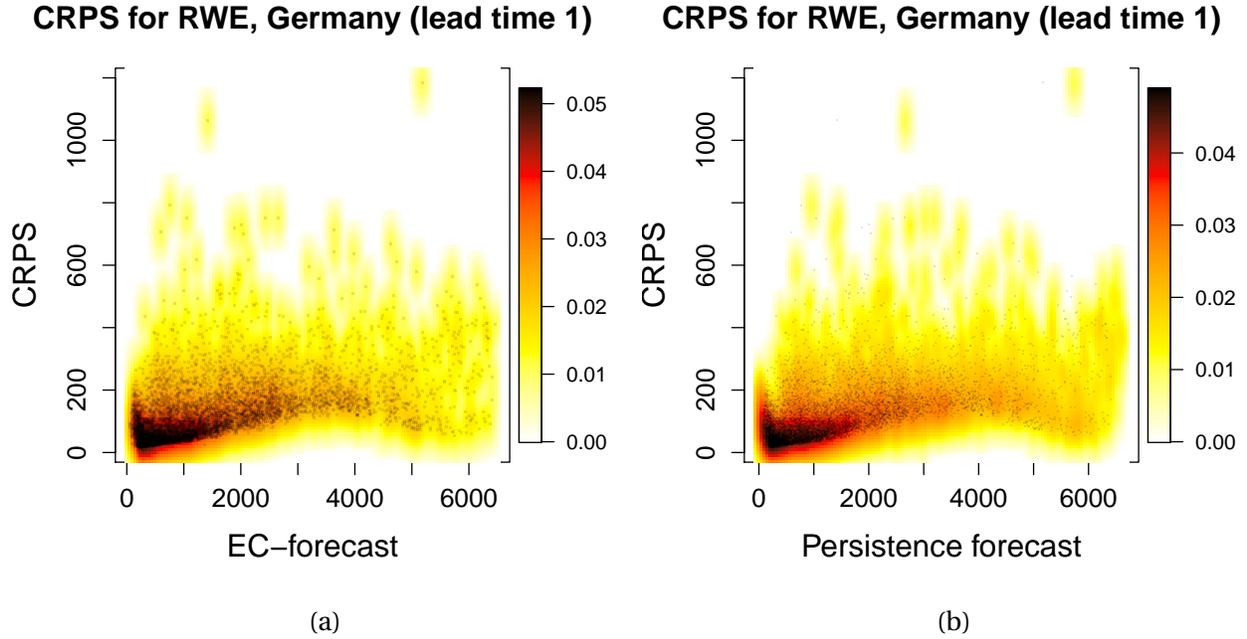


Figure B.1: CPRS for a previous build of the CPR-LP model where the precision parameter  $\nu$  was modelled as a constant. The CRPS is plotted against the EC-forecast in (a), and against the persistence forecast in (b). The colors represent the density of points through the transformation function  $f(x) = x^{0.25}$ , where  $x$  is a 2D kernel density of points, and  $f(x)$  is the density transformed to the color scale.

beta transformation in the CPR-LP model is related to the precision parameter such that the variance decreases when the precision parameter increases. To attain a sharp forecast in the intermediate domain, i.e. a beta transformation of low variance, we need a large value for the precision parameter in this region.

We choose to model the precision parameter through a function which supports this property. Let us define  $z(\cdot)$  as

$$z(x_{t,l}; a_{h,l}) = \exp(a_{h,l}) [G_h(x_{t,l})(1 - G_h(x_{t,l}))], \quad (\text{B.2})$$

where  $x_{t,l}$  is either the EC-forecast or the persistence forecast, and  $a_{h,l}$  is some parameter to be estimated. Figure B.2 displays (B.2) with  $a_{h,l} = 0$  and  $G_h(\cdot) = U_h(\cdot)$ , over the domain  $[0, 6700]$ . The full model of the precision parameter in (B.1) includes (B.2) in the second and third term, where  $a_{h,l} = a_{h,l}^{EC}$ ,  $x_{t,l} = x_t^{EC}$  for the second term, and  $a_{h,l} = a_{h,l}^P$ ,  $x_{t,l} = x_{t,l}^P$  for the third term. In addition to (B.2), the dispersion modelling in (B.1) includes an intercept and a fourth term

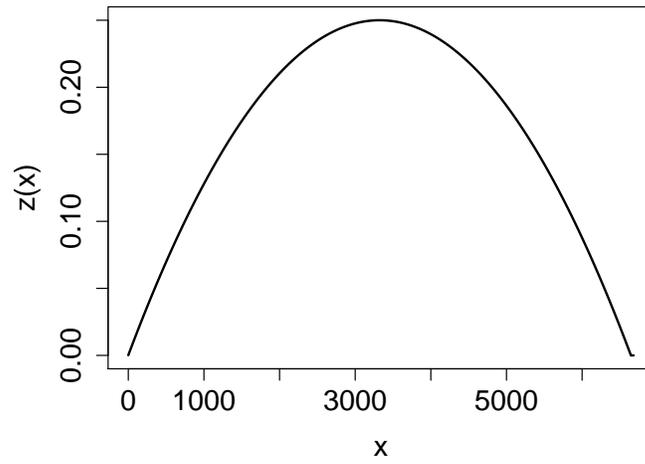


Figure B.2: Visualization of (B.2), with  $a_{h,l} = 0$  and  $G_h(\cdot) = U_h(\cdot)$ , over the domain  $[0, 6700]$ .

representing the absolute difference between the EC-forecast and persistence forecast.

