



HAL
open science

Probabilistic forecasting of the wind energy resource at the monthly to seasonal scale

Bastien Alonzo, Peter Tankov, Philippe Drobinski, Riwal Plougonven

► **To cite this version:**

Bastien Alonzo, Peter Tankov, Philippe Drobinski, Riwal Plougonven. Probabilistic forecasting of the wind energy resource at the monthly to seasonal scale. *International Journal of Forecasting*, 2020, 36 (2), pp.515-530. 10.1016/j.ijforecast.2019.07.005 . hal-01614920

HAL Id: hal-01614920

<https://hal.science/hal-01614920>

Submitted on 11 Oct 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Probabilistic forecasting of the wind energy resource at the monthly to seasonal scale

Bastien Alonzo^{a,b}, Philippe Drobinski^a, Riwal Plougonven^a, Peter Tankov^c

^a*IPSL/LMD, CNRS, Ecole Polytechnique, Université de Paris-Saclay, Palaiseau, France*

^b*Laboratoire de Probabilités et Modèles Aléatoires, Université Paris Diderot - Paris 7, Paris, France.*

^c*CREST - ENSAE ParisTech, Palaiseau, France*

Abstract

We build and evaluate a probabilistic model designed for forecasting the distribution of the daily mean wind speed at the seasonal timescale in France. On such long-term timescales, the variability of the surface wind speed is strongly influenced by the atmosphere large-scale situation. Our aim is to predict the daily mean wind speed distribution at a specific location using the information on the atmosphere large-scale situation, summarized by an index. To this end, we estimate, over 20 years of daily data, the conditional probability density function of the wind speed given the index. We next use the ECMWF seasonal forecast ensemble to predict the atmosphere large-scale situation and the index at the seasonal timescale. We show that the model is sharper than the climatology at the monthly horizon, even if it displays a strong loss of precision after 15 days. Using a statistical postprocessing method to recalibrate the ensemble forecast leads to further improvement of our probabilistic forecast, which then remains sharper than the climatology at the seasonal horizon.

Keywords: Wind energy, Wind speed forecasting, Seasonal forecasting, Probabilistic forecasting, Ensemble forecasts, Ensemble model output statistics

Email addresses: bastien.alonzo@lmd.polytechnique.fr (Bastien Alonzo),
philippe.drobinski@lmd.polytechnique.fr (Philippe Drobinski),
riwal.plougonven@lmd.polytechnique.fr (Riwal Plougonven),
peter.tankov@ensae.fr (Peter Tankov)

1. Introduction

In the recent years, energy transition has been on the forefront of political and societal issues, mainly due to the increasing awareness of the need to act against the climate disruption. This has led many countries to encourage the use of renewable energy. Since 2008, the European Union (EU) targets 20% of renewable energy contribution to the total energy mix by 2020, and 27% by 2030. Consequently, wind energy has seen a major growth in Europe. To give an idea of this sharp increase, the total installed wind power capacity in the EU has changed from 12.9 GW in 2000 to 141.6 GW in 2015 (EWEA (2016)). The actual share in the final consumption met by wind energy in the EU was 11.4% in 2015 (EWEA (2016)).

At such scales, the variability of the wind power production due to the natural intermittency of the wind resource becomes a critical issue for successful network integration of this source of energy (Albadi and El-Saadany (2010)). As a consequence, the interest and demand for near-surface wind speed forecasts has seen a major boost. Numerous methods exist for forecasting wind speed at different horizons, motivated by different applications (Chang (2014); Soman et al. (2010)). Many studies focus on the short-term scale ranging from several minutes to one day (Carpinone et al. (2015); Gomes and Castro (2012); Stesfos (2002)). Medium-term forecasting methods, ranging typically from 3 days up to 10 days, have also been investigated in depth (Barbounis et al. (2006); Taylor et al. (2009); Wytock and Kolter (2013)). On much longer timescales and with very different motivations, the impact of the climate change on wind speeds has also been addressed (Najac et al. (2009); Pryor and Barthelmie (2010); Sailor and M. Smith (2008)).

Whereas both relatively short and very long timescales have been thoroughly studied, the intermediate timescale going from monthly to seasonal horizon is a research topic for which not so many studies exist. This timescale is becoming very important for the transmission system operators (TSOs) as the proportion of intermittent resources in the energy mix increases. The TSOs are responsible for balancing the supply and demand of energy and they are required to make seasonal projections, e.g., to guarantee the security of energy supply during the coming winter, which becomes more difficult with the the increased variability of energy production. The risk of not being able to satisfy the energy demand may be quantified in terms of the notion of Loss of load expectation (LOLE). Quoting from NationalGrid (2016), the LOLE is a “measure of the risk across the whole winter of demand exceeding

38 supply under normal operation. It gives an indication of the amount of time
39 across the whole winter that the System Operator may need to call on a
40 range of emergency balancing tools to increase supply or reduce demand.”
41 For instance, a cold winter characterized by weaker winds than normal may
42 in some cases lead to a lack of energy if not enough other production means
43 have been made available upstream to meet the energy demand.

44 Among the few existing studies of long term wind speed forecasting, (Azad
45 et al. (2014); Bilgili et al. (2007)) advocate the use of Artificial Neural Net-
46 works (ANN) for forecasting average monthly wind speed. These studies give
47 an accurate estimate of the trend of the wind speed at the yearly horizon
48 but provide limited information on the wind variability at higher frequencies.
49 ANN models have also been used for forecasting daily mean wind speed at
50 the seasonal scale providing more information on the wind variability within
51 a given season for energy production evaluation (Guo et al. (2012); J. Wang
52 et al. (2015); More and Deo (2003)).

53 These studies provide 'point forecasts', which give one value for the wind
54 energy production at the specified horizon, but do not consider the uncer-
55 tainty on the forecast (as a rule, forecast uncertainty is difficult to quantify
56 with neural networks since the underlying probabilistic model is not easy
57 to define). At such timescales, the idea of point forecast can be very ques-
58 tionable due to the dominant chaotic nature of the atmospheric system at
59 the timescales exceeding typically 10 days. At this long-term horizon, the
60 idea of probabilistic forecasting therefore gains sense. Indeed, forecasting
61 centers such as the European Center for Medium-range Weather Forecasts
62 (ECMWF) use ensemble forecasts to take into account the uncertainty due
63 to the growing small errors in the atmospheric system. Surface wind speed
64 is a variable that is not provided by such prediction models because of its
65 complexity and interaction with the surface, but valuable information on
66 the general circulation of the atmosphere can be retrieved from such fore-
67 casts. Indeed, several works have confirmed the predictability at seasonal
68 timescales of recurrent oscillating patterns in the atmosphere, such as the El
69 Nino (Cassou (2008); Owen and Palmer (1987)) or the North Atlantic Os-
70 cillation (NAO) (Davies et al. (1997); Rodwell et al. (1999)). On the other
71 hand, large-scale atmospheric patterns have already been shown to partly
72 explain the surface wind speed in France at such timescales (Alonzo et al.
73 (2017)).

74 In this paper, our aim is to use long-term forecasts of large-scale cir-
75 culation patterns provided by ECMWF to obtain probabilistic long-term

76 forecasts of local surface wind speeds. Our approach is to build a prob-
77 abilistic model describing the relationship between the local surface wind
78 speed and the large-scale circulation of the atmosphere, summarized by a
79 single purpose-built index. To this end we estimate the conditional proba-
80 bility density function of the wind speed given the index by gaussian kernel
81 density estimation over 20 years of daily data. We next use the ECMWF
82 seasonal forecast ensemble to predict the large-scale situation of the atmo-
83 sphere and the index at the seasonal timescale. The prediction of the index is
84 then plugged into our model, to obtain probabilistic forecasts of the surface
85 wind speed. The ensemble forecast displays a growing uncertainty with time
86 leading to an increase of the confidence interval width predicted by the prob-
87 abilistic model. We show that the model is sharper than the climatology at
88 the horizon of one month, even if it displays a strong loss of precision after 15
89 days. Using the statistical postprocessing method EMOS (Ensemble Model
90 Output Statistics) to recalibrate the ensemble forecast leads to a further im-
91 provement of our probabilistic forecast, which then remains sharper than the
92 climatology at the seasonal horizon.

93 This paper is structured as follows. Section 2 describes the method to
94 build the probabilistic model as well as the data used in this study. In section
95 3, the performance of the model is assessed. In section 4, the probabilistic
96 model is used to forecast the wind speed at the monthly and seasonal hori-
97 zon by applying it to seasonal ensemble forecasts of large scale circulation
98 patterns of the atmosphere.

99 **2. Data & methods**

100 *2.1. Data : ECMWF reanalysis and forecasts*

101 In this paper, we use the so called “perfect model” approach meaning that
102 the ECMWF ERA-I reanalysis is considered as the reality. This is justified
103 by the comparison of ECMWF products and observation, in particular for
104 surface wind speed (Jourdier (2015)). The model is estimated and evaluated
105 the surface wind speed retrieved from this data. We use ECMWF reanalysis
106 for 37 years between 1 January 1979 and 31 December 2015¹.

