# Space-time calibration of wind speed numerical forecasts from computer models

Luiz E. S. Gomes  Thaís C. O. Fonseca
Kelly C. M. Gonçalves  Ramiro R. Cárdenas

November 18, 2018

**Abstract**

Forecasts of meteorological variables from numerical models are systematically subject to errors. Such errors are mainly due to deterministic simulations of thermodynamic processes of the atmosphere from their current conditions through systems of differential equations. In addition, these systems are solved in a discrete grid, presenting uniform forecasts for every region belonging to the same grid cell. Consequently, forecasts from numerical models may not be representative at specific locations. In this context, statistical post-processing techniques are appropriate for calibration of these forecasts, minimizing possible distortions. This work aims to minimize the errors of Eta mesoscale model's forecasts for the wind speed through the development of improved spatiotemporal extensions of the main statistical post-processing models for meteorological variables. The proposed models were structured by the data augmentation technique and Bayesian Dynamic Linear Models. In particular, it considers trans-Gaussian random fields to account for asymmetric behaviour of meteorological variables such as wind speed. An optimized MCMC scheme based on robust adaptive Metropolis is employed for statistical inference and prediction. Three case studies illustrates the usefulness of our proposed correction in comparison with the usual postprocessing alternatives such as EMOS and GOP.

Wind speed Eta mesoscale model calibration dynamical spatiotemporal linear models data augmentation technique

## 1 Introduction

Numerical forecasts of meteorological variables are often based on mathematical models which make deterministic predictions from current atmospheric conditions. These are basically differential equations that have no analytic solution and use the numerical integration to simulate physical, dynamic and thermodynamic processes of the atmosphere depending on their current conditions. From this solution it is possible to obtain the solution of the system for any future time of interest. (Krishnamurti, 1995).

The Eta mesoscale model's numerical forecasts (Mesinger et al., 1988; Black, 1994) are useful in this context and, operationally, have been used by Centro de Previsão de Tempo Estudos Climáticos (CPTEC) of Instituto Nacional de Pesquisas Espaciais (INPE) since 1996 in order to provide short-term and long-term weather forecasts in Brazil. Its prognostic variables are air temperature, zonal and meridional components of the wind, specific humidity and surface pressure. Its current horizontal resolution is 5 km and the model operates twice a day (00 UTC and 12 UTC), providing hourly outputs for a forecast horizon of up to 72 hours (INPE/CPTEC, 2018).

These numerical systems are solved in a discrete grid, i.e., they present uniform predictions for every region belonging to the same cell of this grid. In particular each forecast is obtained based on average data of the region (e.g. average altitude and predominant vegetation). Notice that the representativeness of the predictions in places with complex orography and dense vegetation becomes deficient due to the differences in the real characteristics of the surface with the homogenization made by this model. Therefore, predictions generated by the Eta model may not be representative at a specific location (Chou et al., 2007), thus producing systematic errors.

In order to produce predictions at different points, minimizing these and other limitations of numerical models, statistical postprocessing techniques are appropriate and can improve the the estimates accuracy in a probabilistic context (Glahn and Lowry, 1972). In this context, spatial models may be useful in correcting predictions errors in regions where the numerical models have smoothed important characteristics of the terrain such as orography and land use.

Very often the estimation of parameters in postprocessing statistical methods consists of defining a training period based on a moving temporal window accounting for the effect of past observations and predictions from numerical models. If these training periods are reasonably long it becomes easier to estimate the uncertainty in predictions (Gneiting, 2014). However, these longer training periods may introduce distortions due to seasonal effects. Gneiting et al. (2005) and Raftery et al. (2005) analyze the effects of the training window size in the uncertainty estimation of parameters. The authors highlight that gains are obtained for windows as large as 25 days and that the window size must be tailored for the specific application to achieve better uncertainty measurements.

To account for seasonality the Kalman Filter (Kalman, 1960) and Bayesian dynamical models West and Harrison (1997) are natural alternatives to non-temporal calibration models allowing for seasonality and dynamics in the bias parameters.

This work proposes a unifying approach to statistical postprocessing by accounting for temporal dependencies overcoming the temporal window definition besides accommodating spatial features from the terrain which might be essential for wind prediction and uncertainty quantification. The proposal is based on spatial dynamical models and considers data augmentation to allow for an efficient estimation algorithm. An optimized MCMC scheme based on robust

adaptive Metropolis is employed for statistical inference and prediction. Three case studies illustrates the usefulness of our proposed correction in comparison with the usual postprocessing alternatives such as EMOS and GOP.

## 1.1 Motivation: Minas Gerais wind speed data

The state of Minas Gerais is located in the Southeast region of Brazil and recently interest has grown in the installation of wind turbines in order to expand the wind power use in the State. Throughout Minas Gerais and its surroundings, there are 59 meteorological monitoring stations, where hourly information is collected on the wind speed at 10 meters high, relative humidity, surface temperature, atmospheric pressure and precipitation. Figure 1 presents the distribution of these stations.

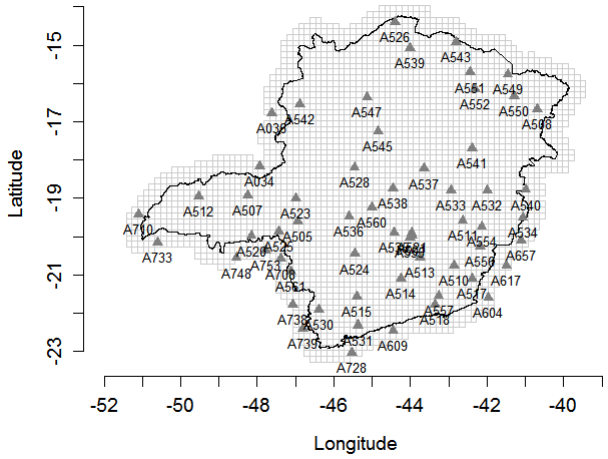
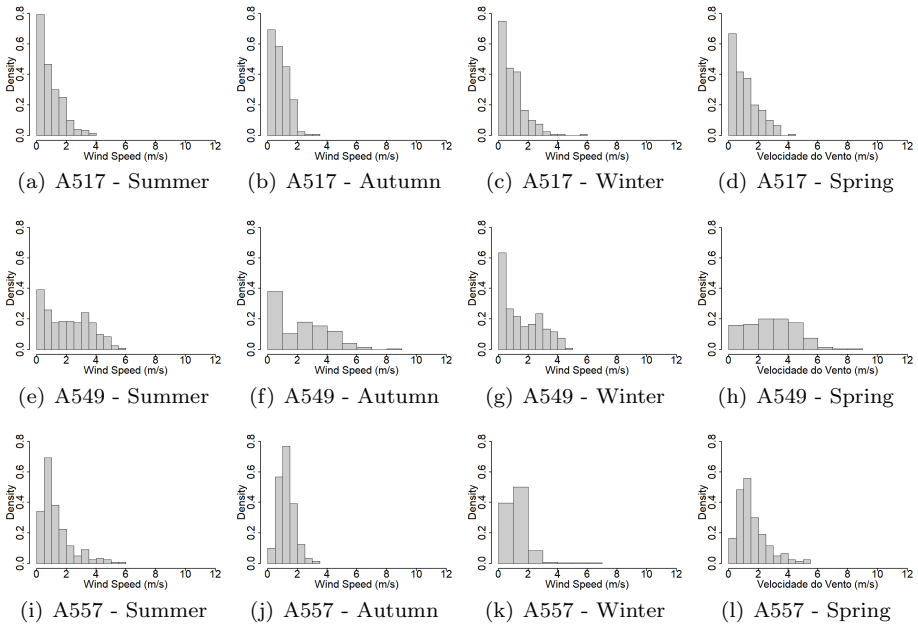



Figure 1: Meteorological stations locations in Minas Gerais. The solid triangles represent the stations and the lines represent the discrete grid used by the Eta mesoscale model.