# **Appendix C**

## **Additional Figures**

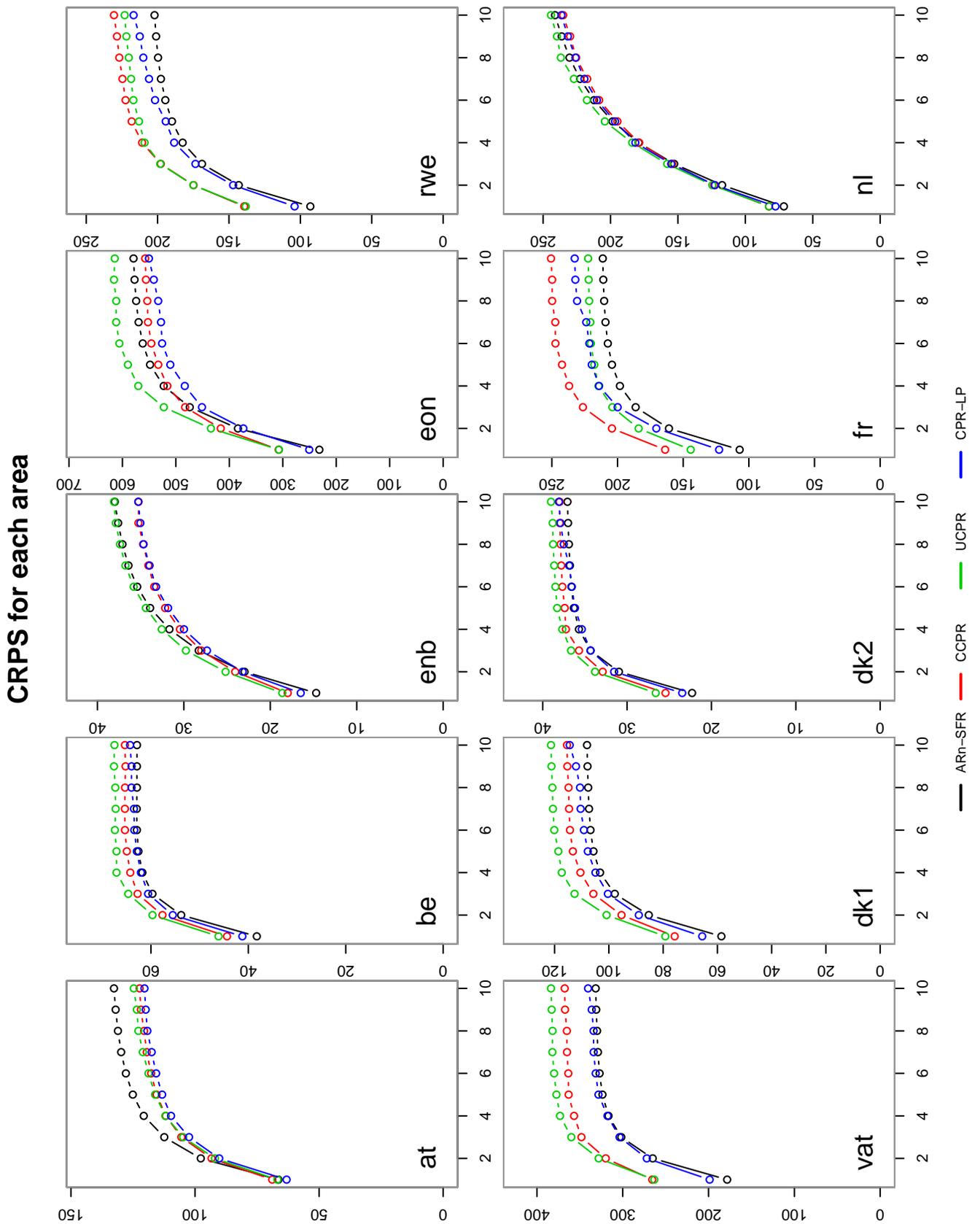
