Potential of Ensemble Copula Coupling for Wind Power Forecasting

Kaleb Phipps, Nicole Ludwig, Veit Hagenmeyer, Ralf Mikut

Institute for Automation and Applied Informatics, Karlsruhe Institute of Technology
Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen
E-Mail: kaleb.phipps@kit.edu

Abstract

With the share of renewable energy sources in the energy system increasing, accurate wind power forecasts are required to ensure a balanced supply and demand. Wind power is, however, highly dependent on the chaotic weather system and other stochastic features. Therefore, probabilistic wind power forecasts are essential to capture uncertainty in the model parameters and input features. The weather and wind power forecasts are generally post-processed to eliminate some of the systematic biases in the model and calibrate it to past observations. While this is successfully done for wind power forecasts, the approaches used often ignore the inherent correlations among the weather variables. The present paper, therefore, extends the previous post-processing strategies by including Ensemble Copula Coupling (ECC) to restore the dependency structures between variables and investigates, whether including the dependency structures changes the optimal post-processing strategy. We find that the optimal post-processing strategy does not change when including ECC and ECC does not improve the forecast accuracy when the dependency structures are weak. We, therefore, suggest investigating the dependency structures before choosing a post-processing strategy.

1 Introduction

As the share of renewable energy sources in the energy system increases, wind power forecasts become essential to guarantee balanced supply and demand. However, wind power highly depends on the chaotic weather system as well as other stochastic features and thus modelling uncertainty in these forecasts is important [1]. Probabilistic wind power forecasts aim to capture the uncertainty inherent in the model parameters and input features. Capturing this uncertainty is not easy and weather predictions, for example, are known to be biased and underdispersed. Meteorologists have therefore been post-processing ensemble weather predictions to describe the uncertainty more accurately [2, 3, 4, 5, 6, 7, 8, 9, 10]. In this post-processing the ensemble weather predictions are calibrated to the actual historical weather observations, eliminating some of the model biases. Transferring this approach to wind power forecasts yields promising results for handling the uncertainty in wind power forecasting models with uncertain weather features and forecast horizons of 3h-24h [11].

Phipps et al. [11] show that post-processing only the resulting wind power forecast is the best strategy with respect to forecast accuracy. However, their approach ignores dependencies between the weather variables as the post-processing is done using Ensemble Model Output Statistics (EMOS). This method involves fitting parametric distributions to the ensemble forecasts and sampling from them to generate post-processed forecasts. Random sampling causes inherent correlations, so-called dependency structures, between variables to be lost, which could affect the forecast performance. The present paper, therefore, extends the previous post-processing strategies by including Ensemble Copula Coupling (ECC) to restore the dependency structures between the variables and investigates whether including the dependency structures changes the optimal post-processing strategy. In contrast to other approaches, we do not investigate the temporal [12] or spatiotemporal [13] dependencies. We also use a parametric approach unlike the non-parametric methods used in [14].

The extended post-processing strategy is evaluated on two data sets with both a linear regression and an artificial neural network used as forecasting models. Both the ability of ECC to restore the dependency structures between the variables, as well as the resulting forecast accuracy are considered as performance measures.

The remainder of the present paper is structured as follows. Firstly we introduce theoretical concepts including ECC in Section 2. We then discuss the altered post-processing strategies (Section 3) before evaluating the effect of ECC in Section 4. We discuss the results in Section 5 before Section 6 concludes.

2 Background

Before evaluating the effect of ECC on wind power ensemble post-processing, we introduce EMOS, ECC, and also discuss the forecasting models we use.

2.1 Ensemble Post-Processing

The weather ensembles from the Ensemble Prediction System (EPS) are known to be biased and underdispersed, and thus need to be calibrated [2]. Not accounting for this bias or under estimated variance could lead to false wind power forecasts which affect the stability of the energy system. The present paper applies EMOS developed by Gneiting et al. [6] to perform this calibration. EMOS is based on non-homogenous regressions and performed for each weather variable individually given a single origin and set forecast horizon.

The standard EMOS approach is designed for ensemble members that are individually distinguishable. Ensemble members from the European Centre for Medium-Range Weather Forecasts (ECMWF) are, however, classified as exchangeable, thus representing equally likely future scenarios without distinguishing features [15, 16]. Given exchangeable ensembles $x_1, \dots, x_M$, we apply EMOS from Gneiting and Katzfuss [5], where the weather variable y with an assumed normal distribution is modelled as

$$y|x_1, \dots x_M \sim \mathcal{N}\left(a + b\sum_{m=1}^{M} x_m, c + dS^2\right), \tag{1}$$

with a, b, c and d being regression coefficients, and the variance S^2 being a linear function of the ensemble spread with

$$S^2 = \frac{1}{M} \sum_{m=1}^{M} \left(x_m - \frac{1}{M} \sum_{m=1}^{M} x_m \right)^2 . \tag{2}$$

We also use the same approach with a truncated normal distribution. Since we post-process each weather variable for a single origin and set forecast horizon, the EMOS coefficients change when any one of these three parameters are altered. We apply EMOS using a rolling calibration window of 40 days with the help of the *scoringRules*[1] package. For more information on EMOS and its application in meteorology see Gneiting et al. [6] or Gneiting [17].