The observed dataset in this stations could be used in order to decrease the systematic errors produced by the Eta mesoscale model's for the wind speed forecasts. According to Ailliot et al. (2006) some common features observed in wind dataset are: intermittent atmospheric regimes with predominance of a certain direction of wind in certain regions; spatial and temporal correlation; no Gaussianity; no stationarity; conditional heteroscedasticity - the variance of the wind speed changes frequently in time; seasonal annual and diurnal components due to the effects of the sun and seasons; possible trends. Thus, statistical postprocessing models should take into account these characteristics.

However, in addition to the behaviors previously mentioned, Eta mesoscale model's forecasts in Minas Gerais usually overestimate wind speed at 10 meters. This is due to the fact that models usually considered for wind modelling are

restricted to positive valued winds and do not allow for zero speed predictions. This is an important limitation of most of postprocessing models, as low and zero speeds are very often observed in the region of Minas Gerais as evidenced in Figure 2 which presents histograms of the wind speed at 10 meters for some weather stations. In general, the wind speed distribution at 10 meters in Minas Gerais is asymmetric, with great variability, with a mass point at 0 and reaching a maximum velocity of 12 $m/s$.




(a) A517 - Summer (b) A517 - Autumn (c) A517 - Winter (d) A517 - Spring

(e) A549 - Summer (f) A549 - Autumn (g) A549 - Winter (h) A549 - Spring

(i) A557 - Summer (j) A557 - Autumn (k) A557 - Winter (l) A557 - Spring

Figure 2: Histograms of wind speed collected at three different stations during each season.

In this way, the main aim is to propose statistical post-processing models which consider the spatial-temporal correlation and accommodate the fairly skew and censoring behaviour present in the dataset, in order to minimize the systematic error presented in the Eta mesoscale model's numerical predictions due to its intrinsic limitations. In particular, the proposed models are based on Box-Cox transformation and spatial dynamic linear models. The data augmentation approach for models with censored variables allow for an efficient estimation algorithm in this case.

# 2 Statistical postprocessing models

One numerical prediction does not describe a meteorological phenomena, thus, ensembles are often considered to allow for some sort of uncertainty quan-

tification. The ensembles might be interpreted as a Monte Carlo experiment aiming to produce a range of future states of the atmosphere from different initial conditions (Epstein, 1969). Besides, the simulation of scenarios might indicate extreme events which would not be identified by one run of the numerical model (Grimit and Mass, 2007).

The statistical postprocessing techniques are an attractive bias correction alternative to larger grid resolutions. The pioneer work in this context is the Model Output Statistics (MOS, Glahn and Lowry, 1972) which considers a multiple linear regression relating the responses (observed wind) with $m$ ensemble members $F_1, ..., F_m$ (numerical predictions) assuming constant variance.

An extension of the MOS known as Ensemble MOS (EMOS, Gneiting et al., 2005), allows for a relation between the dispersion of ensemble members and the response variance, the spread-skill relationship (Whitaker and Loughe, 1998).

In the context of spatial calibration the Geostatistical Output Perturbation (GOP, Gel et al., 2004) allows for spatial dependence resulting in meteorological fields for a fixed temporal horizon. The Spatial Ensemble Model Output Statistics (SEMOS, Feldmann et al., 2015) combines the EMOS method with GOP in its formulation.

As follows, an alternative spatial model is presented which accounts for temporal dynamics, a mass at zero and asymmetry through a spatiotemporal model with censoring.

## 2.1 Proposed models

Let $\left\{Y_t(s), s \in S \subset \mathbb{R}^2, t = 1, ..., T\right\}$ be a spatial random field in discrete time $t$. The observed response vector in $n$ locations $\mathbf{Y}_{t,s} = (y_t(s_1), ..., y_t(s_n))'$ is composed by censured variables, $y_t(s_i) \geq c$, $i = 1, \ldots, n$, $t = 1, \ldots, T$. Assume that $\mathbf{Y}_{t,s}$ follows a Truncated Gaussian distribution as follows.

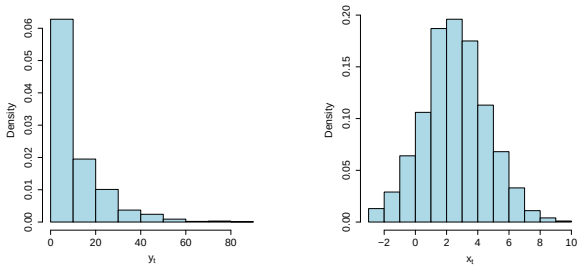
$$Y_t(s) = \begin{cases} BC^{-1}\left(X_t(s); \lambda\right), & \text{se} \quad BC^{-1}\left(X_t(s); \lambda\right) \geq c, \\ c^*, & \text{se} \quad BC^{-1}\left(X_t(s); \lambda\right) < c. \end{cases} \tag{1}$$

with $c$ and $c^*$ known constants, $\lambda$ is the unknown parameters of the Box-Cox transformation, $X_t(s)$ is a Gaussian process and $BC(.; \lambda)$ represents the Box-Cox transformation (Box and Cox, 1964) defined as:

$$BC(y; \lambda) = \begin{cases} \left(y^\lambda - 1\right)/\lambda, & \text{se} \quad \lambda \neq 0 \quad \text{e} \quad y > 0, \\ \log y, & \text{se} \quad \lambda = 0 \quad \text{e} \quad y > 0, \end{cases}$$

Thus, $X_t(s)$ is a latent Gaussian process which allows for asymmetry in the resulting process of interest $Y_t(s)$, that is, after transforming the possibly asymmetric process $Y_t(s)$ the resulting field follows a Gaussian process model. This behavior is illustrated in Figure 3 which presents a simulation of a transformed Gaussian field at a single location.

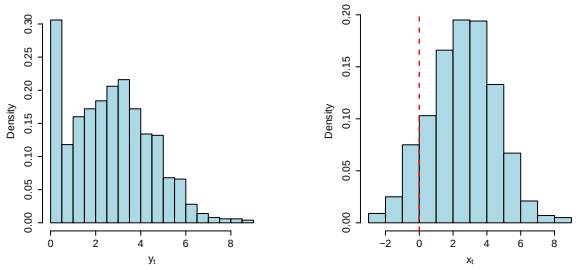
Furthermore, the constant $c$ allows for censoring, that is, winds observed from a minimum value which could be 0, but could be larger than 0 depending on the configuration of the turbine. Figure 4 illustrates the practical effect of




(a) Observed wind. (b) Transformed wind.