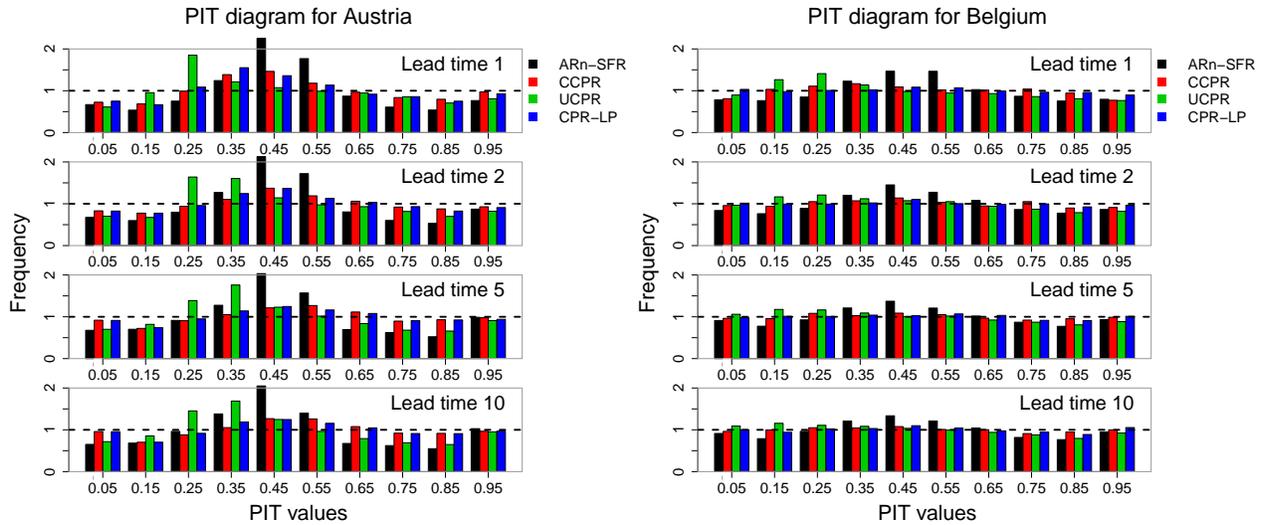
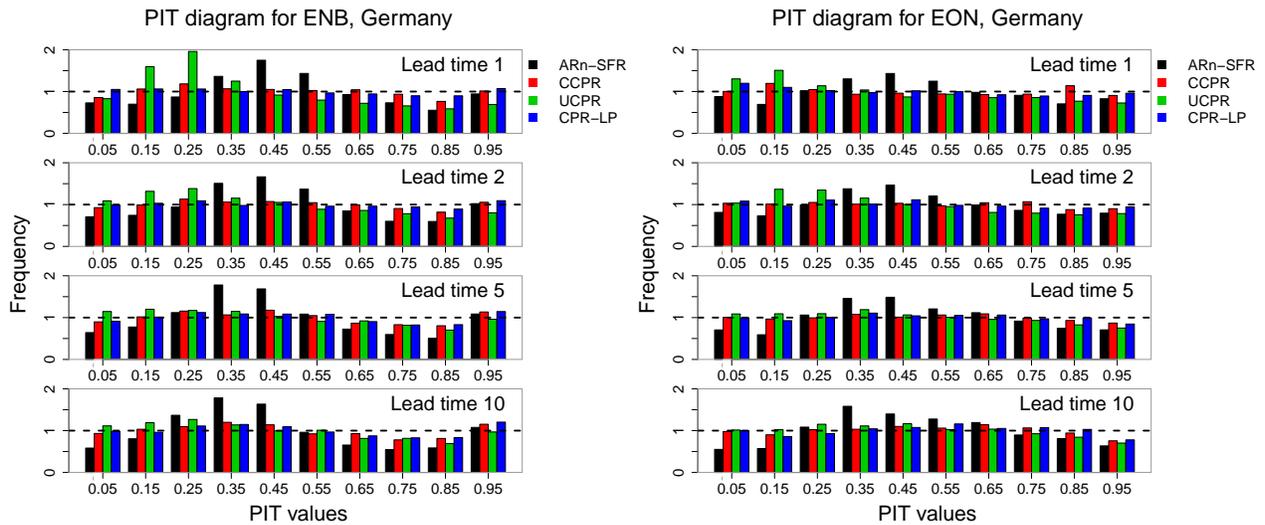