2.2 Ensemble Copula Coupling

With the EMOS method introduced above, we now have ensembles that are calibrated to the past data. All weather variables are modelled with a univariate distribution. We could sample from each of these distributions independently and input the information into the wind power forecasting scheme. However, the original Numerical Weather Prediction (NWP) includes information on the weather variables dependencies among each other as well as in space and time. This information gets lost with the univariate EMOS approach. In order to retain these dependencies, several empirical copula-based approaches have been developed, which we want to introduce in more detail in this section.

A d-dimensional Copula is a multivariate cumulative distribution on the unit cube $[0,1]^d$ with uniform margins [18, 19]. The importance of copulas in restoring dependency structures is based on the theorem of Sklar [20]. He states that for any multivariate cumulative distribution function F with margins $F_1, \ldots, F_M$ there exists a copula C, that is unique on the range of the margins and has the form

$$F(x_1 \ldots, x_M) = C(F_1(x_1), \ldots F_M(x_M)), \tag{3}$$

[1] https://cran.r-project.org/web/packages/scoringRules/scoringRules.pdf

for $x_1,\ldots,x_m \in \mathbb{R}$. With regards to ensemble calibration, we already have the uniform margins $F_1,\ldots,F_M$ in the form of the EMOS univariate distribution. Therefore, Sklar's theorem states that as long as an appropriate copula is defined, univariate ensemble post-processing techniques can be used to accommodate any dependency structure.

The ECC approach, is based on the mathematical concepts defined above with the appropriate copula in the form of a *reordering* process. The idea is that given a dependence structure "template" [21], the samples drawn from the univariate EMOS distributions can be reordered in such a way that they resemble the initial correlation structures that were given in the NWP. The templates are based on the raw ensembles, where we assume that the ensembles capture the correlations. While several variants exist, we take a closer look at random ECC (ECC-R), quantile ECC (ECC-Q), and transformation-based ECC (ECC-T). These methods differ from each other in the way they sample from the distributions and whether they include a reordering step. Firstly, we take a look at an ECC approach based on random draws. Here, we sample with the independent standard uniform random variates $u_1,\ldots,u_M$, such that

$$\tilde{x}_1 = F^{-1}(u_1),\ldots,\tilde{x}_M = F^{-1}(u_M). \tag{4}$$

In this ECC-R approach, the samples have to be reordered according to the template depicted by the raw ensembles. This template is based on the rank structure. Given each time horizon, location and variable and following the notation by Schefzik et al. [21], the raw ensembles $x_1,\ldots x_M$ and their order statistics $x_{(1)} \leq \ldots \leq x_{(M)}$ construct a ranking permutation π, with $\pi(m) := \mathrm{rank}(x_m)$ for $m \in \{1,\ldots,M\}$. After drawing the samples from the distribution, they are reordered according to π. The ECC approach based on quantiles involves the same reordering step as in the ECC-R approach, the sampling method is however different. In the case of ECC-Q, the samples are drawn from equally spaced quantiles, such that

$$\tilde{x}_1 = F^{-1}\left(\frac{1}{M+1}\right),\ldots,\tilde{x}_M = F^{-1}\left(\frac{M}{M+1}\right) \tag{5}$$

and then reordered as before. In contrast to the ECC-R and ECC-Q approach, ECC-T relies on a transformation and does not require an additional reordering step. The samples are drawn such that

$$\tilde{x}_1 = F^{-1}(S(x_1)), \dots, \tilde{x}_M = F^{-1}(S(x_M)), \tag{6}$$

where S is the fit of a cumulative distribution function to the raw ensembles [21]. The choice of S depends on the variables in question and in the case of temperature, pressure and wind component vectors, S can be assumed as normal with mean equal to the ensemble mean and variance equal to the ensemble variance [3].

2.3 Forecasting Models

The present paper focuses on the effect of ECC on optimal post-processing strategies and not developing state-of-the-art wind power forecasts. Therefore, we use the same forecasting models (a linear regression model and a neural network), and the same forecasting strategy as in [11]. Both are described in the following together with how we measure the forecasting accuracy.

2.3.1 Linear Regression

The simplest models, which we use to forecast wind power, are linear regression models. These linear regression models can be described with

$$y_{t+h} = \beta_0 + \alpha y_{t+h-24} + \sum_{k=1}^{K} \beta_k W_{t+h}^k + \sum_{j=1}^{J} \gamma_j D_{t+h}^j + \varepsilon_{t+h}, \tag{7}$$

where y_t is the dependent variable, which in this case is the wind power, y_{t-24} is the actual wind power a day before, W^k are weather time series, such as wind speed and temperature, and D^j are dummy variables, such as the season, the month and the year. The models are fitted for each forecast horizon $h = h_1 \dots h_H$ with $h \leq 24$ using actual historical weather data in order to describe the real relationship among the variables and remove any bias fitting on historical weather forecasts or ensembles could introduce. Each ensemble $x_1 \dots x_M$ from the EPS is then used in a separate prediction run for each forecast horizon to generate an ensemble of wind power predictions with the previously fitted regression coefficients

$$\hat{y}_{t+h}(x_1,\dots x_M) = \hat{\beta}_0 + \hat{\alpha} y_{t+h-24} + \sum_{k=1}^{K} \hat{\beta}_k \widehat{W}^k_{t+h}(x_1,\dots x_M) + \sum_{j=1}^{J} \hat{\gamma}_j D^j_{t+h} + \hat{\varepsilon}_t, \tag{8}$$

where $\widehat{W}^k_{t+h}$ is the weather forecast made at time t for the forecast horizon h.