107 The basic idea of this work is to link the large-scale circulation of the at-
108 mosphere with the daily mean surface wind speed distribution in France. The

¹ECMWF Data are available at <http://apps.ecmwf.int/datasets/>

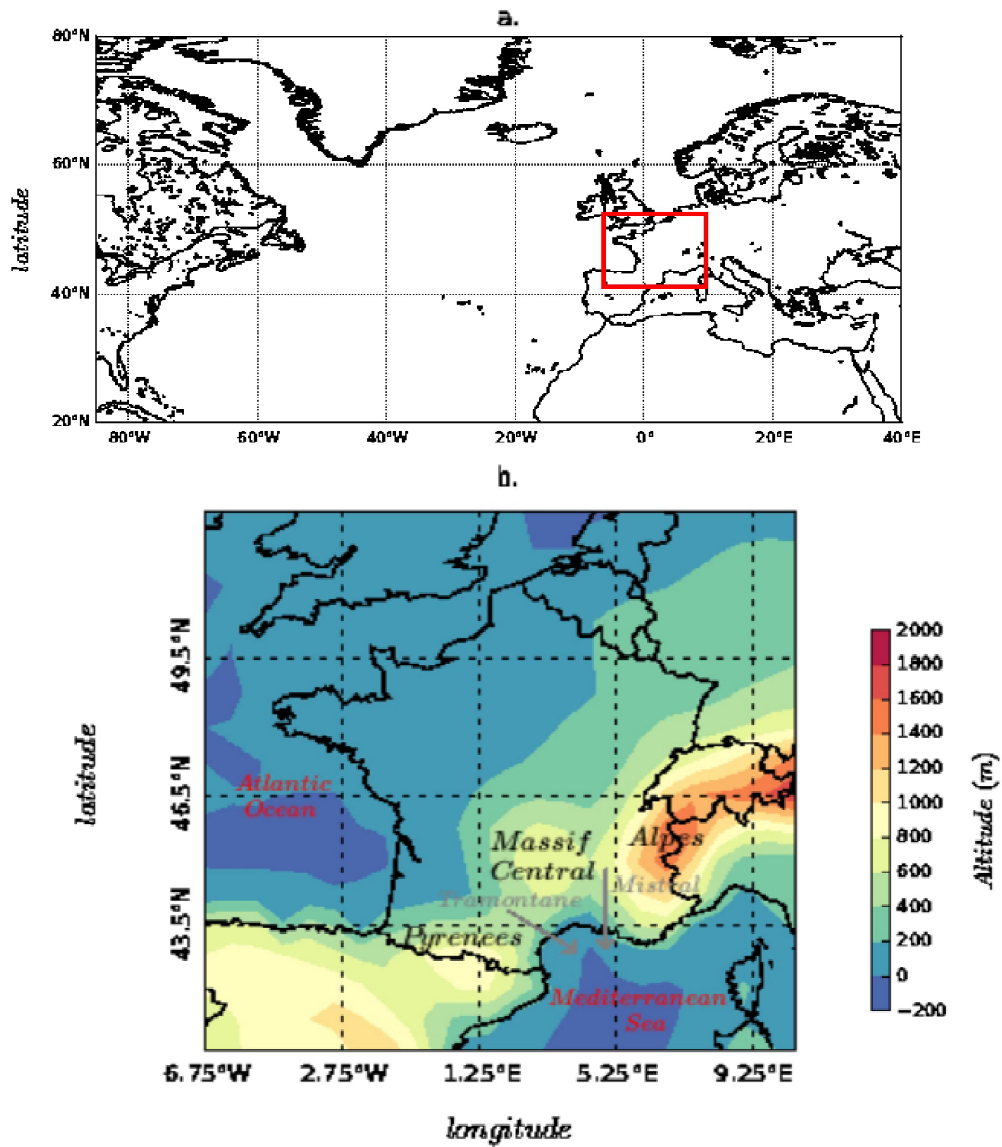


Figure 1: a. Domain used to perform the PCA on the Z500. The red box corresponds to the domain covering France on which the model is built and assessed. b. Domain covering France and part of its neighbouring countries. The colors represent the altitude above the sea level.

109 large-scale circulation is well described by the 500-hPa geopotential height
 110 (Z500) over the North Atlantic/European region (Michelangeli et al. (1995)).

111 We therefore retrieve from the ECMWF reanalysis the daily time series of
112 500-hPa geopotential height (Z500) over a large domain that spans over North
113 Atlantic Ocean and Europe (20°N to 80°N and 90°W to 40°E), (Figure 1,
114 a), with grid size of 0.75°. The daily surface wind speed used to build and
115 evaluate the probabilistic model is also retrieved from ECMWF reanalysis.
116 The data spans the same period, but over a smaller domain which covers
117 France and parts of neighbouring countries (40.5°N to 52.5°N and -6.75°W
118 to 10.5°E), (Figure 1, b).

119 The 37 years of data are split into three periods. The first 20-year period
120 (1 January 1979 to 31 December 1998), is used to build and estimate the
121 model. This period is referred to as the fitting period. On the subsequent
122 13-year period (1 January 1999 to 31 December 2011), the probabilistic model
123 is evaluated and compared to the past seasonal climatology of the wind speed,
124 considered as the benchmark for wind speed forecasting at such long-term
125 horizon. The seasonal climatology is defined as the empirical distribution of
126 the daily average wind speed computed over all days in a given season of the
127 fitting period. This period is referred to as the validation period. The results
128 of the validation of our model are described in section 3.

129 On the 4 remaining years, we use the model to build probabilistic fore-
130 casts of the surface wind speed at the seasonal horizon. 48 ECMWF seasonal
131 ensemble forecasts of the Z500 field over the large domain are retrieved. Sets
132 of forecasts are retrieved from 2012 to 2015, beginning on every first day of
133 each month. A major change of the assimilation system and forecast model
134 limits the use of seasonal forecasts before November 2011. Seasonal forecasts
135 provide a prediction of the Z500 at more than three months horizon, allow-
136 ing to predict the surface wind speed at either monthly or seasonal horizon.
137 The seasonal ensemble forecasts consist of 41 members. Each member has
138 a slightly different initial state, so that the uncertainty on the atmospheric
139 circulation grows with the forecast horizon giving a range of different possi-
140 ble states of the atmosphere. The forecasting performance of our model is
141 analyzed in section 4.

142 *2.2. Statistical methods*

143 In the first step of building our model we apply the Principal Component
144 Analysis (PCA) to the Z500 variable to reduce its dimension. The outputs
145 of the PCA are the Empirical Orthogonal Functions (EOF) describing the
146 prevalent spatial patterns in the data, and the associated Principal Compo-
147 nents (PC) time series which show how the state of the atmosphere projects

148 onto these patterns. The first EOFs may be identified with the classical
 149 large-scale weather patterns (NAO, SCA, ...) which control the European
 150 climate variability (Casanueva et al. (2014); Folland et al. (2008); Wallace
 151 and Gutzler (1980)). We expect the PCs to be well predicted in the seasonal
 152 ensemble forecasts.

153 In the second step, we build a model giving the probability distribution
 154 of the daily mean wind speed knowing the first n PCs. In other words, we
 155 want to compute the conditional density $p(y|X_1, \dots, X_n)$ of the daily mean
 156 surface wind speed Y given the PCs of Z500 X_1 to X_n . Computing this
 157 conditional density directly is difficult due to the high dimension of the vector
 158 (X_1, \dots, X_n) . To overcome this issue, we use the single index approximation
 159 (Delacroix et al. (2003)): we assume that the information about the PCs
 160 (X_1, \dots, X_n) may be summarized by a single scalar index

$$I = \beta_0 + \sum_{i=1}^N \beta_i X_i + \sum_{i=1}^N \beta_{ii} X_i^2 + \sum_{i=1}^{N-1} \sum_{j>i}^N \beta_{ij} X_i X_j, \quad (1)$$

161 where the coefficients β_0 , β_i and β_{ij} are computed by least-squares regression
 162 of the surface wind speed Y on the principal components X_1, \dots, X_n for each
 163 location. A test of optimization of the index parameters β_i by minimization
 164 of the continuous ranked probability score (CRPS – see below) has been
 165 performed at several locations, but did not produce a significant improvement
 166 (only of the order of 0.1% of the initial CRPS).

167 The conditional probability density function $p(y|I)$ is given by the stan-
 168 dard formula

$$p(y|I) = \frac{p(y, I)}{p(I)}, \quad (2)$$

169 where $p(y, I)$ is the joint density of the surface wind speed Y and the index I
 170 and $p(I)$ is the marginal density of the index. A gaussian kernel density esti-
 171 mator (KDE) is used to estimate the joint density and the marginal density
 172 over the period of length T :

$$\hat{p}(y|I = i) = \frac{\sum_{t=1}^T K_{h_1}(y - Y_t) K_{h_2}(i - I_t)}{\sum_{t=1}^T K_{h_2}(i - I_t)}, \quad (3)$$

173 where K_h is the gaussian kernel function written as :

$$K_h(x) = \frac{1}{h\sqrt{2\pi}} \exp\left(-\frac{x^2}{2h^2}\right). \quad (4)$$

174 While the estimated density is not very sensitive to the choice of the kernel
175 function, the bandwidth parameters h_1 and h_2 have a significant impact
176 on the resulting probability density function. In our study, the bandwidth
177 parameters have been computed by cross-validation.