Figure 3: Histograms of observed wind speed and transformed wind speed obtained after taking the squared root transformation.

this modelling setup. Notice that the latent process corresponding to values of the Gaussian process below c naturally induces a mass at zero at the process of interest $X_t(s)$ after being truncated at $c$. The truncated Gaussian model has been useful in applied setting, such as precipitation modelling Bardossy and Plate (1992)). The complete model with Box-Cox transformation and censoring is facilitated by the use of data augmentation techniques (Tanner and Wong, 1987) and naturally allows for missing data.




(a) Observed wind. (b) Potential wind.

Figure 4: Histograms of observed wind speed and transformed wind speed obtained after censoring the latent process.

In the usual setup for statistical calibration of numerical models training sets, time windows are defined to account for observed and predicted temporal variation. Although larger training windows result in smaller uncertainty, it introduces distortions due to seasonal effects.

In general, the seasonal patterns of meteorological variables are well defined, e.g. radioactive forcing - 24 hours and seasons - 3 months.

In this context, the Kalman filter (Kalman, 1960) allows for both longer temporal training periods and inclusion of temporal dynamics in the bias parameters. As follows we consider the Bayesian dynamical approach which considers the Kalman filter for forward filtering and the backwards posterior smoothing for full Bayesian inference as proposed by West and Harrison (1997).

Therefore, we propose the use of dynamic linear spatial models which allow for regression coefficients to vary over time besides accounting for spatial dependence in the region of interest which might be oversmoothed by the numerical model.

### 2.1.1 Dynamical Geostatistical Output Perturbation

The proposed model adds a temporal dynamic to the GOP formulation by allowing the coefficients to vary over time. The covariance modeling is stochastic over time with Beta-Gamma evolution depending on discount factors. The dynamics are introduced by defining the observation and state equations as follows.

$$\mathbf{X}_{t,s} = \mathbf{F}'_{t,s}\boldsymbol{\theta}_t + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim N(\mathbf{0}_n, \varphi_t^{-1}\Sigma), \tag{2a}$$

$$\boldsymbol{\theta}_t = \mathbf{G}_t\boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t, \quad \boldsymbol{\omega}_t \sim T_{n_{t-1}}(\mathbf{0}_p, \mathbf{W}_t), \tag{2b}$$

$$\varphi_t = \gamma_t\varphi_{t-1}/\delta^*, \quad \gamma_t \sim \text{Beta}(\kappa_t, \bar{\kappa}_t), \tag{2c}$$

$$\kappa_t = \delta^* n_{t-1}/2 \text{ e } \bar{\kappa}_t = (1-\delta^*)n_{t-1}/2 \tag{2d}$$

where $\boldsymbol{\varepsilon}_t = (\varepsilon_t(s_1), ..., \varepsilon_t(s_n))'$ follows a zero mean multivariate normal distribution with correlation matrix $\Sigma_t$, with elements $\Sigma_{i,j} = C(s_i, s_j), \quad i, j = 1, \ldots, n$ and $\varphi_t = 1/\sigma_t^2$. $C(., \phi)$ is a valid correlation function depending on an unknown parameter $\phi$. In particular, we assume $C(s_i, s_j) = \exp(-\phi\|s_i - s_j\|)$, the exponential correlation function, with $\phi > 0$ representing the exponential decay parameter and $\|s_i - s_j\|$,the Euclidean distance between locations $s_i$ and $s_j$, $i, j = 1, ..., n$. $\mathbf{X}_{t,s} = (x_t(s_1), ..., x_t(s_n))'$, $\mathbf{F}'_{t,s}$ matrix with dimension $n \times r$ ($r \geq m$) composed by covariates (e.g. predicted ensembles, latitude, longitude and altitude), $\boldsymbol{\theta}_t$ represents the state variables with dimension $r$.

For the purely temporal components in (2b) and (2c), $\mathbf{G}_t$ is a evolution matrix with dimension $r$, $\boldsymbol{\omega}_t$ are mutually independent, zero mean, Student-t distributed, with $n_{t-1}$ degrees of freedom and unknown scale matrix $\mathbf{W}_t$, which might be estimated using discounting factors. The degrees of freedom parameters $n_{t-1}$ are defined through a Beta-Gamma stochastic evolution (West and Harrison, 1997). The parameter $\delta^* \in [0, 1]$ behaves as a discount factor , that is, the larger the discount smaller the random chock in the observational covariance. When $\delta^* = 1$, the covariance is static in time, that is, $\sigma_t^2 = \sigma^2$, $\forall t$. The initial information at time $t = 0$ assumes $\boldsymbol{\theta}_0|\mathbf{D}_0 \sim T_{n_0}(\mathbf{m}_0, \mathbf{C}_0)$ and $\varphi_0|\mathbf{D}_0 \sim G(n_0/2, d_0/2)$.

### 2.1.2 Spatiotemporal Ensemble Model Output Statistics

Analogously, the proposed model combines the spatial EMOS with DLMs. Assuming the structure stated in 1 for the original process $Y_t(s)$, the spatiotemporal Gaussian model for $X_t(s)$ is given by:

$$\mathbf{X}_{t,s} = \mathbf{F}'_{t,s}\boldsymbol{\theta}_t + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim N(\mathbf{0}_n, \Sigma^*_t), \tag{3a}$$

$$\boldsymbol{\theta}_t = \mathbf{G}_t\boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t, \quad \boldsymbol{\omega}_t \sim N(\mathbf{0}_p, \mathbf{W}_t), \tag{3b}$$

where $\boldsymbol{\varepsilon}_t = (\varepsilon_t(s_1), \ldots, \varepsilon_t(s_n))'$ follows a zero mean multivariate normal distribution with covariance matrix $\Sigma$, with elements $\Sigma^*_{t_{i,j}} = D_{t_{i,i}} C(s_i, s_j) D_{t_{j,j}}, \quad i, j = 1, \ldots, n$, with also $C(s_i, s_j) = \exp(-\phi\|s_i - s_j\|)$, the exponential correlation function and $D_t = diag(\sqrt{\beta_0 + \beta_1 S^2_{1,t}}, ..., \sqrt{\beta_0 + \beta_1 S^2_{n,t}})$ a n-dimensional diagonal matrix, such that $S^2_{i,t}$ is the sample variance of the *ensemble* for the location $i$ in time $t$. Thus, its main difference from EMOS is to insert the ensemble dispersion to the variance, since there is a positive relation between its width and the forecast absolute error. This is called spread-skill relationship. Different from the model (2), $\boldsymbol{\omega}_t$ is now normally distributed, with unknown covariance matrix $\mathbf{W}_t$, which might also be estimated using discounting factors.

## 2.2 Inference procedure

Let $\mathbf{y} = (\mathbf{y}_1, ..., \mathbf{y}_T)$ be the collection of $T$ observed time series at $n$ spatial locations in $\mathbb{R}^2$ and, $\Theta = (\boldsymbol{\theta}_{0:T}, \sigma^2_{0:T}, \phi, \lambda)'$ and $\boldsymbol{\Theta}^* = (\boldsymbol{\theta}_{0:T}, \boldsymbol{\beta}, \phi, \lambda)'$ be the parameter vector in (2) and (3), respectively, such that $\boldsymbol{\theta}_{0:T} = (\boldsymbol{\theta}_0, ..., \boldsymbol{\theta}_T)$ and $\sigma^2_{0:T} = (\sigma^2_0, ..., \sigma^2_T)$.