2.3.2 Neural Network

In order to better forecast non-linear dependencies, we implement a neural network. We tested multiple neural network configurations before selecting a configuration with two hidden layers of 10 and 7 neurons respectively and trained it with the resilient backpropagation algorithm. This network architecture was selected because it is the simplest we found, that still returns accurate forecasts. The chosen activation function is a hyperbolic tangent given by

$$\sigma(x) = \frac{e^{2x} - 1}{e^{2x} + 1} \in [-1, 1]. \tag{9}$$

The input features remain the same as for the linear regression model explained above. Again, the parameters (i. e.weights) are fitted using the actual historical weather data and each ensemble member is passed through the network to get an ensemble wind power prediction. The neural networks are implemented in R with the *neuralnet* package[2].

[2] https://cran.r-project.org/web/packages/neuralnet/neuralnet.pdf

2.3.3 Forecasting Accuracy

To evaluate the forecasting approaches, we use the Continuous Ranked Probability Score (CRPS). This error measure is used to assess the calibration and sharpness of the probabilistic forecast and can be described as follows

$$\mathrm{CRPS}(F,y) = \int_{\mathbb{R}} \left(F(z) - \mathbb{1}\{y \leq z\}\right)^2 dz, \tag{10}$$

with F being the wind power generations predictive cumulative distribution function, y the verifying observation and $\mathbb{1}$ denoting an indicator function. We report the score over all time steps $t = 1, \ldots N$ in the test set

$$\mathrm{CRPS} = \frac{1}{N} \sum_{t=1}^{N} \mathrm{CRPS}(F_t, y_t). \tag{11}$$

3 Post-Processing Strategies

The present paper focuses on determining whether including ECC into the post-processing of wind power forecasts affects the performance of these strategies and changes the optimal post-processing strategy. Four post-processing strategies were identified by Phipps et al. [11] and these are shown with the addition of ECC in Figure 1. We identify two strategies that can be extended to include ECC, whilst two remain the same:

Raw. Using the raw weather ensembles directly in the forecast is not affected by ECC. Here we take all available M ensemble members from the EPS for multiple weather variables to generate the wind power forecast. Thus, the resulting wind power forecast consists of M members.

One-Step-P. The second strategy identified by Phipps et al. [11] is also unchanged by ECC. The output of the wind power forecasting model is post-processed, without any previous calibration of the input variables. This assumes that post-processing the wind power ensembles also accounts for the biases in the weather ensembles. For a detailed description of the One-Step-P approach see Phipps et al. [11].

One-Step-W. This strategy involves calibrating the raw weather ensembles before they are used as inputs in the wind power forecast model. We initially post-process the weather ensembles analogue to Phipps et al. [11]: Each weather variable (temperature, wind speed, wind component vectors etc.) is considered separately and post-processed using EMOS with a rolling calibration window. This results in probability distributions for each weather variable and we form new post-processed weather ensembles by sampling from these distributions. It is in this resampling stage that we apply ECC. Depending on the method (see Section 2.2) we either reorder the EMOS samples or also consider a different sampling strategy or transformation. As a result the strategy now has multiple variants. We use the resulting post-processed ensembles with restored dependency structures as inputs into the wind power forecast model. The resulting ensemble of wind power forecasts is not processed further.

Two-Step-WP. The final strategy is also altered when we include ECC. In this strategy both one-step post-processing approaches are coupled together. We first post-process the weather ensembles using the altered One-Step-W strategy including ECC. These ensembles are used to generate a wind power ensemble forecast and then we again post-process the result as in the One-Step-P method. Since there is only one set of wind power ensembles, therefore not dependencies present, we are unable to apply ECC after the first EMOS application step.

In the present paper we compare these approaches, focusing on discovering if the inclusion of ECC in One-Step-W and Two-Step-WP improves the forecast performance. Specifically through an evaluation on two data sets, we investigate if ECC leads to one of these two strategies outperforming One-Step-P, which was previously found to perform best [11].

4 Evaluation

We use two different data sets to evaluate the effect of ECC on the performance of the post-processing strategies described above. In this section we briefly introduce those data sets and present the results of our evaluation.