178 3. Evaluation and optimization of the model

179 3.1. Criteria for model evaluation: calibration and sharpness

180 The performance of a probabilistic forecasting model is typically assessed
181 in terms of calibration and sharpness (Carney and Cunningham (2006); Fos-
182 ter and Vohra (1998); Gneiting et al. (2007); Thorarinsdottir (2013)). While
183 calibration refers to the statistical consistency between the model and the
184 actual values of the variable to predict, sharpness is a property of the model
185 only and measures the width of the confidence intervals. Different modes
186 of calibration exist and must be considered for the model to be fully cali-
187 brated. In the following, we evaluate probabilistic calibration and marginal
188 calibration. Consider a probabilistic forecast at time t in the form of a pre-
189 dictive distribution function $F_t(x)$, and corresponding to the realization x_t .
190 Probabilistic calibration (Gneiting et al. (2007)) measures the compatibility
191 of the probabilistic forecast $F_t(x)$ with the actual realization x_t by means of
192 the probability integral transform (PIT) defined by $p_t = F_t(x_t)$. The fore-
193 cast is said to be probabilistically calibrated if the PIT follows a uniform
194 distribution.

195 On the other hand, the marginal calibration (Gneiting et al. (2007))
196 compares the long-run distribution of the probabilistic forecast $\overline{F}(x) :=$
197 $\frac{1}{T} \sum_{t=1}^T F_t(x)$ to the long-run (climatological) distribution of the data, pro-
198 vided that the data is stationary. In meteorological terms, the assumption
199 of stationarity of the data corresponds to the common assumption of the
200 existence of a stable climate.

201 To evaluate the model performance we consider as the benchmark the
202 seasonal climatology, which is often used within the wind energy industry
203 for such long-term wind energy prediction (Pinson and Kariniotakis (2009)).
204 Indeed, the persistence and autocorrelation of the wind disappear after 5 days
205 at most so that the seasonal pattern is the only information that remains in
206 absence of additional data.

207 *Probabilistic calibration.*

208 Probabilistic calibration is assessed using the Probability Integral Transform (Gneiting et al. (2007)). By applying, at each time step, the predicted

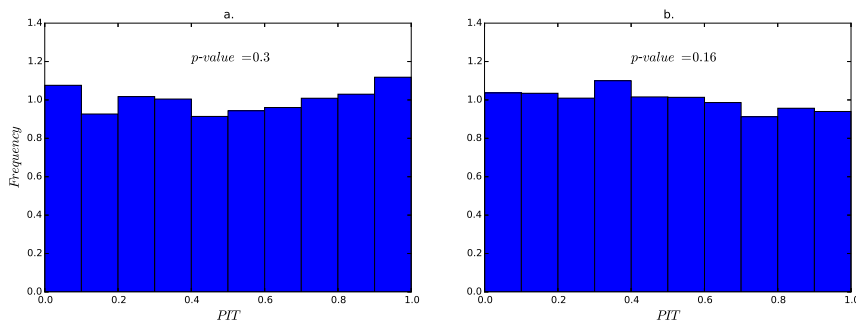


Figure 2: Example of a PIT histogram for one point in France (49.5°N/2.25°E), for the model (a.) and the climatology (b.). The p-value of the KS test performed on the 3 days sampled PIT is indicated. In this particular example the null hypothesis of uniformly distributed PIT is not rejected at the 5% confidence interval for neither the model nor the climatology.

209

210 Cumulative Distribution Function (CDF) $F_t(\cdot)$, to the actual value y_t , we
211 obtain a sample $(F(y_t))_{t=1}^T$ of values in $[0, 1]$, which must follow a uniform
212 distribution on $[0, 1]$ if the forecast is probabilistically calibrated. Uniformity
213 of the PIT can be evaluated visually by plotting its histogram, usually
214 referred to rank histogram in meteorology, (Figure 2), or more rigorously by
215 performing a Kolmogorov-Smirnov (KS) test on the sample. The KS test
216 is to be performed on independent and identically distributed random vari-
217 ables. Hamill (Hamill (2000)) shows that the correlated errors of samples can
218 lead to misinterpretation of the PIT while testing uniformity. The samples
219 thus have to be spaced far enough in space and time to be reasonably close
220 to being independent. As the daily mean wind speed is autocorrelated up to
221 time scales of about 3 to 5 days, so is the PIT. Figure 3 shows the autocor-
222 relation of the entire sample of the PIT (Fig 3 a) and of the PIT resampled
223 every 3 days (Fig 3 b). After 3 days, the sampled PIT shows little or no
224 autocorrelation and the KS test is thus performed on a 3 day sampled PIT.

225 *Marginal calibration.*

226 Marginal calibration can be seen as a way to ensure that the actual cli-
227 matology of the wind speed over the validation period is well represented by

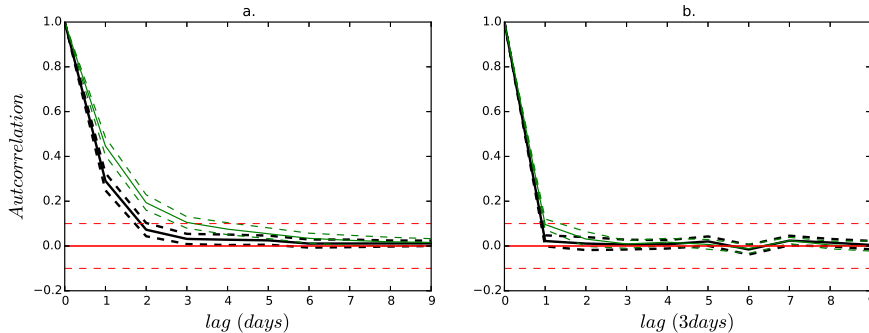


Figure 3: Autocorrelation of the PIT of the model (bold black curves) and the climatology (green curves), as function of the time lag : for the entire PIT sample (a) and for the PIT sampled every 3 days (b). The straight line represents the spatially averaged autocorrelation, and the dashed lines represent the 20th and the 80th percentiles.

228 the model. Actual climatology refers here to the probability density function
 229 of the wind speed over the validation period (and should not be confused
 230 with the past seasonal climatology computed on the fitting period, taken as
 231 a predictive distribution of reference). Marginal calibration can be assessed
 232 visually by plotting the difference between the climatological CDF on the
 233 validation period and the mean predicted CDF given by the model.

234 Figure 4 shows the marginal calibration computed using the probabilistic
 235 model (black dashed line) and the past seasonal climatology (black solid
 236 line) on the validation period at one grid point in the center of France
 237 (49.5°N/2.25°E).

238 To highlight the fact that part of the deviation comes from the statistical
 239 variations of the samples, we generate fifty random samples from the distribu-
 240 tion of the actual wind speed over the 17 years of validation period (actual
 241 climatology) estimated by KDE. From these random samples we compute
 242 fifty different resampled actual climatologies, and calculate the difference
 243 between the distribution obtained for each one of them and that of the ac-
 244 tual climatology. The red solid line represents the mean difference between
 245 the actual climatology and resampled actual climatologies, and red dotted
 246 lines represent the 20th and 80th percentiles. We see that for this particu-
 247 lar point the curve corresponding to the past seasonal climatology is outside
 248 this bootstrap-style confidence interval, while the curve corresponding to our
 249 probabilistic forecast is well inside it.

250 In order to visualize marginal calibration on a map, we compute at each

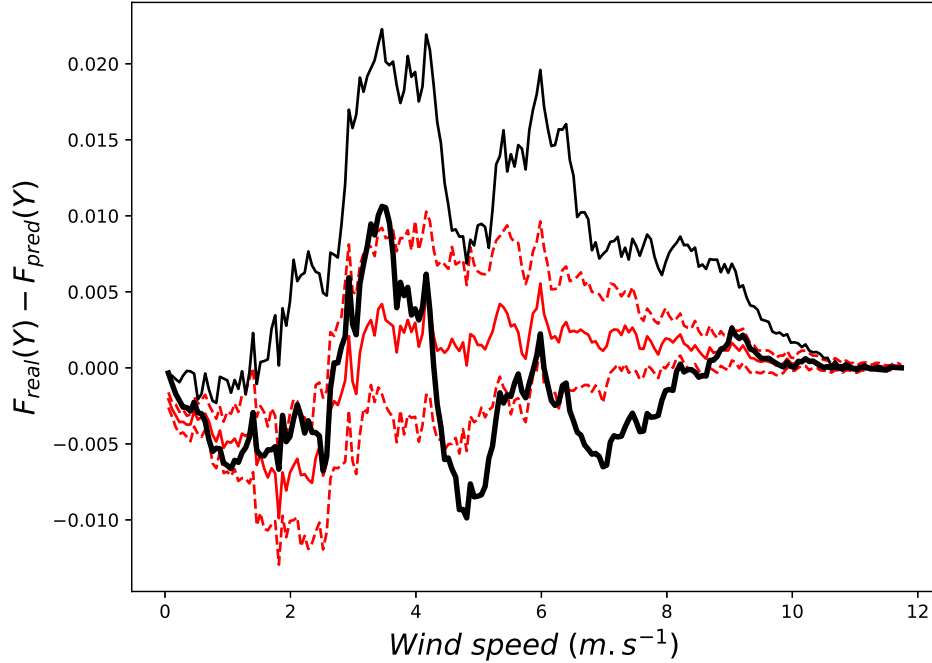


Figure 4: Difference between the actual CDF (actual climatology) on the validation period and : the mean predicted CDF (bold black solid line), the mean past seasonal climatology (black solid line), the mean of the resampled actual climatologies (red solid line), the 20th and 80th percentiles of the resampled actual climatologies (red dotted lines) for one point in France (49.5°N/2.25°E).

251 grid point the Mean Absolute Errors (MAE) between those distributions.
 252 MAE is calculated following the equation.

$$MAE = \int_{-\infty}^{\infty} |F_{real}(Y) - F_{pred}(Y)| dy \quad (5)$$

253 The model is considered marginally calibrated if the computed MAE defined
 254 above is less than the 95th percentile of the MAE computed for the so called
 255 resampled actual climatologies.