Figure C.1: CRPS for all models for each bidding area. CRPS values on the y-axis and lead times on the x-axis.



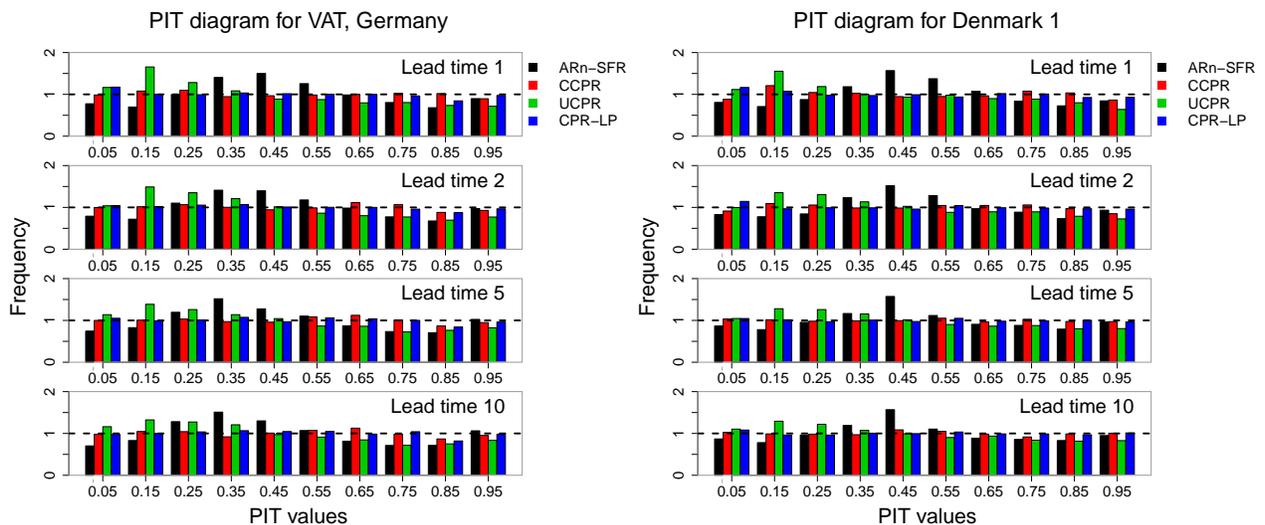
(a) PIT diagrams for bidding area Austria(AT).

(b) PIT diagrams for bidding area Belgium(BE)



(c) PIT diagrams for bidding area ENB, Germany.

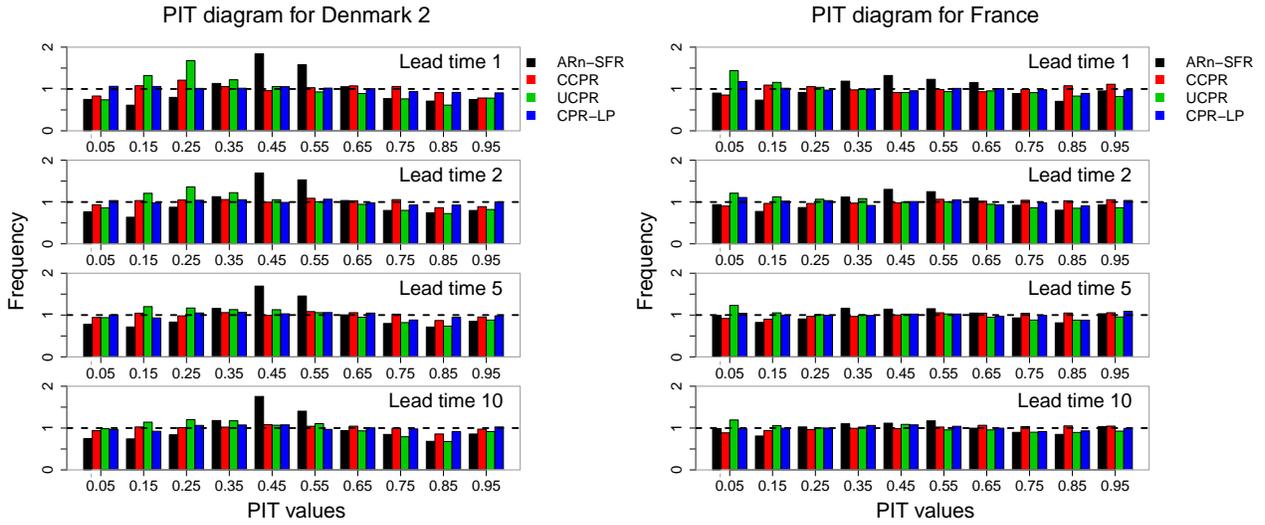
(d) PIT diagrams for bidding area EON, Germany.



(e) PIT diagrams for bidding area VAT, Germany.

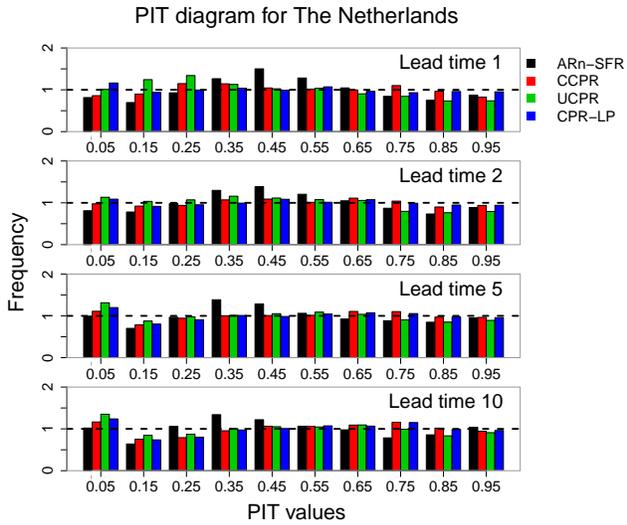
(f) PIT diagrams for bidding area Denmark 1(DK1).

Figure C.2



(a) PIT diagrams for bidding area Denmark 2(DK2).