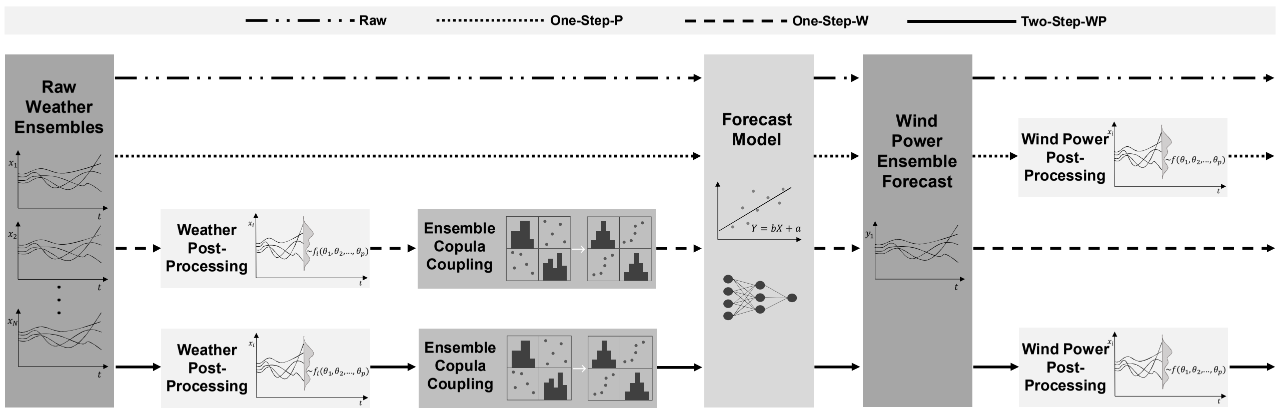


Figure 1: An overview of the post-processing strategies compared. Whilst no post-processing (Raw) and post-processing only the power ensembles (One-Step-P) are unaffected by adding ECC, we extend the other two strategies. Therefore we add ECC after the EMOS stage, when we post-process only the weather ensembles (One-Step-W). We also include ECC after the EMOS stage for the weather variables but before the post-processing of the power ensembles in the Two-Step-WP strategy.

4.1 Data

We evaluate the post-processing strategies on two data sets: A benchmark data set including both an onshore and offshore wind park, and real data from bidding zones 3 and 4 in Sweden. This section briefly introduces the data used.

4.1.1 Benchmark Data

The benchmark data set is based on simulated wind power data based on real wind parks located in Germany. This data was simulating using the *renewable ninjas*[3] API, with the input data being selected to mimic real onshore and offshore wind parks in Germany as closely as possible. Staffell and Pfenninger [22] verify that the simulation and bias-corrections implemented in the renewable ninjas API are capable of reproducing accurate wind power time series. A detailed description of the parameters used in the simulation is provided by Phipps et al. [11].

We access open source weather ensemble data through The International Grand Global Ensemble (TIGGE) archive[4]. TIGGE archive is a result of *The Observing System Research and Predictability Experiment* which aimed to combine ensemble forecasts from leading forecast centres to improve probabilistic forecasting capabilities [23]. Due to damaged tapes in TIGGE we only use data from February 2017 until August 2018, including the parameters two-meter temperature, surface pressure, 10m-U-Component of wind, 10m-V-Component of wind and wind speed. Weather data is downloaded for the same location as the synthetic wind park generated through renewable ninjas. We use the ERA5 reanalysis data for the ground truth historical weather data [24]. We download the identical weather parameters for the same timespan via the Copernicus Climate Data Store (CDS) API [5]. Data from 2017 is used for training the forecast models and from 01.2018-08.2018 for evaluation. For

[3] www.renewables.ninja.
[4] https://apps.ecmwf.int/data sets/data/tigge/
[5] https://cds.climate.copernicus.eu/home

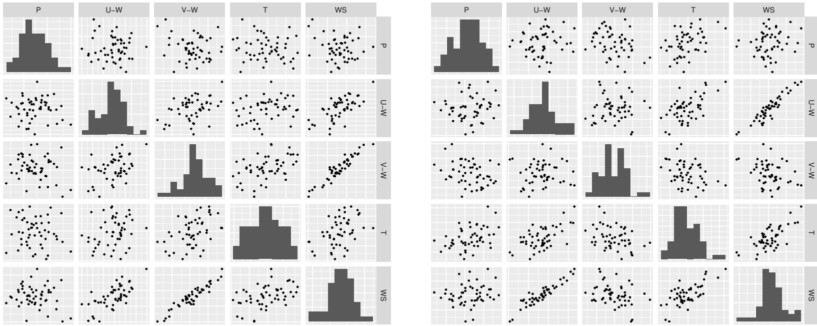
more detailed information on the benchmark data set, including information on how to replicate it, see Phipps et al. [11]

4.1.2 Swedish Data

The Swedish electricity system is divided into four sub-areas or bidding zones with the present paper focusing only on the area contained in bidding zones 3 and 4. We download the wind power generation data aggregated on a bidding zone level through the open-source transparency platform which is operated by the European Network of Transmission System Operators (ENTSO-E) [25]. This data is available at an hourly resolution, but due to limitations in the weather data we can only forecast every 3h. The weather data for bidding zone 3 and 4 is made up of the ECMWF EPS Molteni et al. [15] and the ERA5 reanalysis data C3S [24]. The ERA5 data again serves as the ground truth for post-processing, whilst we use the EPS for the ensemble forecasts. We download the parameters two-meter temperature, surface pressure, 100m-U-Component of wind, 100m-V-Component of wind and wind speed. Since the weather data only comes in a grid-based format and we perform forecasts for an entire bidding zone, this weather data must be aggregated. We use a weighted average method to perform this aggregation. We use data from 2015-2017 for training our forecast models and from 01.2018-08.2019 for evaluation. A detailed description of this data set, including the weighted average aggregation method is provided by Phipps et al. [11].

4.2 Results

We evaluate both the ability of ECC to restore the dependency structures and the effect this has on forecast performance. This section presents the results of the analysis for both data sets introduced above. When analysing performance of ECC with scatter plots we only consider April 14, 2018 at 6:00am. The results for all other dates and forecasts horizons are similar and therefore not discussed in detail.