256 *Sharpness.*

257 Sharpness refers to the width of the predictive distribution, that is to
 258 say, the accuracy of the forecast. Confidence interval widths are therefore

259 good diagnostics of the sharpness of a probabilistic forecasting model. In this
 260 paper, the 90% confidence interval width is used as measure of sharpness.

261 *Continuous Ranked Probability Score (CRPS).*

262 The Continuous Ranked Probability Score (CRPS) is a widely used scor-
 263 ing rule in meteorological probabilistic forecasts (Candille et al. (2007); Can-
 264 dille and Talagrand (2005)). It aims to evaluate both calibration and sharp-
 265 ness simultaneously. The CRPS for a single predictive distribution F and
 266 realization y_t is defined by:

$$CRPS(F, y_t) = \int_{-\infty}^{\infty} (F(y) - 1_{(y \leq y_t)})^2 dy \quad (6)$$

267 with $1_{(y \leq y_t)}$ being defined as :

$$1_{(y \leq y_t)} = \begin{cases} 1, & \text{if } y \geq y_t. \\ 0, & \text{otherwise.} \end{cases} \quad (7)$$

For the entire sample of size T we define the CPRS by

$$CPRS = \frac{1}{T} \sum_{t=1}^T CPRS(F_t, Y_t).$$

268 3.2. Optimization of the model

269 In this section we discuss the choice of the number of principal compo-
 270 nents to be used in the model. By adding more PCs, the variability of the
 271 large scale circulation is better accounted for, but too many PCs can also
 272 lead to overfitting and thus poor calibration of the model. Depending on the
 273 region, the optimal number of PCs can be estimated. For example, although
 274 the onshore wind variability can be partially explained by the large-scale at-
 275 mosphere circulation, smaller scale phenomena such as topography effects,
 276 can have a significant influence on the wind speed. Conversely, offshore wind
 277 speed is more regular and obviously not impacted by orography so that large-
 278 scale atmosphere circulation is the main driver of its variability at those long
 279 timescales.

280 To determine the optimal number of PCs we increase their number from
 281 5 to 30 with an increment of 5 (which corresponds to 6 different models)
 282 and evaluate the probabilistic and marginal calibration and sharpness on the

283 validation period of 17 years for the model based on the index computed
 284 with the corresponding number of PC.

285 Unexpectedly, marginal calibration shows no significant variation depend-
 286 ing on the number of PC (no more than 10% of the statistical error), probably
 287 because of the CDF averaging effect. Conversely, probabilistic calibration
 288 and sharpness show high sensitivity to the number of PC (Figure 5). Unfor-
 289 tunately, on average, adding PCs sharpens the model, but also decalibrates
 290 it (Figure 5 a and c).

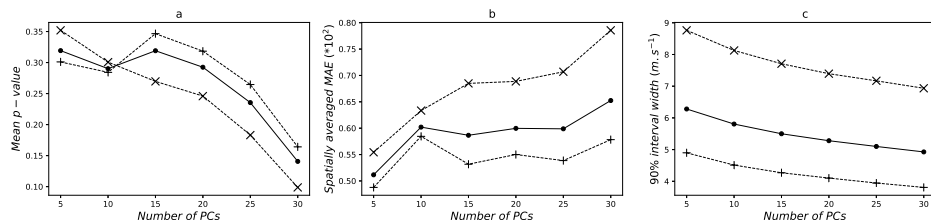


Figure 5: **a.** Spatially averaged p-value of the KS test performed on 3 days sampled PIT used to assess probabilistic calibration ; **b.** Spatially averaged MAE between actual climatology and the predicted climatology given by the model used to assess marginal calibration ; **c.** Spatially averaged 90% confidence interval width used to assess sharpness. All three graphs are plotted as function of the number of PCs used to fit the index. The black line with point markers is the average over the entire domain, the black dotted line with 'x' markers is the average over the offshore part of the domain, and the black dotted line with '+' markers over the onshore part of the domain.

291 In our final model, we use the following methodology to choose the opti-
 292 mal number of PCs for each location. We first test the null hypothesis that
 293 the PIT follows a uniform distribution with a 95% confidence level using the
 294 3 days sampled PIT for each model. If the hypothesis is not rejected for
 295 any of the 6 models corresponding to different numbers of PCs, we keep the
 296 model which maximizes the sharpness. If the null hypothesis is rejected for
 297 all 6 models, we keep the model that maximizes the p-value of the KS test,
 298 with the risk to have a non-calibrated model.

299 Figure 6 shows the result of the choice described above. Over the northern
 300 half of the domain and along the western coast of France, a large number of
 301 PC (> 15) is required to build the index meaning that the variability can be
 302 explained by shorter scale phenomena without compromising the calibration
 303 quality of the model. This results in a sharper model than when using less
 304 PCs. Conversely, for offshore wind taking a large number of PCs reduces the
 305 calibration quality.

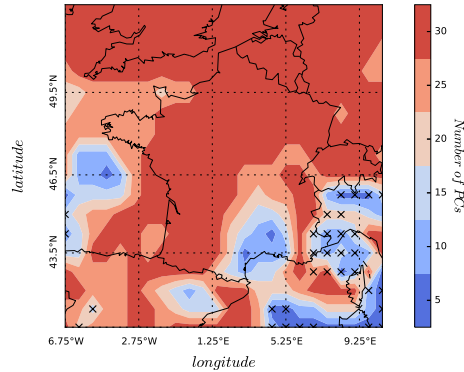


Figure 6: Optimal number of PCs used to fit the index of the model determined using the optimization process described in the text. 'x' markers show points where the model is not calibrated (Figure 7).

306 In the southeast of France, over the Mediterranean coast and the sea, we
 307 can find a clear signature of the orography. Offshore, the Mistral, which refers
 308 to the strong wind blowing over the Mediterranean sea after being channeled
 309 in the valley formed by the Alps and the Massif Central (Drobinski et al.
 310 (2017)), can be identified by an intermediate number of PC (20-30). The
 311 Tramontane also refers to an orographic wind blowing over the same region
 312 but channeled in the valley formed by the Pyrenees and the Massif Central
 313 (Brossier and Drobinski (2009)) (Fig 1 b). South of the Alps, the model is
 314 not calibrated, and south of the Massif Central only 5 PCs are used resulting
 315 in model that is less sharp.

316 3.3. Evaluation of the optimized model

317 Figure 7 shows the results of the KS test performed on the 3 days sampled
 318 PIT given by the optimized model (Fig 7 a and b) and the climatology (Fig
 319 7 c and d). The p-value for the climatology ranges between 0 and 0.8, while
 320 it ranges between 0 and 0.5 for the model. The null hypothesis of adequate
 321 calibration is not rejected in the North part of the domain for the model,
 322 while for the climatology this hypothesis is rejected over the North part of
 323 the domain. The climatology does not represent the law of the wind well
 324 in those regions but the probabilistic model represents it quite well (Fig
 325 7 b and d). This can be surprising as the climatology is built using 20
 326 years of data which may seem to be sufficient to ensure calibration over a

327 period of the same length. Nevertheless, it has been shown that annual wind
328 trends can be significant over 1 to 2 decades in this region (Jourdier (2015)).
329 Using only the past five years of wind speed data to build the empirical
330 seasonal CDF allows to follow those trends. This sliding CDF displays a
331 null hypothesis of adequate calibration which is not rejected over the entire
332 domain. Nevertheless, it performs as well as the seasonal climatology in terms
333 of sharpness (*Not shown*). In the South of the domain, the model and the
334 climatology perform similarly in terms of probabilistic calibration showing
335 large non-calibrated areas. Indeed, the region is very complex and strongly
336 influenced by orography. This complexity seems to be hard to recover with
337 the information on the season only (climatology) or the information on the
338 large-scale circulation (model).

339 Figure 8 shows the MAE between the real climatological CDF over the 13-
340 year validation period and, on the one hand, the averaged CDF predicted by
341 the model (Fig 8 a.) and on the other hand the climatological CDF based on
342 the 20 year fitting period (Fig 8 b). We can clearly see a strong correlation
343 between marginal calibration and probabilistic calibration. The model is
344 considered marginally calibrated if the computed MAE is inferior to the
345 95th percentile of the MAE computed for the resampled actual climatologies.
346 Applying this criterion to MAE computed for the model and for the past
347 seasonal climatology gives a map (not shown) which is very similar to those
348 in Figures 7c and 7. If the model or the climatology is probabilistically
349 calibrated in a given location, it is also marginally calibrated there. Overall,
350 for both calibration criteria, the calibration of the model is at least as good
351 as that of the climatology and often much better.

352 Figure 9 displays the 90% confidence interval width averaged over the vali-
353 dation period for the model (IC_{90mod}) (Fig 9 a) and the climatology (IC_{90clim})
354 (Fig 9 b), and the ratio of IC_{90clim} to IC_{90mod} (Fig 9 c). The first striking
355 observation is that the 90% confidence interval is much larger offshore than
356 onshore for both the model and the climatology. This highlights the fact that
357 even if the wind may be more regular, it can also be much stronger because
358 of the low roughness, so that the difference between weak and strong wind
359 events is by far larger than onshore. The signature of the Mistral and Tra-
360 montane is clear (Fig 9 a, b), with even larger interval width that may come
361 from the bimodal distribution of the wind speed in this region (Drobinski
362 et al. (2015)).