The inference procedure is performed under the Bayesian paradigm, and model specification is complete after assigning a prior distribution for the parameter vector $\Theta$ and $\Theta^*$. An advantage of following the Bayesian paradigm is that the inference procedure is performed under a single framework, and uncertainty about parameters estimation is naturally accounted for. Moreover, uncertainty about spatial interpolations, and temporal predictions is naturally described through the credible intervals of the respective posterior predictive distributions.

Assuming some components of $\Theta$ and $\Theta^*$ are independent *a priori*, we get by Bayes' theorem, the following posterior distribution for $\Theta$:

$$p(\Theta|\mathbf{y}) \propto \prod_{t=1}^{T} |\Sigma_t|^{-1/2} \exp\left\{-\frac{1}{2}\sum_{t=1}^{T}(\mathbf{X}_{t,s} - \mathbf{F}'_{t,s}\boldsymbol{\theta}_t)'{\Sigma_t}^{-1}(\mathbf{X}_{t,s} - \mathbf{F}'_{t,s}\boldsymbol{\theta}_t)\right\} \tag{4}$$

$$\times \exp\left\{-\frac{1}{2}\sum_{t=1}^{T}(\boldsymbol{\theta}_t - \mathbf{G}_t\boldsymbol{\theta}_{t-1})'{\mathbf{W}_t}^{-1}(\boldsymbol{\theta}_t - \mathbf{G}_t\boldsymbol{\theta}_{t-1})\right\} \tag{5}$$

$$\times \prod_{\{i,t:Y_{it}>c\}} Y_{it}^{\lambda-1} \times p(\boldsymbol{\theta}_0, \sigma^2_0, \phi, \lambda). \tag{6}$$

The posterior distribution for $\Theta^*$ is obtained analogously.

The kernel of this distribution does not result in a known distribution. Markov chain Monte Carlo (MCMC) methods are considered to obtain samples from the posterior distribution of interest. In particular, Gibbs sampler algorithm is used for $Z_t(s)$ and $U_t(s)$, the forward filtering backward sampling algorithm (FFBS, Frühwirth-Schnatter, 1994; Carter and Kohn, 1994) for $\boldsymbol{\theta}_{0:T}$ and $\sigma^2_{0:T}$ and the Robust adaptive Metropolis algorithm (RAM, Vihola, 2012) for the remaining parameters which is detailed in Appendix B.

# 3 Application

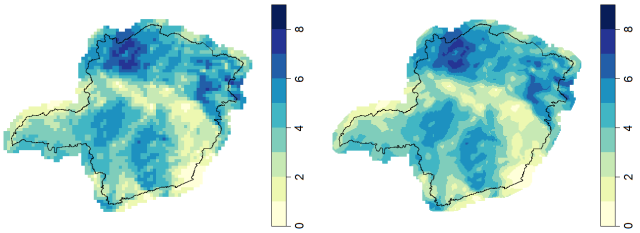
The state of Minas Gerais is located in the Southeast region of Brazil, being entirely formed by plateaus. The rugged relief gives to the state a privileged water resource, sheltering great hydroelectric potential. The predominant vegetation is Cerrado, consisting of large variations in the landscape between the rainy and dry seasons, resulting in a seasonal influence of the aerodynamic roughness of the terrain in the displacement of the winds. The climate in Minas Gerais varies from hot semiarid to humid mesotherm. In general, the distribution of rainfall is unequal with the north presenting long periods of drought. In the higher altitude areas of the south, the pluviometric regime is more intense. Seasonality also influences on temperatures, with high averages predominantly in the summer. Periodically, in most of the territory of Minas Gerais, there are more intense winds in winter and spring (Amarante et al., 2010).

In this case study, interest lies in analyzing the error in the Eta numerical predictions, specially in places that are affected by aspects of geographic characteristics such as latitude, longitude and altitude, proximity to water bodies and regional vegetation. This heterogeneous behaviour is illustrated by figures 11 and 12 of Appendix A which show, respectively, the time series of the wind speed at 10 meters and its numerical predictions and the autocorrelation function (ACF) for the time series at the same stations, throughout the four seasons. Notice that the wind at different sites across this region do not follow a similar pattern. Stations such as A507 and A547 show higher errors during the spring and lowest errors in the fall. In addition, there are well-defined periodic patterns during some seasons. Stations A537, A543 and A547 do not record this pattern during the spring, unlike A507 and A530. These periodic patterns are seasonal effects due to solar forcing which has a direct influence on wind speed.

The dataset considered for model calibration consists of the ensemble of hourly numerical forecasts initiated at 12 UTC of the Eta mesoscale model for the instantaneous wind speed at 10 meters in the State of Minas Gerais and its surroundings from November 1, 2015, 12 UTC to November 30, 2016, 11 UTC, totalizing over 10 thousand ensembles, one for each hour of forecast. The maximum forecast horizon for this model is 255 hours. Moreover, the dataset contains measurements of the meteorological variable under study in the 59 monitoring stations and its surroundings, as well as georeferenced data (e.g. latitude, longitude and relief height). The recorded numerical predictions and weather variables are available for the same time period. The numerical

accuracy of the data is one decimal place, with a lower non-zero value of 0.1. This information will be important for the definition of the censorship constants $c$ and $c^*$ in (1).

The meteorological stations are irregularly spaced and the grid is considerably fine in this region, as presented in Figure 1. However, the proposed calibration models require that numerical predictions and observed wind are available at the same spatial locations. In this context, numerical forecasts for the observation sites are obtained by bilinear interpolation. Figure 5 illustrates the interpolated wind speed at 10 meters in Minas Gerais.




(a) Numerical forecast (b) Bilinear interpolation

Figure 5: Bilinear interpolation in the discrete grid with cells 15km $\times$ 15km used by mesoscale Eta model to get numerical forecasts in the observed sites.

For each fitted model, the same mean structure is defined depending on the averaged ensemble members for each location; an auto-regressive component with values observed in the recent past, due to the lack of local information, such as roughness and relief; and the static auxiliary variables latitude, longitude and altitude. To avoid the eventual unavailability of a member of the ensemble and multicollinearity problems, the ensembles which are currently available under a given calibration horizon are averaged. This procedure does not entail a great loss of information, since there is a high linear correlation between the members, as shown in Figure 6, implying a good stability of the Eta model. According to Grimit and Mass (2007), the ensemble average may capture a possible point anomaly (e.g. cold front) and thus its use is also indicated by theoretical aspects.

In this application, we fitted the following three models described in Table 1. All them considered the left censoring with Box-cox transformation described in model 1.

In particular, we fixed $\delta^* = 1$ in DGOP model described in (2), this is, $\sigma_t^2 = \sigma^2, \forall t$. Thus, $\sigma^2$ in (2) would coincide with $\beta_0$ in STEMOS model described in (3) and the main difference between these models is due to the spread-skill relationship accounted by the effect of $\beta_1$ in (3). Therefore, as $\beta_1$ goes to zero, STEMOS model is closer to DGOP with fixed variance.