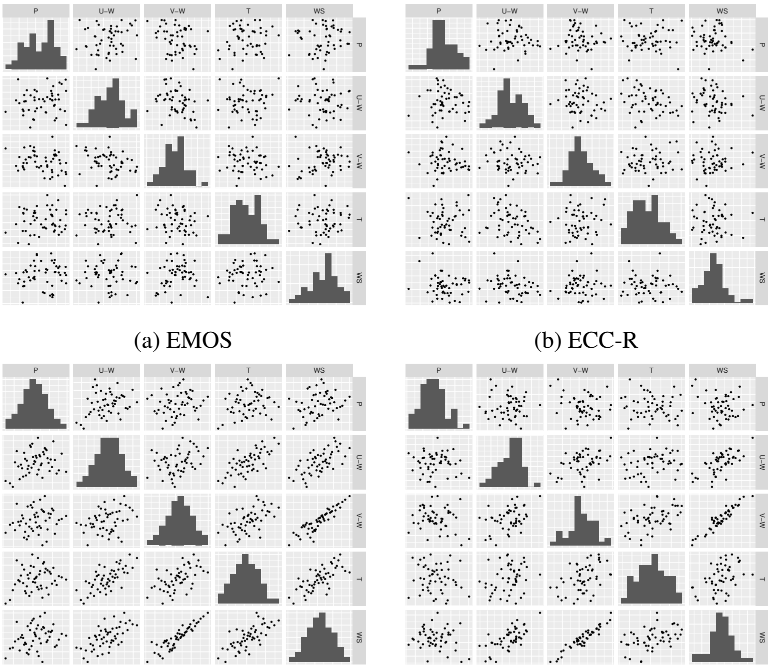
(a) Onshore Benchmark (b) Offshore Benchmark

Figure 2: Scatter plots showing the raw dependency structures for both benchmark data sets on April 14, 2018 at 6am. Pressure (P), the U-component of wind (U-W), the V-component of wind (V-W), temperature (T) and wind speed (WS) are shown. Despite correlation between wind components and wind speed, the dependencies between the weather variables are not particularly strong.

4.2.1 Benchmark Data

First we consider the effect of ECC on restoring the dependency structures in the data. We see the dependencies between various weather variables in the form of scatter plots in Figure 2. These plots show a histogram of the empirical distribution of individual weather ensembles along the diagonal and scatter plots of their dependencies in the other positions. In the onshore benchmark there is a clear correlation between the V-component of wind and wind speed, and also a slight correlation between wind speed and the U-component of wind. There are no strong dependencies shown between any other weather variables. For the offshore benchmark data the only noticeable correlation is between the U-component of wind and wind speed.

Figures 3 and 4 show how these dependency structures change after applying ensemble post-processing and various ECC methods. Figure 3 details the onshore benchmark and we see, that only applying EMOS removes all dependency structures. ECC-R is also not effective, with no dependencies being visible. Both the ECC-Q and ECC-T methods show improvement. The quantile sampling based ECC-Q leads to an almost symmetric marginal distribution,

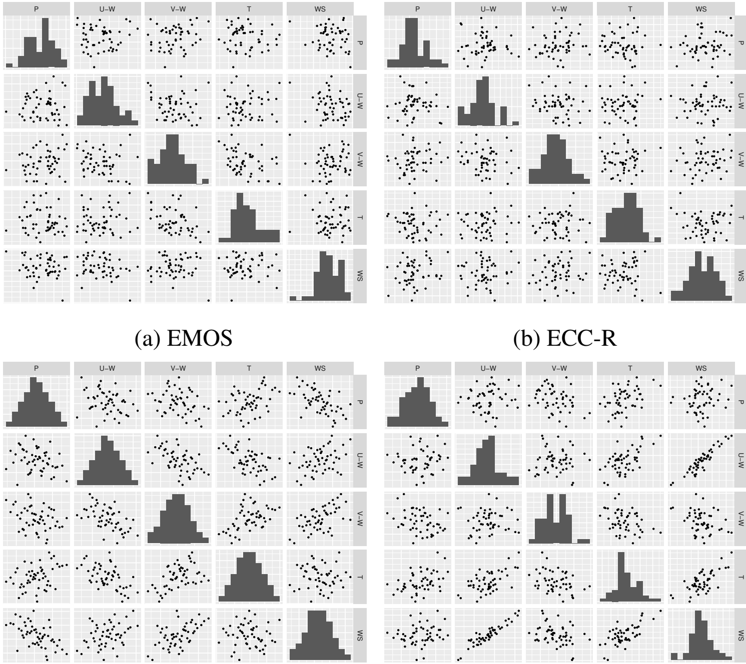


(a) EMOS
(b) ECC-R
(c) ECC-Q
(d) ECC-T

Figure 3: Scatter plots showing the dependency structures with various post-processing strategies that aim to restore the raw dependency structure of the onshore benchmark on April 14, 2018 at 6am. Pressure (P), the U-component of wind (U-W), the V-component of wind (V-W), temperature (T) and wind speed (WS) are shown. We see that EMOS destroys all dependencies and ECC-R is not effective in restoring the structures. ECC-Q and ECC-T both restore the dependency structures effectively.

but also accurately recreates the dependencies between the weather variables. Since ECC-T is based on a transformation it is not surprising that this method recreates the dependencies with the most accuracy. The results are similar for the offshore benchmark in Figure 4, with the exception of the ECC-Q method. Although we see some dependency structures being recreated, ECC-Q is not as effective here as in the onshore data set.

In order to assess the effect of ECC on forecast accuracy we consider plots of the mean CRPS for each forecast horizon on each benchmark data set. Figure



(a) EMOS (b) ECC-R

(c) ECC-Q (d) ECC-T

Figure 4: Scatter plots showing the dependency structures with various post-processing strategies that aim to restore the raw dependency structure of the offshore benchmark on April 14, 2018 at 6am. Pressure (P), the U-component of wind (U-W), the V-component of wind (V-W), temperature (T) and wind speed (WS) are shown. We see that EMOS destroys all dependencies and ECC-R is not effective in restoring the structures. ECC-Q shows some improvements, but ECC-T restores the dependency structures the best.