363 Over the entire domain, on average, the model is sharper than the clima-
364 tology. The model does not perform more than 50% better than the clima-

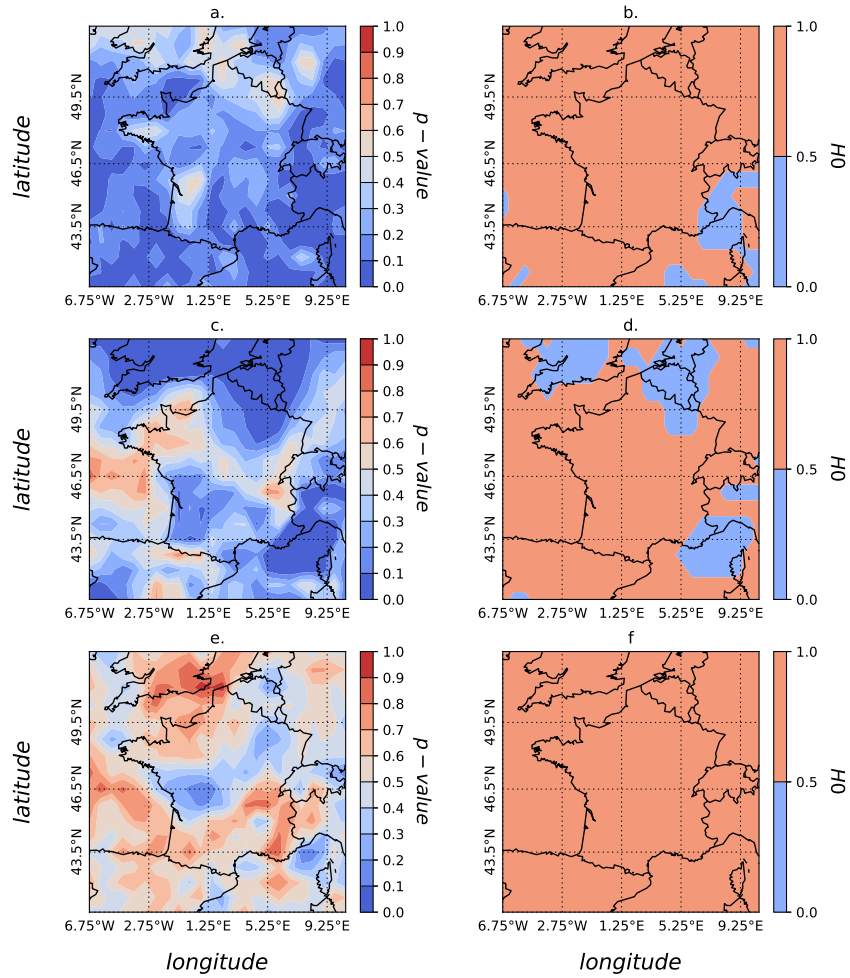


Figure 7: Left graphs: p-value of the KS test performed on the 3 days sampled PIT of the model (a.) and the climatology (b.) and for the empirical seasonal CDF based on the last five years of wind speed (e). Right graphs: the blue area (0 value) shows the regions where the null hypothesis of adequate calibration is rejected for the model (b) and the climatology (d) and for the empirical seasonal CDF based on the last five years of wind speed (f).

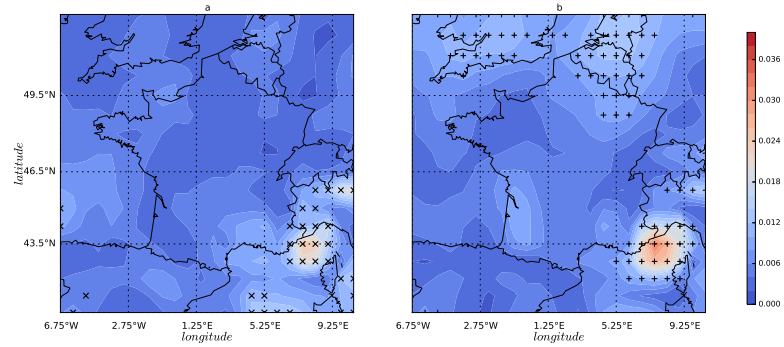


Figure 8: Mean Absolute Error between the real climatological CDF over the 15 years of validation and the averaged CDFs predicted by the model (a.) and the climatological predictive CDFs based on the 20 years of calibration period (b.). 'x' marker on panel a. and '+' markers on panel b. show respectively points where the model and the climatology are not probabilistically calibrated.

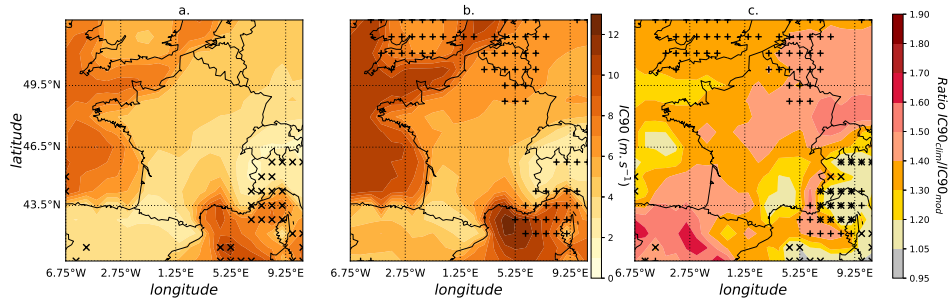


Figure 9: 90% confidence interval width averaged over the validation period, for the model (a.) and the climatology (b.). Panel c. displays the ratio of the confidence interval width of the climatology over the model. 'x' and '+' markers indicate places where respectively the model or the climatology are not calibrated.

365 tology, except in the northeast of the domain. Over the north of France, the
 366 model is sharper than the climatology by more than 40% which is encourag-
 367 ing because of the high wind energy potential in those regions. Unfortunately,
 368 over the west Atlantic ocean, the model is not as sharp as expected compared

369 to climatology, but still, it performs 20% to 30% better. Over the south of
 370 France, in addition to the bad calibration of the model, its performance in
 371 terms of sharpness is not as spectacular as in other regions. Again, this can
 372 be due to the complexity of the wind variability in this region.

-	-	All seasons	Winter	Spring	Summer	Fall
Model	Mean IC90	5.2	5.5	5.1	4.8	5.4
	σ	2.1	2.3	2.0	1.9	2.2
Climatology	Mean IC90	6.9	7.9	6.4	5.8	7.7
	σ	2.4	2.7	2.1	2.1	2.6
Ratio	IC_{90clim}/IC_{90mod}	1.3	1.4	1.2	1.2	1.4

Table 1: 90% confidence interval width (IC90) ($m.s^{-1}$) averaged on the validation period and on the whole domain, for all seasons, and every season separately, for the model and the climatology.

373 By averaging the interval width separately for each season, we can high-
 374 light a strong seasonal variability of the interval given by the climatology,
 375 which is not so noticeable for the model (Table 1). Thus, the model shows
 376 even better performance compared to the climatology in winter and fall (40%
 377 sharper than climatology on average over the domain) which are the seasons
 378 when the risk of high LOLE may be larger because of low temperature. The
 379 model is 80% sharper than the climatology in the northeast regions in winter
 380 and fall (*not shown*). Differences between land and sea are present for all
 381 seasons, and the Mediterranean region is always more problematic.

382 The CRPS of the model and the climatology should inform us on both
 383 the calibration and sharpness. It is expressed in the same units as the pre-
 384 dicted quantity ($m.s^{-1}$ in the case of wind speeds) and reduces to the MAE
 385 for point forecasts. Figure 10 shows the mean CRPS on the validation pe-
 386 riod, for the model (Fig 10a), the climatology (Fig 10b) and the ratio of
 387 $CRPS_{clim}/CRPS_{mod}$ (Fig 10c). All panels are very comparable to those of
 388 Figure 9. Even if there is no doubt that the CRPS addresses both calibration
 389 and sharpness, on average, it appears to put too much weight on sharpness.
 390 For instance, over the Alps, the model is not calibrated, nevertheless the
 391 CRPS has very low values, indicating that the model has good performance.
 392 The CRPS values do display a very significant difference between land and
 393 sea, and so does the confidence interval width. It is thus clear by comparing
 394 figures 9 and 10 that the average CRPS is informative about the sharpness
 395 of the model more than about its calibration quality.

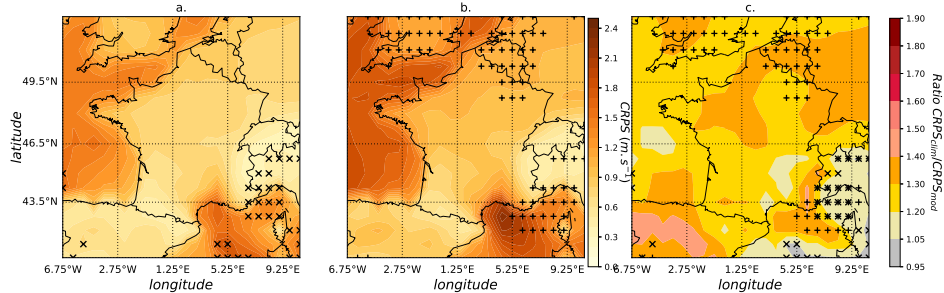


Figure 10: CRPS of the model (a.) and the seasonal climatology (b.). 'x' marker on panel a. and '+' markers on panel b. show respectively points where the model and the climatology are not calibrated. c. is ratio of the CRPS

396 4. Forecasting the wind at the monthly and seasonal horizon

397 4.1. Methodology

398 To make monthly / seasonal forecasts with our model, we must take into
 399 account the uncertainty of the Z500 forecast, and thus also of the index. The
 400 seasonal ensemble forecasts of ECMWF are based on 41 members displaying
 401 a large range of possible Z500 fields. For each member, we first calculate
 402 the values of the principal components by projecting the corresponding Z500
 403 field onto the EOFs identified during the stage of model calibration. Next,
 404 for each member of the ensemble forecast, and for each location where sur-
 405 face wind speed forecast is needed, we compute the corresponding index
 406 value using equation (1), where the coefficients β_i were identified during
 407 the stage of model calibration. This gives us an ensemble of index values
 408 I_1, \dots, I_n . From this ensemble we construct the predictive distribution of in-
 409 dex values, denoted by μ . This can be done in two different ways. The first
 410 method (raw forecast) consists in taking simply the empirical distribution of
 411 I_1, \dots, I_n , that is, $\mu = \frac{1}{n} \sum_{k=1}^n \delta_{I_k}$, where δ_x is the point mass at point x .
 412 The second method uses statistical post-processing of the ensemble forecast
 413 to construct a distribution μ with better calibration / sharpness properties
 414 than the raw forecast. In this paper, we use the Ensemble Model Output
 415 Statistics (EMOS) method, described below, for forecast post-processing.