For models (2) and (3) we set

$$\mathbf{F}'_{t,s} = \left(1, \bar{f}(s_i), Y_{t-h}(s_i), \text{height}(s_i), \text{latitude}(s_i), \text{longitude}(s_i), 1, 0\right),$$

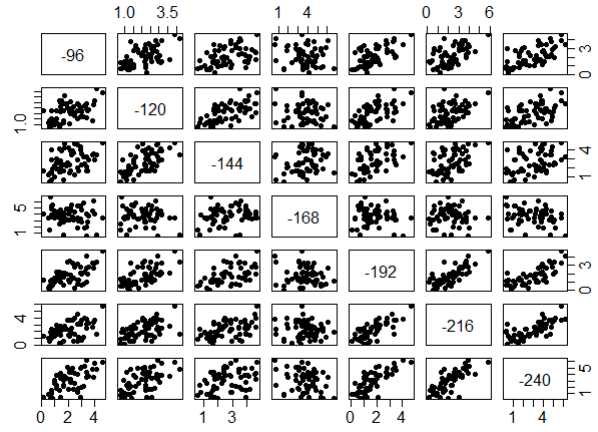



Figure 6: Pairwise scatterplot for different ensemble members of weather forecasts at 10 meters in April 1, 2016, 12 UTC.

| Model | Features |
|---|---|
| Dynamical GOP (DGOP) | spatio-temporal component<br>DLM<br>Beta-Gamma evolution for variance |
| Spatiotemporal EMOS (STEMOS) | spatio-temporal component<br>spread-skill relationship in variance |
| Spatial EMOS (SEMOS) | spatial component<br>Normal DLM<br>spread-skill relationship in variance |

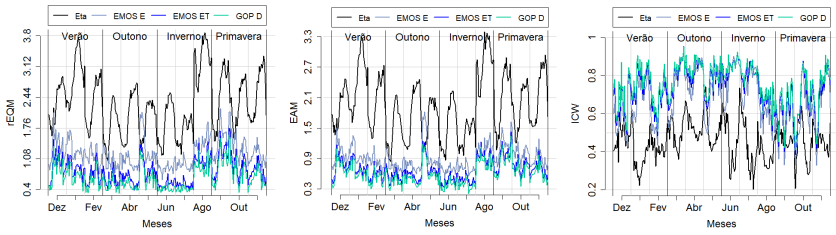
Table 1: Brief summary of models fitted in study case I with its main features.

$\mathbf{G}_t = \begin{pmatrix} \mathbf{I}_6 & \cos(2\pi/24) & \text{sen}(2\pi/24) \\ 0 & -\text{sen}(2\pi/24) & \cos(2\pi/24) \end{pmatrix}$, where $\mathbf{I}_k$ is the identity matrix of dimension k× k. The covariance matrix of the evolution equation, $\mathbf{W}_t$, is estimated using the ideas of discounting factors, and we set the discount factors equal to 0.99 to the intercept,0.99 to the components referring to geographic location and 0.95 to the seasonality component, respectively.

We assign reasonably vague priors to model parameters, more specifically, we assume $\lambda \sim N(1, 10)$. For the parameters in the exponential correlation function $\phi$, we assume $G\left(2, \frac{\max(d)}{6}\right)$, which prior mean is such that the practical range (when the correlation is equal to 0.05) is reached at half of the maximum distance (dmax/2) between geographical locations, and the variance is infinite. Moreover, we assume $\boldsymbol{\theta}_0 \sim T_1(\mathbf{0}, \mathbf{I}_8)$ and $\varphi \sim G(1; 0, 1)$ in GOPD, $\boldsymbol{\theta}_0 \sim N(\mathbf{0}_8, \mathbf{I}_8)$ and $\boldsymbol{\beta} \sim NT_{(0,\infty)}(\mathbf{0}_2, 10\mathbf{I}_2)$ in STEMOS, and, $\boldsymbol{\theta} \sim N(\mathbf{0}_8, \mathbf{I}_8)$ and $\boldsymbol{\beta} \sim NT_{(0,\infty)}(\mathbf{0}_2, 10\mathbf{I}_2)$

Table 2: Computational time (in minutes) for each calibration model fitted.

| Model | GOPD | STEMOS | SEMOS |
|---|---|---|---|
| | 34,78 | 42,74 | 34,43 |




(a) srMSE (b) MAE (c) WIA

Figure 7: *Model comparison criteria computed under DGOP, STEMOS, SEMOS and Eta model for each season.*

in SEMOS model.

For each model, we ran two parallel chains starting from very different values and we used the diagnostic tools to check convergence of the chains. The MCMC algorithm was implemented in the R programming language, version 3.4.1 (R Core Team, 2017), in a computer with an **Intel(R) Core(TM) i5-4590 processor 3.30 GHz, 8GB RAM memory**. Some particular functions were implemented in `C++` language using the library `Armadillo` (Sanderson and Curtin, 2016) through the package `Rcpp` (Eddelbuettel et al., 2011). Table 2 displays the computational time (in minutes) for each calibration model fitted.

Model comparison is performed using the square root of the mean square error (srMSE), mean absolute error (MAE), Willmott's index of agreement(WIA, Willmott, 1981) and interval score (IS, Gneiting and Raftery, 2007). The WIA is a standardized measure ranging from 0 (absence of agreement) to 1 (perfect agreement). The IS takes into account the amplitude and coverage of the prediction intervals in a parsimonious way. In particular, the only first three criteria seems to be appropriate to compare numerical predictions from the Eta model, which provides only point estimates, with the proposed postprocessing statistical models. Each of these criteria are described in more detail in Appendix C.

Figure 7 shows the different model comparison criteria computed for each proposed model and the Eta mesoscale model. All three criteria point that the proposed models performed significantly better than the Eta model. Furthermore, between the three calibration models fitted, all three criteria point to DGOP and STEMOS as the best.

In order to compare the interval forecasts, Figure 8 presents the IS throughout the seasons. Note that in this criterion, potential disparities have been highlighted. The proposed SEMOS has a long peak sequence, while the pro-

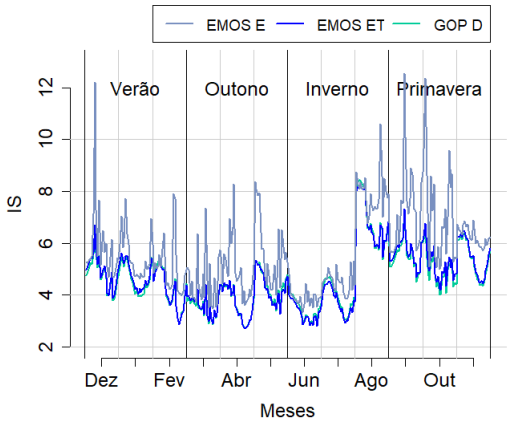
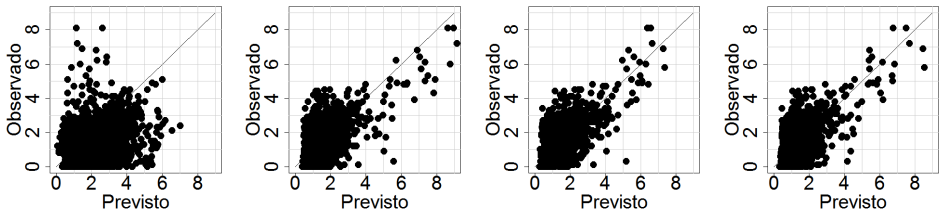



Figure 8: *Interval Score criteria computed under DGOP, STEMOS, SEMOS and Eta model for each season.*

posed STEMOS and GOPD are more stable.