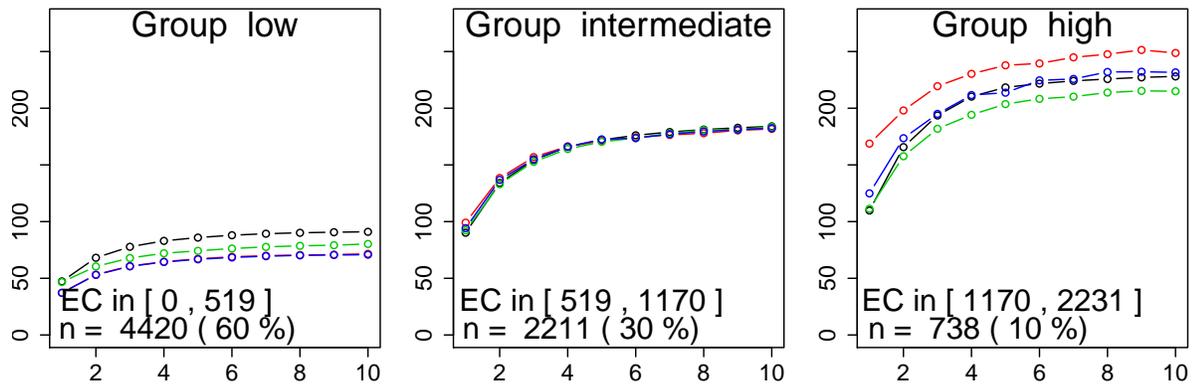
(b) PIT diagrams for bidding area France(FR).



(c) PIT diagrams for bidding area Netherlands(NL).

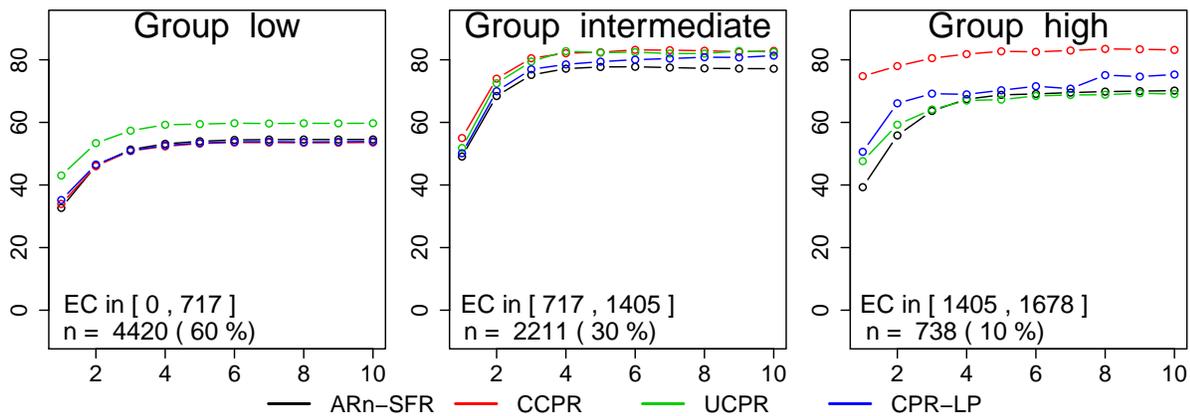
Figure C.3

### CRPS for Austria in groups



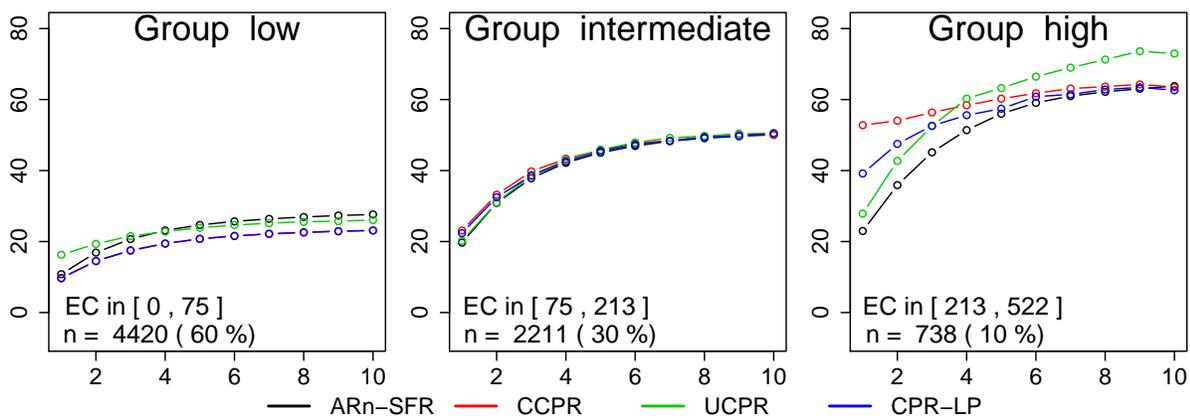
(a)

### CRPS for Belgium in groups



(b)

### CRPS for ENB, Germany in groups



(c)

Figure C.4: CRPS divided into three groups with respect to different values of EC-forecasts for various areas: Austria(a), Belgium(b) and ENB in Germany(c). CRPS on the y-axis and lead time on the x-axis. The groups are defined in Section 6.2.2. The minimum and maximum EC-forecast in MW are shown in the brackets, and  $n$  is the number of data points belonging to this group.