416 Once the predictive distribution for the index has been constructed, the

417 density of the predictive distribution for the surface wind speed given the
 418 forecast $p(y|F)$ is obtained by integrating the density of the conditional dis-
 419 tribution of the wind speed given the index with respect to the predictive
 420 distribution of the index:

$$p(y|F) = \int_{-\infty}^{\infty} p(y|I = x)\mu(dx), \quad (8)$$

421 This should produce a less sharp model with a higher chance to be calibrated
 422 than if only the mean of the forecast ensemble is used.

423 *Ensemble Model Output Statistics - EMOS.*

424 To recalibrate and sharpen a forecast ensemble different statistical post-
 425 processing methods exist such as the Bayesian Model Averaging (BMA)
 426 (Möller et al. (2013); Raftery et al. (2005); Slughter et al. (2013)) or the En-
 427 semble Model Output Statistics (EMOS) (Gneiting et al. (2005); N.Schuhen
 428 et al. (2012); Thorarinsdottir and Gneiting (2010)). EMOS aims at recal-
 429 ibrating the distribution of ensemble forecasts, but also at sharpening it. This
 430 method is inspired by Gneiting et al. (2005) apart from the optimization al-
 431 gorithm. This method is based on the assumption that μ has a normal
 432 distribution $N(m_I, \sigma_I)$, where m_I is a weighted linear combination of the
 433 index values of the ensemble,

$$m_I = b_0 + \sum_{m=1}^n b_m I_m, \quad (9)$$

434 and σ_I is parameterized by

$$\sigma_I = c + d \text{Var}(I), \quad (10)$$

435 where $\text{Var}(I)$ is the empirical variance of the ensemble.

436 The parameters of the EMOS method b_0, \dots, b_n, c and d are estimated
 437 as follows. In the first step of the estimation procedure, on the training
 438 period, set to three years in this study, we perform a linear regression of
 439 the index I computed from the actual ERAI-reanalysis on the index values
 440 I_1, \dots, I_n computed from the ECMWF seasonal forecasts. This gives us
 441 a first estimate of b_0, \dots, b_n . In this first step we set $c = 0$ and $d = 1$.
 442 Then, in the second step, we improve the first-step estimates by minimizing
 443 the Continuous Ranked Probability Score (CRPS) of the forecasts, averaged

444 over the training period, seen as function of the parameters b_0, \dots, b_n, c and
445 d using the Powell algorithm (Powell (1964)).

446 In the end, we obtain a set of parameters b_1, \dots, b_m, c and d that min-
447 imize the CRPS score. The minimization of the CRPS must optimize the
448 calibration and the sharpness. We apply the obtained parameters on the
449 remaining year of forecasts to estimate the gaussian distribution $N(m_I, \sigma_I)$
450 of the index and then integrate over this distribution as in eq (8). The pro-
451 cedure is repeated 4 times by training on three different years and testing
452 on the remaining year. This results in 48 EMOS forecasts of the daily mean
453 wind speed distribution at the seasonal horizon.

454 4.2. Results

455 Figure 11 displays the p-value of the KS test performed on the 3 days
456 sampled PIT of the 4 years of forecasts², for the climatology (Fig 11 a),
457 raw forecasts (Fig 11 b), and EMOS forecasts (Fig 11 c). The colorbar is
458 designed to test the null hypothesis at 95% confidence. Regarding this test,
459 forecasts are not calibrated in the northeast part of France which disagrees
460 with the test performed on the validation period. As this behaviour is quite
461 comparable to the climatology, this suggests that the fitting period of 20
462 years used to build the model and climatology may not be representative
463 enough of the wind in the forecasting years in those regions.

464 Figure 12 shows the ratio IC_{90clim}/IC_{90mod} averaged at 15 days (Fig 12
465 a.), monthly (Fig 12 b.) and seasonal horizon (Fig 12 c.) for raw forecasts
466 (black dots) and EMOS forecasts (green dots), for all forecast years.

467 The accuracy decreases with the forecast horizon. It appears that the
468 forecast is quite sharp within 15 days and deteriorates significantly for larger
469 horizons. The largest deterioration occurs between November and February.
470 In spring and summer, forecasting performance does not seem to be highly
471 sensitive to the forecast horizon.

472 Moreover, EMOS forecasts only significantly improves the accuracy with
473 respect to raw forecasts, in winter and fall, especially at monthly and seasonal
474 horizon. Raw and EMOS forecasts at 15 days horizon are almost always

²Note that we only have 48 independent forecasts which is not enough to test calibration. To get around this difficulty, we use forecasts obtained at the same date for different horizons (3 days, 6 days, 9 days, ..., 30 days) as if they were independent. Autocorrelation analysis shows that they are indeed uncorrelated. Each forecast thus corresponds to 10 data points.

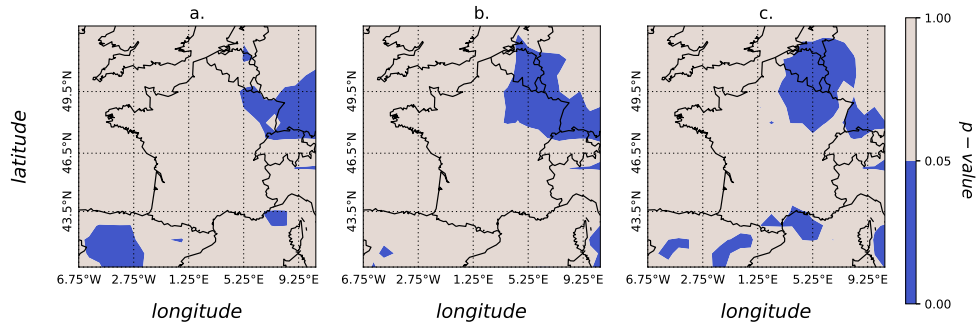


Figure 11: p-value of the KS test performed on the 3 days sampled PIT of the 4 years forecasts, for the climatology (a.), raw forecasts (b.), and EMOS forecasts (c.). Blue areas correspond to regions where the null hypothesis of calibration is rejected at the 5% confidence level.

475 better than the climatology. The improvement with EMOS optimization
 476 does not seem to be very large at this horizon, most probably because the
 477 distribution of the index is already very sharp. EMOS forecasts still slightly
 478 improve raw forecasts by about 5% in the beginning and at the end of the
 479 year.

480 The seasonal variability described in Table 1 is recovered at the 15 days
 481 and monthly horizon, which is encouraging. At the seasonal horizon, raw
 482 forecasts performance does not display a strong seasonal variability and the
 483 ratio is close to one, so the model does not perform better than the clima-
 484 tology. EMOS forecasts performance displays an even lower intra-annual
 485 variability but the ratio is systematically around 1.10. This is a very inter-
 486 esting result as it shows that there is a valuable statistical information
 487 on the local surface wind speeds in the seasonal forecasts of the large-scale
 488 circulation post-processed using the EMOS method, which leads to a 10%
 489 improvement over the climatology on average even at this long timescale.

490 Figures 13, 14 and 15 show the ratio IC_{90clim}/IC_{90mod} for each year of
 491 forecasts, respectively at 15 days, monthly and seasonal horizons, for raw
 492 forecasts (top) and EMOS forecasts (bottom). A sharp decrease of the ratio
 493 can be seen between figures 13 and 14. Comparatively, the accuracy decrease

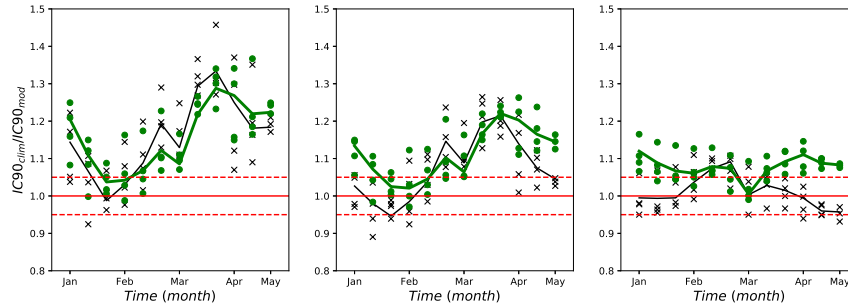


Figure 12: Ratio $IC_{90_{clim}}/IC_{90_{mod}}$ at 15 days (a.), monthly (b.) and seasonal horizon (c.) for every raw forecasts (black cross) - the black slight line represents the mean ratio ; and EMOS forecasts (green dots) - the green bold line represents the mean ratio. Four forecasted years are 2012, 2013, 2014, 2015.