5 compares the means CRPS of One-Step-P against variants of One-Step-W. Figure 6 also plots the mean CRPS scores, but this time for variations of Two-Step-WP against One-Step-P. The One-Step-P strategy is almost always slightly more accurate than the variations of One-Step-W and very similar to the Two-Step-WP variations. We also see, that there is almost no difference between the various variations One-Step-W and Two-Step-WP based on different ECC methods. The neural networks perform worse than the linear models on the benchmark data. This is mainly due to a lack of training data but

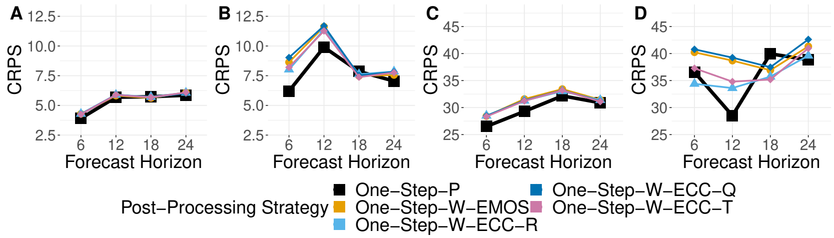


Figure 5: Plot comparing the average CRPS score for the test data on the benchmark data set for each forecast horizon using the One-Step-W variants to the One-Step-P strategy. In (A) we see the linear model for the onshore data and (B) the neural network. (C) and (D) show the evaluation of the linear model and neural network on the offshore data set.

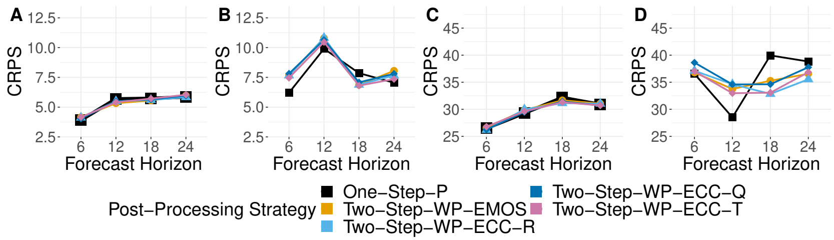


Figure 6: Plot comparing the average CRPS score for the test data on the benchmark data set for each forecast horizon using the Two-Step-WP variants to the One-Step-P strategy. In (A) we see the linear model for the onshore data and (B) the neural network. (C) and (D) show the evaluation of the linear model and neural network on the offshore data set.

also suggest that a linear model is more than capable of delivering accurate forecasts.

4.2.2 Swedish Data

The effect of ECC on the dependency structures in the Swedish data is similar to the benchmark data set and therefore we do not discuss it here in detail. The mean CRPS values for bidding zone 3 are shown in Table 1 and for bidding zone 4 in Table 2. Again all One-Step-W variations deliver similar results and perform noticeably worse than the One-Step-P strategy. The Two-Step-WP performs similarly to the One-Step-P method, but again applying ECC has little effect. On a whole, we see that as with the benchmark data set ECC has

Table 1: Summary of mean CRPS for the test data in bidding zone 3 in Sweden. The best prediction for each strategy and each forecast model is highlighted in bold.

Data Set		6h	12h	18h	24h
Bidding Zone 3	Linear Raw	97.12	84.86	91.97	94.28
	Linear One-Step-P	64.12	68.90	64.62	69.27
	Linear One-Step-W-E	107.95	95.28	103.85	100.43
	Linear One-Step-W-R	107.95	95.26	103.21	100.24
	Linear One-Step-W-Q	108.21	95.48	103.67	100.27
	Linear One-Step-WP-T	107.89	95.43	103.60	100.17
	Linear Two-Step-WP-E	63.70	**67.26**	**63.28**	67.38
	Linear Two-Step-WP-R	64.50	67.40	63.61	**67.02**
	Linear Two-Step-WP-Q	64.55	67.41	63.65	68.10
	Linear Two-Step-WP-T	**63.58**	67.78	64.04	68.03
	Neural Raw	64.35	70.73	72.55	66.99
	Neural One-Step-P	**61.05**	63.02	67.13	**59.54**
	Neural One-Step-W-E	73.40	64.81	63.26	74.45
	Neural One-Step-W-R	74.60	65.29	65.34	74.11
	Neural One-Step-W-Q	74.85	65.52	65.07	75.45
	Neural One-Step-W-T	74.79	66.18	64.52	74.38
	Neural Two-Step-WP-E	65.69	**59.37**	56.79	69.38
	Neural Two-Step-WP-R	67.23	60.81	57.12	69.65
	Neural Two-Step-WP-Q	66.27	61.04	57.29	70.87
	Neural Two-Step-WP-T	67.51	62.23	**56.22**	70.29

almost no effect on the forecast accuracy, with all ECC variants performing similarly to post-processing strategies where only EMOS is used. In the case of the Swedish data we also note, that the neural network performs slightly better than the linear model in bidding zone 3.

