

Statistical Correction of the Result of Weather Forecast by Applying the Model Output Calibration

Carlos Amaral Hölbig¹, Vanessa Lago Machado²,
Willingthon Pavan³

*Graduate Program in Applied Computing (PPGCA)
University of Passo Fundo (UPF)
Passo Fundo, Brazil*

Abstract

This work describes the simulation models calibration method called Model Output Calibration. In order to verify its effectiveness, presents the application of the MOC in weather forecast correction generated by the Eta 15Km model at CPTEC/INPE. Eta is a regional model for numerical weather prediction. The results of the statistical correction of Eta forecast were positive, with satisfactory improvements in the variables tested (temperature and relative humidity). The use of this approach shows the possibility of gains in the results of simulation models of crops and diseases that use as predictive variables the variables generated by weather forecast models.

Keywords: Statistic correction, Weather forecast, MOC, Eta Model.

1 Introduction

Simulation models are important tools that act as forecast tools, enabling decisions to be made before the events occur. But even with all the developments in techniques for simulation modeling, the accuracy found in the data, in some cases, it is not yet satisfactory, and ultimately influences wrongly data verification.

Numerical weather prediction (NWP) models are composed of weather forecast techniques performed through computational modeling, allowing to analyze and decide on various situations related to weather and climate. These models are divided into global and regional models (limited area). Furthermore, numerical

¹ Email: holbig@upf.br

² Email: 154401@upf.br

³ Email: pavan@upf.br

weather prediction models divide their modeling space in boxes, which are defined by horizontal grids boxes with vertical levels. Thus, assume that the atmosphere is homogeneous in each of these boxes, it is necessary to know only the data of one point per box.

For purposes of this research, the numerical weather prediction model used in our work was the regional model Eta [6], used by the Center for Weather Forecasting and Climate Research (CPTEC⁴) of the National Institute for Space Research (INPE) in Brazil.

The forecast of this regional model extends up to 264 hours (11 days) are supplied twice a day (at 00:00 and 12:00, UTC time). Eta has grids with 40 km, 15 km, 5 km and 1 km. This model was developed by Mesinger et al. [15] and made operational at the National Centers for Environmental Prediction (NCEP⁵) [1]. It is a model based on the use of vertical coordinates, which remains approximately horizontal in mountainous areas, making it suitable for studies of steep topography regions, for example, the Andes Mountains [5].

The Eta model is used operationally at CPTEC/INPE since 1997 for weather forecasts [2] and, since 2002, to seasonal climate forecasts [3] in high resolution over South America for horizons of a few days, months or decades, for climate change studies.

The main sources of observed data of Eta model are: meteorological stations, surface stations, airports, ships, aircraft, satellites and radars. These sources are part of the Global Observing System (GOS) [11], as shown in Figure 1. When the observed data are received, some situations may occur, for example, some of the locations may not have measurement data observed at certain hours: caused by a lack of measurement of typing or even failure at the station; another situation is the analysis of unrealistic measurements: sometimes occurs the erroneous data entry in any station. To correct these errors is necessary the treatment of data by, for example, numerical interpolation of missing or unrealistic data available in the databases.

The prognostic variables Eta model are: air temperature, zonal and meridional wind components, specific humidity, surface pressure and turbulent kinetic energy. Thus, the model is proposed to predict atmospheric phenomena in more detail, when associated with these variables, such as storms or fog. As these variables do not have linearity and can vary more frequently in large spaces of time due to climate changes involved, such a prediction is best defined when it occurs in a short time.

This paper describes and analyzes the application of statistical calibration method for simulation models called Model Output Calibration (MOC) with the aim to verify if there is a significant improvement in correcting the result of the forecast the weather forecast model Eta 15km at CPTEC/INPE.

Section 2 will describe some methods for simulation models correction, presenting its main features. Section 3 will address the Model Output Correction, high-