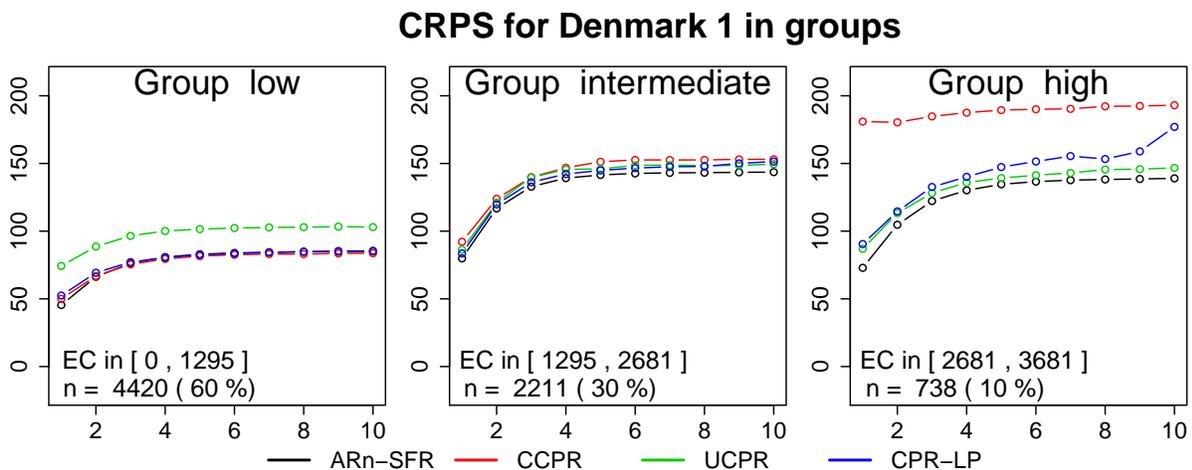
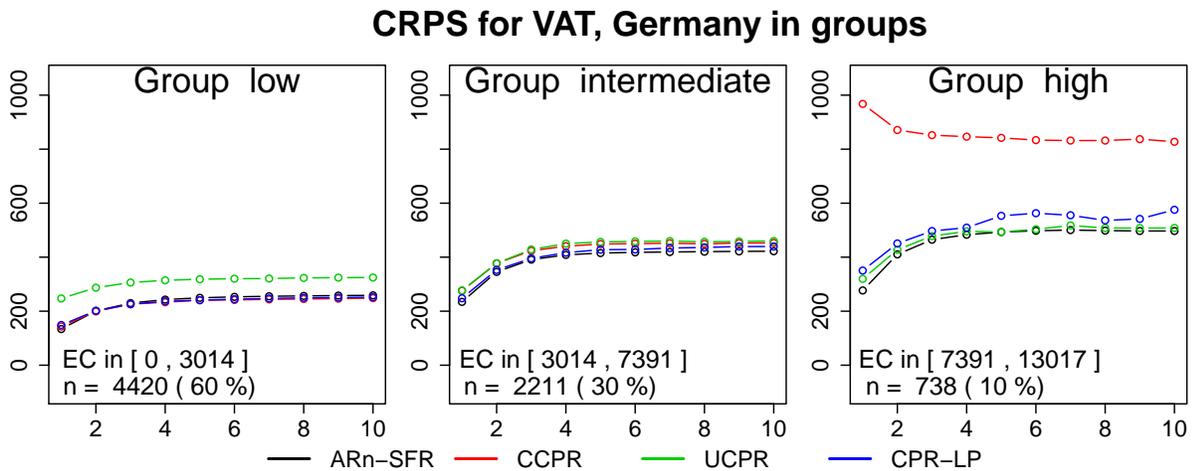
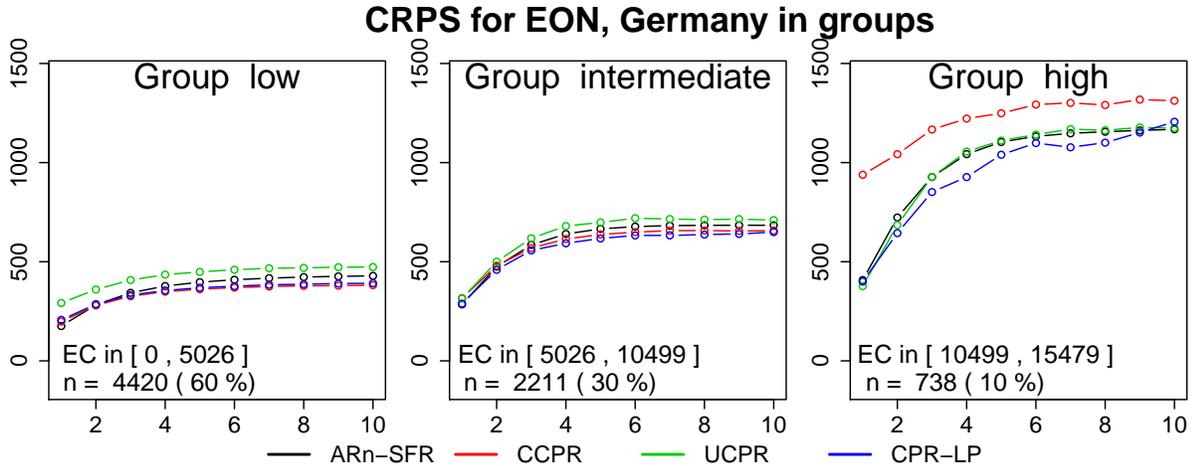