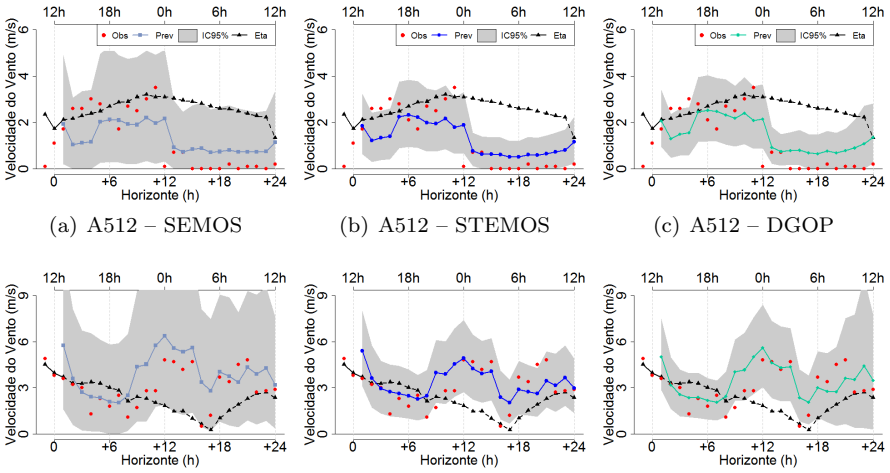
In order to determine the local behavior of the numerical and calibrated forecasts based on the proposed models' fit, we select the winter months of July and August, because they show a good performance. Although they belong to the same season of the year, the srMSE and MAE in August seems more like to the pattern of the following season, appearing to be a period of climatic regime changing. From July 20, 13 UTC to July 21, 2016, 12 UTC, wind speeds were recorded up to 8 m/s, as shown in Figure 9. Although in this application, higher speeds ($> 5$ m/s) appear more frequently, most of the recorded speeds still accommodate up to 5 m/s. The scatterplots in Figure 9 show that the proposed models returns a very close relationship of linear between observed and predicted values, even for speeds above 5 m/s. However, SEMOS seems to overestimate the wind velocity at 10 m, while Eta model predictions demonstrates low representativeness.




(a) Eta (b) EMOS E (c) EMOS ET (d) GOP D

Figure 9: Scatterplot comparing the predicted with the observed values from July 20, 13 UTC to July 21, 2016, 12 UTC for each model.

Panels of Figure 10 show the posterior summary of the predictive distribution for 24 hours ahead for 2 different monitoring stations for SEMOS, STEMOS and GOPD. The first station corresponds to A512 located in Ituiutaba and the second one to A555 in Ibirité. There is a large difference between the numerical and calibrated forecasts, indicating a significant relevance of other covariates. Although most of all the actual observed values fall within the 95% posterior predictive intervals, the range of their 95% credible regions differ considerably. Clearly, the SEMOS provides predictive intervals with the highest ranges. This is probably due to the use of a long training period, which can introduce distortions in the parameter estimation due to seasonality provided by the effect of solar forcing on wind speed at 10 meters. On the other hand, the proposed STEMOS and GOPD benefits from this behavior, presenting the intervals with smallest ranges.




(a) A512 – SEMOS (b) A512 – STEMOS (c) A512 – DGOP

(d) A555 – EMOS E (e) A555 – EMOS ET (f) A555 – GOP D

Figure 10: Temporal prediction of the wind speed for 24 hours ahead based on SEMOS and STEMOS from July 20, 13 UTC to July 21, 2016, 12 UTC.

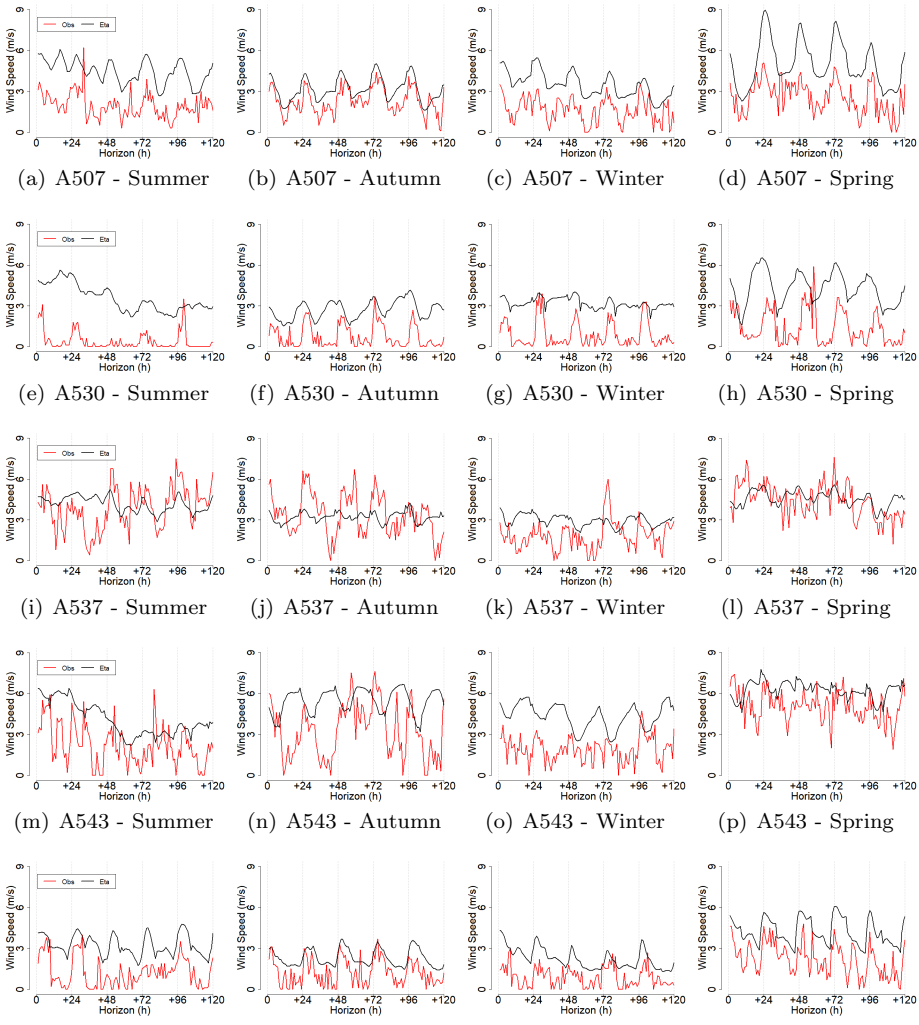
# 4 Concluding remarks

This paper presents a novel statistical calibration model for numerical forecasts of meteorological variables. In particular, we investigate the calibration of wind forecasts which are well known to be locally predicted with error by numerical models. The proposed model allows for spatiotemporal calibration, that is, the coefficients are allowed to dynamically change over time and local characteristics are partially captured by spatial correlation between locations. This overcome the definition of a training period for post-processing. The proposal

introduces data transformation withing the dynamical model which results in a flexible sampling distribution for the errors which could be potentially asymmetric as in the wind example. In the wind application, our approach clearly leads to narrower predictions when compared to simpler models without temporal dependence. As future directions this model might be applied to precipitation and other asymmetric censored variables.