5 Discussion

The fact that the ECC variants perform differently is not surprising as similar results are observed by Schefzik et al. [3]. Across all data sets applying EMOS destroys the dependency structures and ECC-R is relatively ineffective

Table 2: Summary of mean CRPS for the test data in bidding zone 4 Sweden. The best prediction for each strategy and each forecast model is highlighted in bold.

Data Set		6h	12h	18h	24h
	Linear Raw	58.50	67.00	56.51	51.54
	Linear One-Step-P	45.13	51.90	50.35	44.26
	Linear One-Step-W-E	59.48	60.39	55.97	50.73
	Linear One-Step-W-R	59.48	60.78	56.33	51.28
	Linear One-Step-W-Q	59.71	61.01	56.35	50.71
	Linear One-Step-W-T	59.46	60.80	56.33	50.72
	Linear Two-Step-WP-E	45.21	52.69	49.89	43.92
	Linear Two-Step-WP-R	45.00	52.75	50.48	43.48
	Linear Two-Step-WP-Q	44.54	52.38	**50.11**	43.41
Bidding Zone 4	Linear Two-Step-WP-T	**44.43**	**51.44**	50.23	**43.34**
	Neural Raw	52.74	**46.70**	51.02	**43.98**
	Neural One-Step-P	**49.55**	47.80	**48.11**	46.22
	Neural One-Step-W-E	82.29	90.97	100.03	82.26
	Neural One-Step-W-R	81.34	88.97	98.62	80.29
	Neural One-Step-W-Q	78.32	86.29	95.83	78.22
	Neural One-Step-W-T	74.79	82.88	89.79	73.70
	Neural Two-Step-WP-E	52.06	58.04	63.87	49.42
	Neural Two-Step-WP-R	51.79	56.92	62.82	49.39
	Neural Two-Step-WP-Q	52.48	57.19	63.47	49.00
	Neural Two-Step-WP-T	53.27	56.67	61.52	49.35

in rebuilding dependency structures. Since ECC-Q relies on quantile sampling, it produces marginal distributions that are always close to symmetric. With regards to dependency structures, ECC-Q delivers mixed results. For the onshore benchmark (and also bidding zone 3) it recreates the dependency structures almost as accurately as ECC-T, but not for the other data sets. This could be due to ECC-Q still relying on sampling from the EMOS distribution, which we obtain by considering the last 40 days. Therefore it is possible that the EMOS parameters vary in accuracy which could lead to a ECC-Q sampling that is also slightly worse. ECC-T on the other hand is based on a monotonic transformation and it is therefore expected that it always accurately recreates the dependency structures of the raw ensembles. ECC-T therefore unsurprisingly performs the best in this regard, but the trade-off is the marginal distributions, which are not as symmetrical as those from ECC-Q.

The key focus of the present paper is however, evaluating how the inclusion of ECC affects post-processing strategies. Although, as discussed above, the performance of the various ECC variants differs, this appears to have no effect on the forecast performance. The only instance in which the ECC variations noticeably diverge is when the neural network is used in bidding zone 4 in Sweden with the One-Step-W strategy. In this case the One-Step-P method performs far better than any One-Step-W variation, and the EMOS variant of One-Step-W performs better than all ECC strategies. In this case ECC didn't lead to any improvement, but actually caused the forecasts to be slightly worse. For other post-processing strategies and forecast horizons there is no noticeable difference between the One-Step-W or Two-Step-WP variations with and without ECC. For the forecast horizons and data sets considered, we therefore find that dependency structures do not play a significant role with regards to wind power forecast accuracy.

We consider two possible explanations for this. Firstly, the dependency structures present in both our data sets are not strong. Although Schefzik et al. [21] reported improved results when using ECC, they were working with data sets that showed clear dependencies. Since there are no strong dependency structures for ECC to restore in our datasets, it makes sense that this does not lead to an improvement. We therefore suggest to investigate the dependency structures before deciding whether ECC is included in the post-processing strategy. The second explanation is that the wind power forecast model is capable of implicitly restoring these dependencies. When the weather ensembles are used to generate wind power forecasts an implicit calibration (due to the training of the model with historical weather variables and observed/simulated wind power generation) occurs and this could be sufficient to negate the effect of the missing dependency structures.

6 Conclusion

The present paper investigates whether including Ensemble Copula Coupling (ECC) into different post-processing strategies affects which of the strategies is optimal. We show that ECC, particularly ECC-Q and ECC-T, restores the dependency structures effectively, but this does not affect the forecast performance. The strategies post-processing the weather variables only (One-Step-W) or both the weather variables and the wind power (Two-Step-WP) deliver almost identical results, regardless of whether ECC is used or not. Given these results, we conclude, that ECC does not change the optimal post-processing strategy for wind power forecasts. Due to the smaller number of post-processing steps required and superior or similar forecast accuracy, the strategy post-processing only the resulting wind power (One-Step-P) remains optimal. However, our data does not contain strong dependency structures which limits the potential of ECC. Therefore, we recommend investigating the dependency structures before selecting a post-processing strategy.

Future work should focus on investigating data with different dependency structures to understand when ECC plays a role. Additionally, while the present paper focuses solely on dependencies between weather variables, future work should focus on recreating the spatial and temporal dependencies with ECC and analysing their effect on forecast performance.

Acknowledgements

This work was partly funded by the German Research Foundation (DFG) Research Training Group 2153 “Energy Status Data – Informatics Methods for its Collection, Analysis and Exploitation” and Helmholtz AI. The research detailed in the current paper was based on data from the ECMWF obtained through an academic licence for research purposes.