Figure C.5: CRPS divided into three groups with respect to different values of EC-forecasts for various areas: EON in Germany(a), VAT in Germany(b) and Denmark 1(c). CRPS on the y-axis and lead time on the x-axis. The groups are defined in Section 6.2.2. The minimum and maximum EC-forecast in MW are shown in the brackets, and  $n$  is the number of data points belonging to this group.

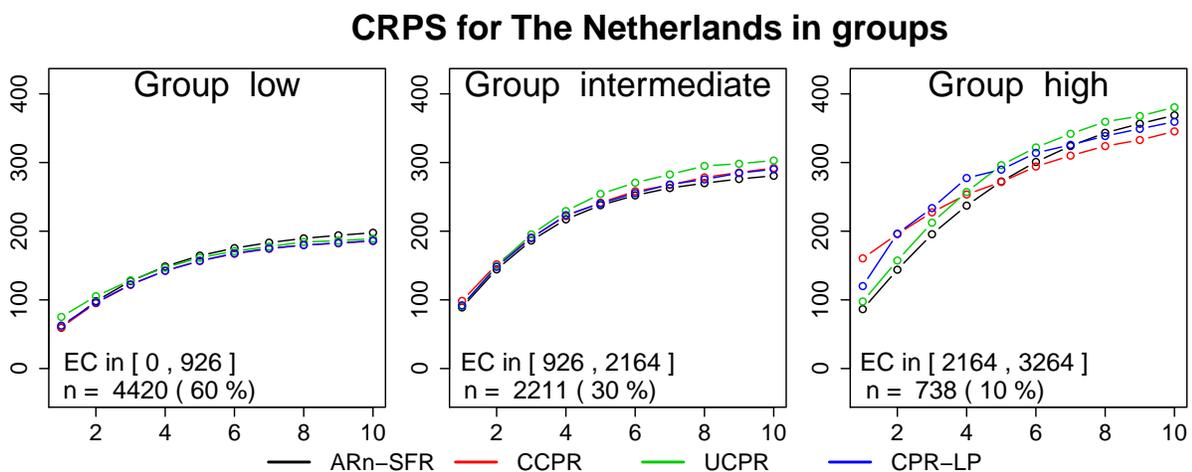
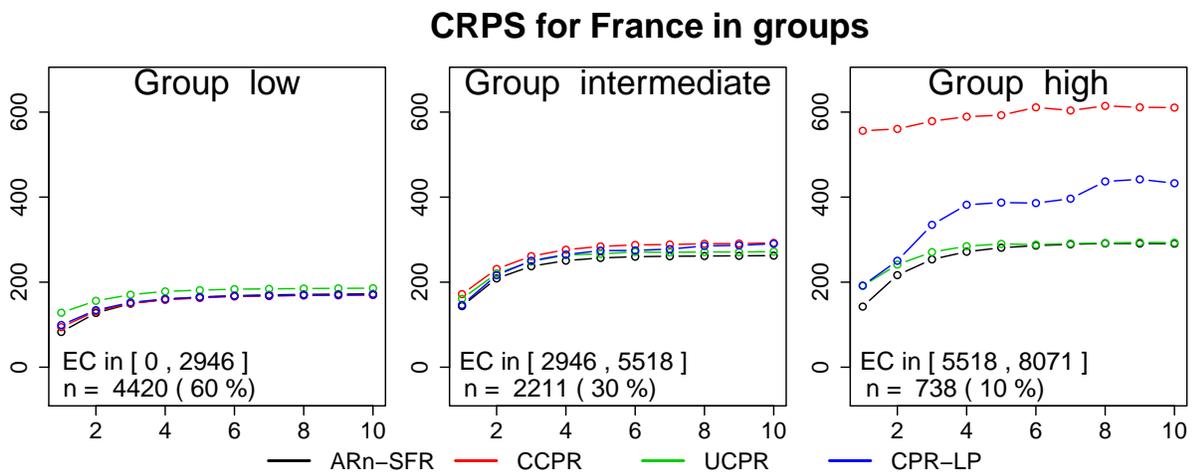
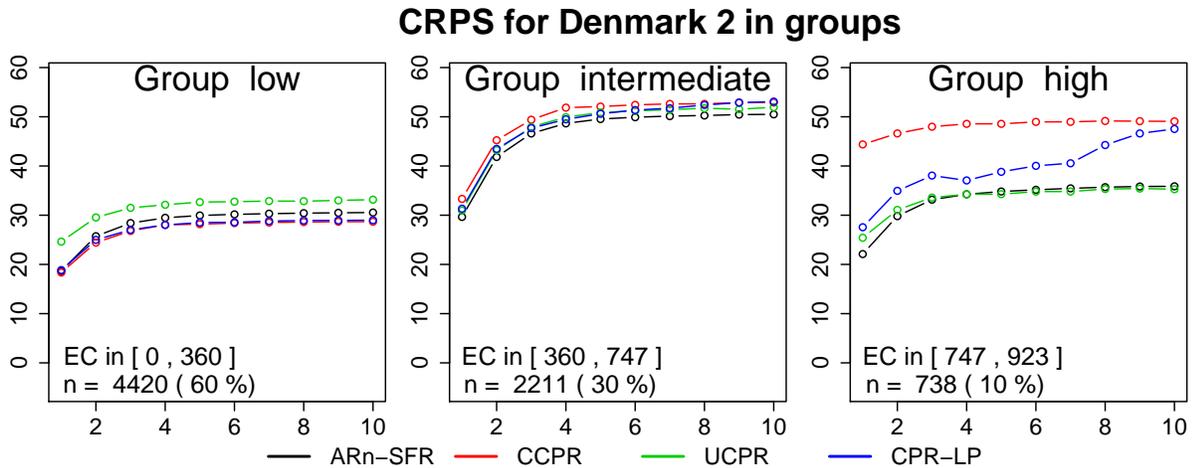