⁴ <http://www.cptec.inpe.br/>

⁵ <http://www.ncep.noaa.gov/>

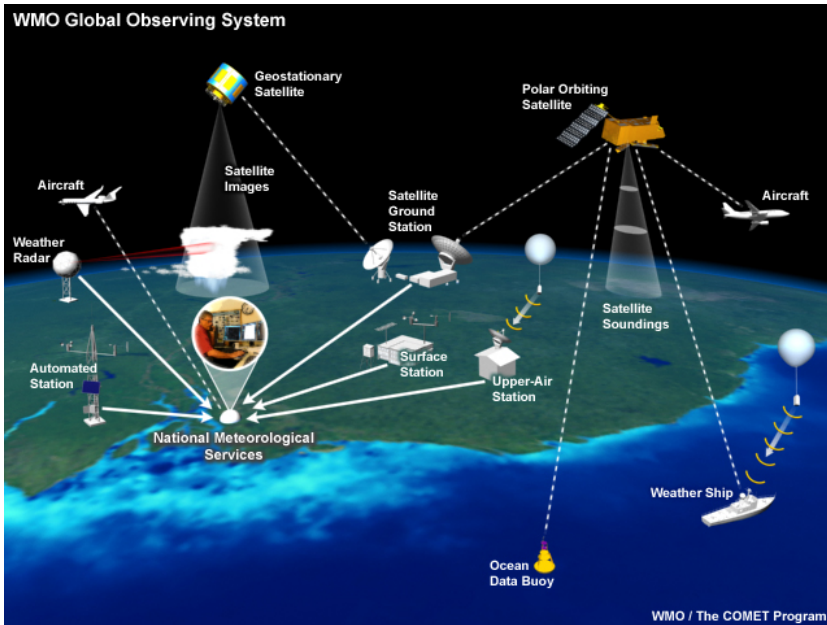


Fig. 1. An illustration of the instruments that comprise the Global Observing System (GOS) [11]

lighting its features and operation. In Section 4, use of the MOC in the correction of forecast generate by Eta model is described and the results obtained with this correction are presented and discussed. Finally, in section 5, the conclusions of this research are presented and aspects of future works to be developed from the results already obtained are discussed.

2 Correction methods of simulation models

Although forecast methods are becoming more sophisticated, with very accurate results, many of the variables provided by the models depend on variables that are susceptible to changes, such as those generated by actions of nature. According to Mao et al. [14], the temperature forecast models are known to have errors due to its coarse representation of the topographic model and a deficient physical model.

These differences are due to the fact that grid point model consider a single point by grid. Therefore, the model assumes that the climatic variables behave homogeneously within each grid. Although the simulation models will be improved, the weather models are far from perfect. However, in order to reduce these errors, besides the use of grid point mesoscale models in certain regions, some statistical methods have been developed for correcting (calibration) these errors.

The calibration data is the process of identifying possible problems in the forecast model and make adjustments in the calculations (codes). The process of calibration and validation are differentiated by the choice of parameters. Thus, the selection criteria used for data calibration may not be relevant to validate the model. Therefore, the selection of the parameters used for data calibration is a technical issue, where each statistical method is a own way. Among the major correction methods

of simulation models we can cite:

- Perfect Prognosis Method (PPM), also known as Perfect Prog: is a method developed by Klein et al. [10] in 1959 and it is the first statistical method for correction of simulated data, which takes into consideration the numerical weather prediction by applying multiple linear regression. An application of PPM is describe in [17];
- Model Output Statistics (MOS): is a method developed by Glahn and Lowry [8]. The MOS consists of determining a statistical relationship between the observed data as predictand and the variables of the numerical model in some time predictions. However, studies have shown that it takes two years of archived data for implementing this model in order to obtain an efficient MOS equation [9]. This fact leads to a deficiency in the model, because the MOS does not offer the possibility of developing a large database, because the numerical models are continuously modified, and for the implementation of MOS, it is necessary that during the filing period for data, the model configuration keep frozen. Comparisons between MOL and PPM are presented in [12,13];
- Calibration Model Output (MOC): was developed by Mao et al. [14] in 1998. It is a method that performs the statistical correction based on multivariate linear regression, as well as MOS and PPM. The main difference in relation to the MOC and the MOS and PPM is their short training period (2-4 weeks) of forecast and observations necessary for performing the correction of the statistical model.

3 Model Output Calibration

The MOC has as its principle the existence of error in the forecast, which may have been influenced by other output variables of the simulation model. The correction model MOC propose in finding the error of forecast through the application of multivariate linear regression for correction of the simulation model, requiring a training period of 2-4 weeks. Thus, the algorithm is used to estimate the necessary adjustments for the model used, predicting errors and making the refinement of the predicates. Therefore, it is obtained as a result of the correction applied by the MOC error of forecast in relation to a given variable, according to the equation (1):

$$\Delta T_{(ij)}(k) = T_{(ij)}(k) - T_{(ij)}^*(k) \quad (1)$$

where $T_{(ij)}(k)$ is a series of forecasts; (k) represents a number of variable values in different periods ($k = 1, 2, 3, \dots, N$); (i, j) refers to the grid point where the values are verified i.e., the location; and $T_{(ij)}^*(k)$ refers to the series of observed data. Soon, with the requirement within a certain period to a certain place, subtracted from the observed value in the same period and place, are obtained then the error of the forecast. Thus, it is assumed that $\Delta T_{(ij)}(k)$ (forecast error) are intrinsically associated with various parameters and model output variables. The forecast error can be a function $X_{ij,l}(k)$, where this is the l - th amount of variables obtained or

derived from a recent model output at the same point grid (i, j) , where the predicted error is correlated with the amount of L model outputs.