# A Further information about the wind speed data

Figure 11 shows the time series of the wind speed at 10 meters and its respectives numerical forecasts for each season. Figure 11 presents the autocorrelation function in the same cases.




(a) A507 - Summer (b) A507 - Autumn (c) A507 - Winter (d) A507 - Spring

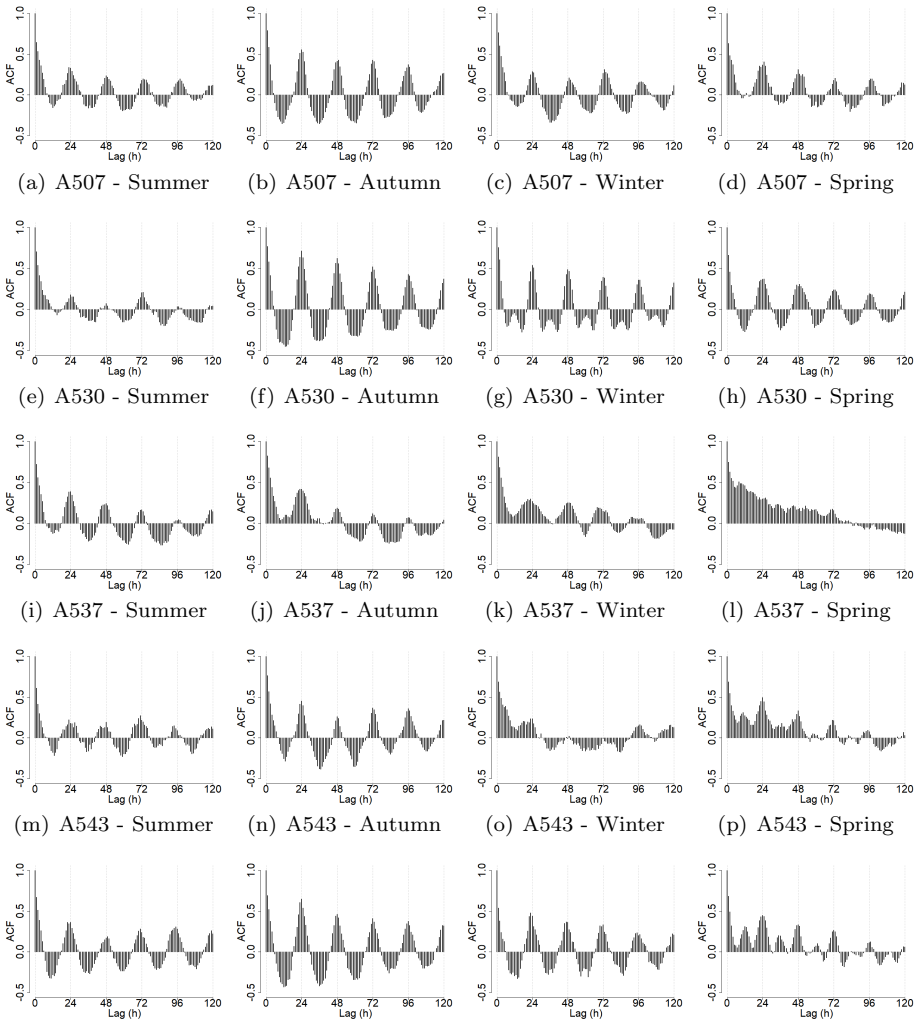
(e) A530 - Summer (f) A530 - Autumn (g) A530 - Winter (h) A530 - Spring

(i) A537 - Summer (j) A537 - Autumn (k) A537 - Winter (l) A537 - Spring

(m) A543 - Summer (n) A543 - Autumn (o) A543 - Winter (p) A543 - Spring

(q) A547 - Summer (r) A547 - Autumn (s) A547 - Winter (t) A547 - Spring

Figure 11: Time series of the wind speed at 10 meters and its respectives numerical forecasts for each season starting at 12 UTC.




(a) A507 - Summer (b) A507 - Autumn (c) A507 - Winter (d) A507 - Spring

(e) A530 - Summer (f) A530 - Autumn (g) A530 - Winter (h) A530 - Spring

(i) A537 - Summer (j) A537 - Autumn (k) A537 - Winter (l) A537 - Spring

(m) A543 - Summer (n) A543 - Autumn (o) A543 - Winter (p) A543 - Spring

(q) A547 - Summer (r) A547 - Autumn (s) A547 - Winter (t) A547 - Spring

Figure 12: Autocorrelation functions for each monitoring location and for each season.

# B Robust adaptive Metropolis

The method was proposed by Vihola (2012) and is used when the full conditional distribution does not have an analytical closed form. It is a more efficient extension of the well-known Metropolis-Hastings algorithm (Metropolis et al., 1953).

In addition to generating a sample of the posterior distribution of the parameters of interest $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_p)$, it also handles the with the acceptance rate $bar\alpha$, by fixing a target rate $\alpha^* \; in(0,1)$ and defining the proposed distribution $\tilde{q}(.)$, independently, only requiring that it must be spherically symmetric and centered in $\mathbf{0}_p$. Define $\left\{\zeta^{(k)}\right\}_{k\geq 1}$ as a sequence which decays to zero. The sampling scheme is described as follows:

1. Initialize $\boldsymbol{\theta}^{(0)} = (\theta_1^{(0)}, ..., \theta_p^{(0)})'$, $S^{(0)} = s_0 I_p$ and $k = 1$;
2. Get a new value to $\boldsymbol{\theta}^{(k)}$ from $\boldsymbol{\theta}^{(k-1)}$ this way:
   (a) Generate a new value for para $\boldsymbol{\theta}^{(k)}$ using that:
   $$\boldsymbol{\theta}^{(k)} = \boldsymbol{\theta}^{(k-1)} + S^{(k-1)} V^{(k)}, \quad V^{(k)} \sim \tilde{q}(.);$$
   (b) The proposed value in (a) is accepted with acceptance probability $\alpha$:
   $$\alpha^{(k)} = min\left\{1, \frac{p(\boldsymbol{\theta}^{(k)}|\boldsymbol{\theta}^{(k-1)}, y) q(\boldsymbol{\theta}^{(k-1)}|\boldsymbol{\theta}^{(k)})}{p(\boldsymbol{\theta}^{(k-1)}|\boldsymbol{\theta}^{(k)}, y) q(\boldsymbol{\theta}^{(k)}|\boldsymbol{\theta}^{(k-1)})}\right\};$$
   (c) Calculate a lower triangular matrix $S^{(k)}$ with positive diagonal elements such that:
   $$S^{(k)} S'^{(k)} = S^{(k-1)} \left(I_p + \zeta^{(k)} (\alpha^{(k)} - \alpha^*) \frac{V^{(k)} V'^{(k)}}{\left\|V^{(k)}\right\|^2}\right) S'^{(k-1)}$$
3. Do $k = k + 1$ and return to 2, until the convergence is achieved.