494 is less pronounced, between figure 14 and 15. In the northeast of the domain,
 495 the ratio is the highest so that the model seems to be very sharp, especially
 496 at the 15 days horizon for the year 2015. This could be the cause of the
 497 decalibration of the model (Fig 11).

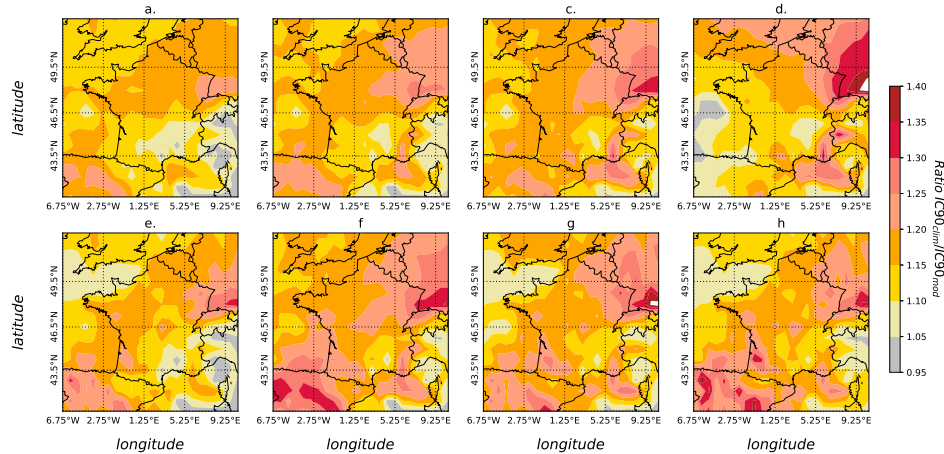


Figure 13: Ratio of $IC_{90_{clim}}$ over $IC_{90_{mod}}$ averaged over 12 seasonal forecasts, for forecasted years 2012 (a., e.), 2013 (b., f.), 2014 (c., g.), and 2015 (d., h.) at 15 days horizon for raw forecasts (top) and EMOS forecasts (bottom)

498 Those figures also show the efficiency of the EMOS method to reduce

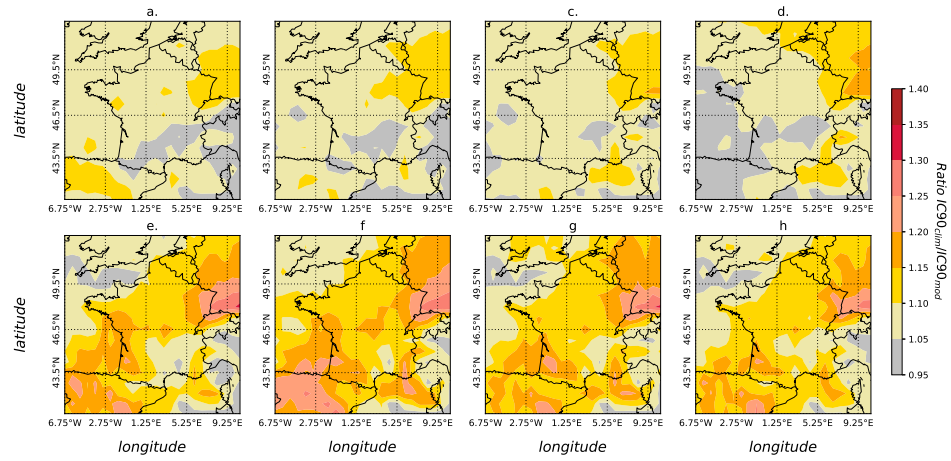


Figure 14: Same as figure 13 for monthly horizon forecasts.

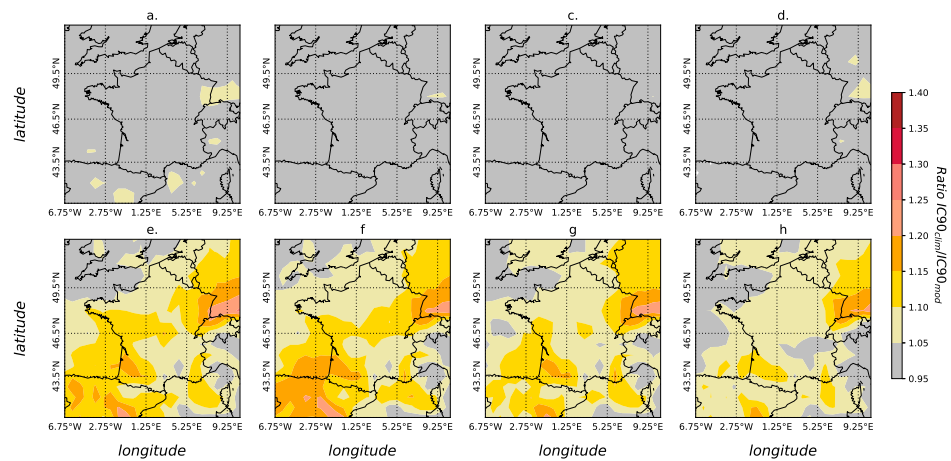


Figure 15: Same as figure 13 for seasonal horizon forecasts.

499 the uncertainty on the index and thus to highly sharpen the model so that
 500 forecasts at the seasonal horizon give more information on the wind than the
 501 climatology which is at this moment widely used for such long-term wind
 502 energy evaluation. The EMOS forecasts display a consistent spatial pattern
 503 in terms of accuracy for any forecast horizon which is not the case for raw

504 forecasts. Indeed, the performance of raw forecasts displays a noticeable
505 inter-annual variability. For instance, at the monthly horizon, forecasted
506 year 2015 is comparable to the climatology in the east of the domain while
507 year 2012, 2013, and 2014 are sharper than the climatology in this part of
508 France. It again highlights the added value of EMOS forecasts compared to
509 raw forecasting method, even if a larger sample of forecasted years should be
510 analysed to confirm this behaviour. The Mediterranean region is the region
511 where the model performs the worst compared to the climatology. This result
512 confirms what was found on the validation period. For EMOS forecasts, the
513 spatial pattern of the ratio is very comparable to the Fig 9 for any forecast
514 horizon and for all years. This is not as clear for raw forecasts. It means
515 that the uncertainty on the ensemble forecast is highly reduced by EMOS
516 method, but moreover that this method reduces the inter-annual variability
517 of the uncertainty of the ensemble.

518 **5. Conclusion**

519 A probabilistic model is proposed to predict daily wind speed distribu-
520 tion from a few days to seasonal timescale. It is compared to the climatology
521 which is often the reference used as the best seasonal forecast for energy man-
522 agement. The study shows that the model is better statistically calibrated
523 than the climatology and is able to follow very long-term trends of the wind
524 speed. On average over France, the model is shown to be 30% sharper than
525 the climatology. It is shown to be more accurate than the climatology espe-
526 cially onshore, in the northwest regions and in winter and fall.

527 We apply the probabilistic model to the seasonal forecast ensemble of
528 ECMWF. We test two methods to forecast wind speed with these ensembles.
529 The first method uses the empirical density of the raw calculated index,
530 and the second estimates the density of the calculated index by optimizing
531 calibration and sharpness of the ensemble using the EMOS statistical post-
532 processing technique (Gneiting et al. (2005)). We show that the model is
533 able to be more precise than the climatology at 15 days and monthly horizon
534 using both methods and that at the seasonal horizon, the EMOS method is
535 systematically more precise than climatology.

536 *Acknowledgement.*

537 This research was supported by the ANR project FOREWER (ANR-14-
538 CE05- 0028). This work also contributes to TREND-X program on energy

- 539 transition at Ecole Polytechnique as well as to the HyMeX program (HYdro-
540 logical cycle in The Mediterranean EXperiment ((Drobinski et al., 2014)))
541 through the working group Renewable Energy.
- 542 Albadi, M., El-Saadany, E., 2010. Overview of wind power intermittency
543 impacts on power systems. *Electric Power Systems Research* 80, 627–632.
- 544 Alonzo, B., Ringkjøb, H.K., Jourdier, B., Drobinski, P., Plougonven, R.,
545 Tankov, P., 2017. Modelling the variability of the wind energy resource on
546 monthly and seasonal timescales. *Renewable energy* 113, 1434–1446.
- 547 Azad, H.B., Mekhilef, S., Ganapathy, V.G., 2014. Long-term wind speed
548 forecasting and general pattern recognition using neural networks. *IEEE*
549 *Transaction on Sustainable Energy* 5, 546–553.
- 550 Barbounis, T.G., Theocharis, J.B., Alexiadis, M.C., , Dokopoulos, P.S., 2006.
551 Long-term Wind Speed and Power Forecasting Using Local Recurrent Neu-
552 ral Network Models. *IEEE Transaction on Energy Conversion* 21, 273–284.
- 553 Bilgili, M., Sahin, B., Yasar, A., 2007. Application of Artificial Neural Net-
554 works for the wind speed prediction of target station using reference sta-
555 tions data. *Renewable Energy* 32, 2350–2360.
- 556 Brossier, C.L., Drobinski, P., 2009. Numerical high-resolution air-sea cou-
557 pling over the Gulf of Lions during two tramontane/mistral events. *Journal*
558 *of Geophysical Research* 114, D10110.
- 559 Candille, G., Cote, C., Houtekamer, P., Pellerin, G., 2007. Verification of
560 an ensemble prediction system against observations. *Monthly Weather*
561 *Review* 135, 2688–2699.
- 562 Candille, G., Talagrand, O., 2005. Evaluation of probabilistic prediction
563 systems for a scalar variable. *Q. J. R. Meteorol. Soc.* 131, 2131–2150.
- 564 Carney, M., Cunningham, P., 2006. Evaluating density forecasting models.
565 Trinity College Dublin, Department of Computer Science .
- 566 Carpinone, A., Giorgio, M., Langella, R., Testa, A., 2015. Markov chain mod-
567 eling for very-short-term wind power forecasting. *Electric Power Systems*
568 *Research* 122, 152–158.