Figure C.6: CRPS divided into three groups with respect to different values of EC-forecasts for various areas: Denmark 2(a), France(b) and the Netherlands(c). CRPS on the y-axis and lead time on the x-axis. The groups are defined in Section 6.2.2. The minimum and maximum EC-forecast in MW are shown in the brackets, and  $n$  is the number of data points belonging to this group.

# Bibliography

- Bassetti, F., Casarin, R., and Ravazzolo, F. (2015). Bayesian nonparametric calibration and combination of predictive distributions. Working Paper 2015/03, Norges Bank.
- Borhaug, J. (2014). Climatology cumulative probability regression. Master's thesis, Norwegian University of Science and Technology.
- Ferrari, S. and Cribari-Neto, F. (2004). Beta regression for modelling rates and proportions. *Journal of Applied Statistics*, 31(7):799–815.
- Gneiting, T., Balabdaoui, F., and Raftery, A. E. (2007). Probabilistic forecasts, calibration and sharpness. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 69(2):243–268.
- Gneiting, T. and Raftery, A. E. (2005). Weather forecasting with ensemble methods. *Science*, 310:248–249.
- Gneiting, T. and Ranjan, R. (2013). Combining predictive distributions. *Electron. J. Statist.*, 7:1747–1782.
- Hamill, T. M. (2001). Interpretation of rank histograms for verifying ensemble forecasts. *Monthly Weather Review*, 129(3):550–560.
- Hersbach, H. (2000). Decomposition of the continuous ranked probability score for ensemble prediction systems. *Weather and Forecasting*, 15(5):559–570.
- Hora, S. C. (2004). Probability judgments for continuous quantities: Linear combinations and calibration. *Management Science*, 50(5):597–604.

- Johnson, N. L., Kotz, S., and Balakrishnan, N. (1995). *Continuous Univariate Distributions Vol. 2*, pages 210–211. Wiley, New York City, second edition.
- Malmgård, A. (2016). Prediction of wind power production combining previous observations and forecasts. Specialization Project, Norwegian University of Science and Technology.
- McCullagh, P. and Nelder, J. A. (1989). *Generalized linear models*, pages 21–32. London: Chapman & Hall, second edition.
- NordPool. Power trading description. <http://www.nordpoolspot.com/How-does-it-work/Intraday-market/>. Accessed: 2016-12-16.
- Ofgem, F. T. (2014). Bidding zones literature review.
- R Core Team (2016). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Raftery, A. E., Gneiting, T., Balabdaoui, F., and Polakowski, M. (2005). Using bayesian model averaging to calibrate forecast ensembles. *Monthly Weather Review*, 133(5):1155–1174.
- Ranjan, R. and Gneiting, T. (2010). Combining probability forecasts. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 72(1):71–91.
- Wallis, K. F. (2005). Combining density and interval forecasts: A modest proposal\*. *Oxford Bulletin of Economics and Statistics*, 67:983–994.
- Walpole, R. E., Meyers, R. H., Meyers, S. L., and Ye, K. E. (2012). *Probability and Statistics*, pages 443–450. Pearson, London, ninth edition.