References

[1] P. Pinson and J. W. Messner, "Application of postprocessing for renewable energy," in *Statistical postprocessing of ensemble forecasts* (S. Vannitsem, D. S. Wilks, and J. Messner, eds.), pp. 241–266, Amsterdam: Elsevier, 2018.

[2] T. L. Thorarinsdottir and T. Gneiting, "Probabilistic forecasts of wind speed: ensemble model output statistics by using heteroscedastic censored regression," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, vol. 173, no. 2, pp. 371–388, 2010.

[3] R. Schefzik, T. L. Thorarinsdottir, and T. Gneiting, "Uncertainty quantification in complex simulation models using ensemble copula coupling," *Statistical Science 2013, Vol. 28, No. 4, 616-640*, 2013.

[4] R. Schefzik, "Ensemble calibration with preserved correlations: unifying and comparing ensemble copula coupling and member-by-member postprocessing," *Quarterly Journal of the Royal Meteorological Society*, vol. 143, no. 703, pp. 999–1008, 2017.

[5] T. Gneiting, *Calibration of medium-range weather forecasts*. 2014.

[6] T. Gneiting, A. E. Raftery, A. H. Westveld, and T. Goldman, "Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation," *Monthly Weather Review*, vol. 133, no. 5, pp. 1098–1118, 2005.

[7] T. Gneiting, K. Larson, K. Westrick, M. G. Genton, and E. Aldrich, "Calibrated probabilistic forecasting at the stateline wind energy center," *Journal of the American Statistical Association*, vol. 101, no. 475, pp. 968–979, 2006.

[8] C. Gilbert, J. W. Messner, P. Pinson, P.-J. Trombe, R. Verzijlbergh, P. Dorp, and H. Jonker, "Statistical post–processing of turbulence–resolving weather forecasts for offshore wind power forecasting," *Wind Energy*, vol. 23, no. 4, pp. 884–897, 2020.

[9] C. Sweeney and P. Lynch, "Adaptive post-processing of short-term wind forecasts for energy applications," *Wind Energy*, vol. 14, no. 3, pp. 317–325, 2011.

[10] A. Bossavy, R. Girard, and G. Kariniotakis, "Forecasting ramps of wind power production with numerical weather prediction ensembles," *Wind Energy*, vol. 16, no. 1, pp. 51–63, 2013.

[11] K. Phipps, S. Lerch, M. Andersson, R. Mikut, V. Hagenmeyer, and N. Ludwig, "Evaluating ensemble post-processing for wind power forecasts," http://arxiv.org/abs/2009.14127v1.

[12] N. Ludwig, S. Arora, and J. W. Taylor, "Probabilistic load forecasting using post-processed weather ensemble predictions," *Journal of the Operational Research Society (submitted)*, 2020.

[13] S. Späth, L. von Bremen, C. Junk, and D. Heinemann, "Time-consistent calibration of short-term regional wind power ensemble forecasts," *Meteorologische Zeitschrift*, vol. 24, no. 4, pp. 381–392, 2015.

[14] M. Taillardat, O. Mestre, M. Zamo, and P. Naveau, "Calibrated ensemble forecasts using quantile regression forests and ensemble model output statistics," *Monthly Weather Review*, vol. 144, no. 6, pp. 2375–2393, 2016.

[15] F. Molteni, R. Buizza, T. N. Palmer, and T. Petroliagis, "The ECMWF ensemble prediction system: Methodology and validation," *Quarterly Journal of the Royal Meteorological Society*, vol. 122, no. 529, pp. 73–119, 1996.

[16] C. Fraley, A. E. Raftery, and T. Gneiting, "Calibrating multimodel forecast ensembles with exchangeable and missing members using bayesian model averaging," *Monthly Weather Review*, vol. 138, no. 1, pp. 190–202, 2010.

[17] T. Gneiting and M. Katzfuss, "Probabilistic forecasting," *Annual Review of Statistics and Its Application*, vol. 1, no. 1, pp. 125–151, 2014.

[18] H. Joe, *Multivariate models and multivariate dependence concepts*. CRC Press, 1997.

[19] R. B. Nelsen, *An introduction to copulas*. Springer Science & Business Media, 2007.

[20] M. Sklar, *Fonctions de Répartition À N Dimensions Et Leurs Marges*. Université Paris 8, 1959.

[21] R. Schefzik, "Ensemble calibration with preserved correlations: unifying and comparing ensemble copula coupling and member-by-member postprocessing," *Quarterly Journal of the Royal Meteorological Society*, vol. 143, no. 703, pp. 999–1008, 2017.

[22] I. Staffell and S. Pfenninger, "Using bias-corrected reanalysis to simulate current and future wind power output," *Energy*, vol. 114, pp. 1224–1239, 2016.

[23] R. Swinbank, M. Kyouda, P. Buchanan, L. Froude, T. M. Hamill, T. D. Hewson, J. H. Keller, M. Matsueda, J. Methven, F. Pappenberger, M. Scheuerer, H. A. Titley, L. Wilson, and M. Yamaguchi, "The TIGGE project and its achievements," *Bulletin of the American Meteorological Society*, vol. 97, no. 1, pp. 49–67, 2016.

[24] Copernicus Climate Change Service (C3S), "Era5: Fifth generation of ecmwf atmospheric reanalyses of the global climate," 2017.

[25] European Network of Transmission System Operators (ENTSO-E), "Transparency platform (tp)," 2019.