- 569 Casanueva, A., Rodríguez-Puebla, C., Frías, M.D., González-Reviriego, N.,
570 2014. Variability of extreme precipitation over Europe and its relationships
571 with teleconnection patterns. *Hydrol. Earth Syst. Sci.* 18, 709–725.
- 572 Cassou, C., 2008. Intraseasonal interaction between Madden-Julian Oscilla-
573 tion and the North Atlantic Oscillation. *Nature* 455, 523–597.
- 574 Chang, W., 2014. A literature review of wind forecasting methods. *Journal*
575 *of Power and Energy Engineering* 2, 161–168.
- 576 Davies, J.R., Rowell, D.P., Folland, C.K., 1997. North atlantic and Eu-
577 ropean seasonal predictability using an ensemble of multidecadal atmo-
578 spheric GCM simulations. *International Journal of Climatology* 17, 1263–
579 1284.
- 580 Delacroix, M., Hardle, W., Hristachea, M., 2003. Efficient estimation in
581 conditional single-index regression. *Journal of Multivariate Analysis* 86,
582 213–226.
- 583 Drobinski, P., Alonzo, B., Basdevant, C., Cocquerez, P., Doerenbecher, A.,
584 Fourri, N., Nure, M., 2017. Lagrangian dynamics of the mistral during
585 the Hymex SOP2. *Journal of Geophysical Research: Atmospheres* 122,
586 1387–1402.
- 587 Drobinski, P., Coulais, C., Jourdier, B., 2015. Surface wind-speed statis-
588 tics modelling: Alternatives to the Weibull distribution and performance
589 evaluation. *Boundary-Layer Meteorol* 157, 97–123.
- 590 Drobinski, P., Ducrocq, V., Alpert, P., Anagnostou, E., Branger, K., Borga,
591 M., Braud, I., Chanzy, A., Davolio, S., Delrieu, G., Estournel, C.,
592 Boubrahmi, N.F., Font, J., Grubisic, V., Gualdi, S., Homar, V., Ivancan-
593 Picek, B., Kottmeier, C., Kotroni, V., Lagouvardos, K., Lionello, P.,
594 Llasat, M., Ludwig, W., Lutoff, C., Mariotti, A., Richard, E., Romero,
595 R., Rotunno, R., Roussot, O., Ruin, I., Somot, S., Taupier-Letage, I.,
596 Tintore, J., Uijlenhoet, R., , H.Wernli, 2014. a 10-year multidisciplinary
597 program on the Mediterranean water cycle. *Meteorol. Soc.* 95, 1063–1082.
- 598 EWEA, 2016. Wind in Power: 2015 European Statistics. Technical Report.
599 European Wind Energy Association.

- 600 Folland, C.K., Knight, J., Linderholm, H.W., Fereday, D., Ineson, S., Hurrell,
601 J.W., 2008. The Summer North Atlantic Oscillation: Past, Present, and
602 Future. *Journal of Climate* 22, 1082–1103.
- 603 Foster, D.P., Vohra, R.V., 1998. Asymptotic calibration. *Biometrika*, , 85,
604 379–390.
- 605 Gneiting, T., Balabdaoui, F., Raftery, A.E., 2007. Probabilistic forecasts,
606 calibration and sharpness. *J. R. Statist. Soc. B* 69, 243–268.
- 607 Gneiting, T., Raftery, A.E., III, A.H.W., Goldman, T., 2005. Calibrated
608 probabilistic forecasting using Ensemble Model Output Statistics and min-
609 imum CRPS estimation. *Monthly Weather Review* 133, 1098–1118.
- 610 Gomes, P., Castro, R., 2012. Wind speed and wind speed forecasting using
611 statistical models: Autoregressive Moving Average (ARMA) and Artificial
612 Neural Networks (ANN). *International Journal of Sustainable Energy
613 Development (IJSED)* 1.
- 614 Guo, Z., Zhao, W., H.Lu, J.Wang, 2012. Multi step forecasting for wind speed
615 using a modified EMD based Artificial Neural Network model. *Renewable
616 Energy* 37, 241–249.
- 617 Hamill, T.M., 2000. Interpretation of rank histograms for verifying ensemble
618 forecasts. *Monthly Weather Review* 129, 550–560.
- 619 J. Wang, S.Q., Zhou, Q., Jiang, H., 2015. Medium-term wind speeds fore-
620 casting utilizing hybrid models for three different sites in Xinjiang China.
621 *Renewable Energy* 76, 91–101.
- 622 Jourdir, B., 2015. Wind resource in metropolitan France: assessment meth-
623 ods, variability and trends. Ph.D. thesis. Ecole Polytechnique.
- 624 Michelangeli, P.A., Vautard, R., Legras, B., 1995. Weather regimes: Recur-
625 rence and quasi stationarity. *Journal of Atmospheric Sciences* 52, 1237–
626 1256.
- 627 Möller, A., Lenkoski, A., Thorarinsdottir, T.L., 2013. Multivariate proba-
628 bilistic forecasting using Bayesian model averaging and copulas. *Q. J. R.
629 Meteorol. Soc.* 139, 982–991.

- 630 More, A., Deo, M., 2003. Forecasting wind with neural networks. *Marine*
631 *Structures* 16, 35–49.
- 632 Najac, J., Boe, J., Terray, L., 2009. A multi model ensemble approach for
633 assessment of climate change impact on surface winds in France. *Climate*
634 *Dynamics* 32, 615–634.
- 635 NationalGrid, 2016. Winter Outlook Report 2016/2017. Technical Report.
636 National Grid.
- 637 N.Schuhen, Thorarinsdottir, T.L., Gneiting, T., 2012. Ensemble Model Out-
638 put Statistics for wind vectors. *Monthly Weather Review* 140, 3204–3219.
- 639 Owen, J., Palmer, T., 1987. The impact of El Niño on an ensemble of
640 extended range forecasts. *Monthly Weather Review* 115, 2103–2117.
- 641 Pinson, P., Kariniotakis, G., 2009. Conditional prediction intervals of wind
642 power generation. *IEEE Transaction On Power Systems* 25, 1845–1856.
- 643 Powell, M.J.D., 1964. An efficient method for finding the minimum of a
644 function of several variables without calculating derivatives. *The Computer*
645 *Journal* 7, 155–162.
- 646 Pryor, S., Barthelmie, R., 2010. Climate change impacts on wind energy: A
647 review. *Renewable and Sustainable Energy Reviews* 14, 430–437.
- 648 Raftery, A.E., Gneiting, T., Balabdaoui, F., Polakowski, M., 2005. Us-
649 ing Bayesian model averaging to calibrate forecast ensembles. *Monthly*
650 *Weather Review* 133, 1155–1174.
- 651 Rodwell, M.J., Rowell, D.P., Folland, C.K., 1999. Oceanic forcing of the
652 wintertime North Atlantic Oscillation and European climate. *Nature* 398,
653 320–323.
- 654 Sailor, D.J., M. Smith, M.H., 2008. Climate change implications for wind
655 power resources in the northwest United States. *Renewable Energy* 33,
656 2393–2406.
- 657 Sloughter, J.M., Gneiting, T., Raftery, A.E., 2013. Probabilistic wind speed
658 forecasting using ensembles and Bayesian model averaging. *Journal of the*
659 *American Statistical Association* 141, 2107–2118.

- 660 Soman, S.S., Zareipour, H., O. Malik, P.M., 2010. A review of wind power
661 and wind speed forecasting methods with different time horizons. North
662 American Power Symposium (NAPS) , 1–8.
- 663 Stesfos, A., 2002. A novel approach for the forecasting of mean hourly wind
664 speed time series. *Renewable Energy* 27, 163–174.
- 665 Taylor, J., McScharry, P., Buizza, R., 2009. Wind power density forecasting
666 using ensemble prediction and time series model. *IEEE, Transactions on*
667 *Energy conversion* 34.
- 668 Thorarinsdottir, T.L., 2013. Calibration diagnostics for point process models
669 via the Probability Integral Transform. *Stat* 2, 150–158.
- 670 Thorarinsdottir, T.L., Gneiting, T., 2010. Probabilistic forecasts of wind
671 speed: Ensemble Model Output Statistics using heteroskedastic censored
672 regression. *J. R. Statist. Soc. A* 173, 371–388.
- 673 Wallace, J.M., Gutzler, D.S., 1980. Teleconnection in the geopotential height
674 field during the northern hemisphere winter. *Monthly Weather Review* 109,
675 784–812.
- 676 Wytock, M., Kolter, J.Z., 2013. Large-scale Probabilistic Forecasting in En-
677 ergy Systems using Sparse Gaussian Conditional Random Fields. *Proceed-*
678 *ings of the IEEE Conference on Decision and Control (CDC)* , 1019–1024.