An optimal acceptance rate to any random walk is $\alpha^* \approx 23,4\%$. It is indicated that $s_0 = \frac{2,4}{\sqrt{p}}$ in order to achieve the convergence in a faster way with the algorithm.

# C Model comparison criteria

In this section we briefly describe the model comparison criteria we use to compare the fitted models in Section 3.

## C.1 Mean absolute error and square root of the mean square error

Standard measures of goodness of fit were also entertained in this study for comparison purposes. The square root of the mean square error (srMSE) and the mean absolute errors (MAE) are given by:

$$\text{srMSE} = \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}(y_t(s_i)-\hat{y}_t(s_i))^2 \quad \text{and} \quad \text{MAE} = \frac{1}{nT}\sum_{i=1}^{L}\sum_{t=1}^{T}|y_t(s_i)-\hat{y}_t(s_i)|,$$

where $\hat{y}_t(s_i)$ is obtained through a Monte Carlo estimate of the posterior mean of the predictive distribution, that is, $E\left[y_t(\mathbf{s}_i) \mid \mathbf{y}\right]$, across $N$ draws. Smaller values of srMSE and MAE indicate the best model among the fitted ones.

## C.2 Willmott's index of agreement

Willmott (1981) introduced a standard measure for assessing the quality of forecasts. The Willmott's index of agreement (WIA) ranges between 0 (absence of agreement) and 1 (perfect agreement) and is given by:

$$\text{WIA} = 1 - \frac{\sum_{i=1}^{n}\sum_{t=1}^{T}\left(y_t(s_i)-\hat{y}_t(s_i)\right)^2}{\sum_{i=1}^{n}\sum_{t=1}^{T}\left(|\hat{y}_t(s_i)-\bar{y}|+|y_t(s_i)-\bar{y}|\right)^2},$$

where $\bar{y} = \frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T} y_t(s_i)$.

## C.3 Interval score

The interval score (IS, Gneiting and Raftery, 2007) is a scoring rule for interval predictions considering the symmetric prediction interval with level $(1-\alpha)\times 100\%$. The score is rewarded by accurate intervals and penalized when there is no coverage of the forecast. If all true values are contained in the prediction interval, this measure is reduced to the range amplitude. The average IS is given by: where $\hat{l}_t(s_i)$ e $\hat{u}_t(s_i)$ are, respectively, the lower limit obtained by the $\frac{\alpha}{2}$ quantile, and the upper bound, obtained by the $1-\frac{\alpha}{2}$ quantile of the based on the predictive distribution.

Smaller IS values indicate probabilistic forecasts more efficient.

# References


Ailliot P, Monbet V, Prevosto M (2006) An autoregressive model with time-varying coefficients for wind fields. Environmetrics 17(2):107–117

Amarante Od, Silva F, Andrade P (2010) Atlas eólico: Minas gerais. Belo Horizonte: CEMIG

Bardossy A, Plate EJ (1992) Space-time model for daily rainfall using atmospheric circulation patterns. Water Resources Research 28(5):1247–1259

Black TL (1994) The new nmc mesoscale eta model: Description and forecast examples. Weather and forecasting 9(2):265–278

Box GE, Cox DR (1964) An analysis of transformations. Journal of the Royal Statistical Society Series B (Methodological) pp 211–252

Carter CK, Kohn R (1994) On gibbs sampling for state space models. Biometrika 81(3):541–553

Chou S, Souza Cd, Gomes JL, Evangelista E, Osório C, Cataldi M (2007) Refinamento estatístico das previsões horárias de temperatura a 2 m do modelo eta em estações do nordeste do brasil. Revista Brasileira de Meteorologia 22(3):287–296

Eddelbuettel D, François R, Allaire J, Ushey K, Kou Q, Russel N, Chambers J, Bates D (2011) Rcpp: Seamless r and c++ integration. Journal of Statistical Software 40(8):1–18

Epstein ES (1969) Stochastic dynamic prediction. Tellus 21(6):739–759

Feldmann K, Scheuerer M, Thorarinsdottir TL (2015) Spatial postprocessing of ensemble forecasts for temperature using nonhomogeneous gaussian regression. Monthly Weather Review 143(3):955–971

Frühwirth-Schnatter S (1994) Data augmentation and dynamic linear models. Journal of time series analysis 15(2):183–202

Gel Y, Raftery AE, Gneiting T (2004) Calibrated probabilistic mesoscale weather field forecasting: The geostatistical output perturbation method. Journal of the American Statistical Association 99(467):575–583

Glahn HR, Lowry DA (1972) The use of model output statistics (mos) in objective weather forecasting. Journal of applied meteorology 11(8):1203–1211

Gneiting T (2014) Calibration of medium-range weather forecasts. European Centre for Medium-Range Weather Forecasts

Gneiting T, Raftery AE (2007) Strictly proper scoring rules, prediction, and estimation. Journal of the American Statistical Association 102(477):359–378

Gneiting T, Raftery AE, Westveld III AH, Goldman T (2005) Calibrated probabilistic forecasting using ensemble model output statistics and minimum crps estimation. Monthly Weather Review 133(5):1098–1118

Grimit EP, Mass CF (2007) Measuring the ensemble spread–error relationship with a probabilistic approach: Stochastic ensemble results. Monthly weather review 135(1):203–221

INPE/CPTEC (2018) Instituto nacional de pesquisas espaciais — centro de previsão de tempo e estudos climáticos — eta model — regional. http://etamodel.cptec.inpe.br/, acessado em 15/05/2018

Kalman RE (1960) A new approach to linear filtering and prediction problems. Journal of basic Engineering 82(1):35–45

Krishnamurti T (1995) Numerical weather prediction. Annual review of fluid mechanics 27(1):195–225

Mesinger F, Janjić ZI, Ničković S, Gavrilov D, Deaven DG (1988) The step-mountain coordinate: model description and performance for cases of alpine lee cyclogenesis and for a case of an appalachian redevelopment. Monthly Weather Review 116(7):1493–1518

Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E (1953) Equation of state calculations by fast computing machines. The journal of chemical physics 21(6):1087–1092

R Core Team (2017) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, URL http://www.R-project.org/

Raftery AE, Gneiting T, Balabdaoui F, Polakowski M (2005) Using bayesian model averaging to calibrate forecast ensembles. Monthly Weather Review 133(5):1155–1174

Sanderson C, Curtin R (2016) Armadillo: a template-based c++ library for linear algebra. Journal of Open Source Software

Tanner MA, Wong WH (1987) The calculation of posterior distributions by data augmentation. Journal of the American statistical Association 82(398):528–540

Vihola M (2012) Robust adaptive metropolis algorithm with coerced acceptance rate. Statistics and Computing 22(5):997–1008

West M, Harrison P (1997) Bayesian forecasting and dynamic models. 2nd ed. Springer Verlag, New York

Whitaker JS, Loughe AF (1998) The relationship between ensemble spread and ensemble mean skill. Monthly weather review 126(12):3292–3302

Willmott CJ (1981) On the validation of models. Physical geography 2(2):184–194