As an example, Chou et al. [4] show that the statistical correction of the temperature forecast there are several associated variables such as soil humidity and surface temperature. Based on these data it is possible to build a multivariate linear regression model, with which it is possible to obtain the error of forecast of model, according to equation (2):

$$\Delta T_{(ij)}(k) = \Lambda_o + \sum_{l=1}^L \Lambda_l X_{ij,l}(k) \quad (2)$$

where $k = 1, 2, \dots, N$ and $l = 1, 2, \dots, L$. The Λ_o and Λ_l coefficients can be defined by the analysis of multivariate linear regression. The $\Delta T_{(ij)}(k)$ is predictand equations, that are the dependent variable, and $X_{ij,l}(k)$ are the variables of the model output representing the predictors. According to Mao et al. [14] the selection predictors of MOC occurs with each new forecast cycle, based on a set of candidate variables, ie, with each new round of correction the predictors are selected (equation (2)), unlike other models of statistical correction where predictors are defined *a priori*.

Therefore, the number of selected predictors (Lc), direct or derivatives, should be less than the number of candidate variables (L). The calibration is applied after correction $N + 1$ where N is the size of the series to calculate the regression coefficients, obtaining thus refined forecast $T_{moc}(k)$ as the equation (3):

$$T_{moc}(k) = T_{(ij)}(k) - \Delta T_{(ij)}(k) \quad (3)$$

where $T_{(ij)}(k)$, is the prediction of the current model and $\Delta T_{(ij)}(k)$ is the error in the same point in the grid. Because the MOC require a short period of data to perform the correction adjustment, you can make adjustments in the models even if environmental changes occurred during the forecast period, making it a excellent tool to regional models.

Mao et al. [14] also showed other advantages of using the MOC: is a simple, flexible and efficient method; it is based on the latest weather conditions; the selection of predictors occurs objective and in a ideal way, because after predictors and candidate variables are chosen, the rest of the process occurs objective and iteratively; the algorithm used for statistical correction of the forecasting model is flexible and easy to implement for any model, because the equation is updated in each forecast cycle and is less sensitive to changes or reconfigurations of the model; information from the predictors selection results can be used in diagnostics and tuning of the physical model, because through this selection you can verify the variables that are influencing the forecast error.

Also according to Mao et al. [14], to evaluate if the calibrations forecasts have a satisfactory result, obtaining a significant gain in the relationship between the predicted and observed value are used three basic statistical measures of error: the mean error (BIAS) - (equation (4)), mean absolute error (MAE) - (equation (5)),

and the mean square error of the square root (RMSE) - (equation (6)). Through these measures it is possible to determine the quality of forecasting and calibration, as well as possible to determine strengths and weaknesses of the system.

$$BIAS = \frac{1}{M_{total}} \sum_{i=1}^{M_{total}} (T_i - T_i^*) \quad (4)$$

$$MAE = \frac{1}{M_{total}} \sum_{i=1}^{M_{total}} |T_i - T_i^*| \quad (5)$$

$$RMSE = \sqrt{\frac{1}{M_{total}} \sum_{i=1}^{M_{total}} (T_i - T_i^*)^2} \quad (6)$$

In equations (4), (5) and (6), M_{total} is the total amount of adjusted forecasts, T_i is the i -th value adjusted by MOC and T_i^* is the i -th observed value. T_i indicates the uncorrected value, thus being equal to that obtained by the forecast model. To analyze the results obtained from BIAS and MAE, as close to zero is the result, the better it is. The result of the RMSE is a non-negative value, however, is more sensitive to large errors because by raising the difference squared, makes negative and positive results will be treated equally [7].

These basic statistical measures of error, presented above, obtain your result in the same unit of the data used in the tests. However, it is often necessary to know the gain obtained with the calibration performed in percentage. For this there is the calculation of Skill Score (SS), according to equation (7), which precisely reflects the improvement percentage of the calibrated result of the forecast model on the outcome of forecast without statistical correction.

$$SS = \left(\frac{E_{forecast} - E_{MOC}}{E_{forecast}} \right) \times 100\% \quad (7)$$

where $E_{forecast}$ and E_{MOC} refer to the error obtained by the model forecast and the error of the calibrated model with the MOC, respectively. Both errors refer to the error obtained by using one of three error statistical measures: BIAS, MAE and RMSE, ie this index (SS) can be calculated for each of the errors measures mentioned above.

4 Use of MOC in the correction of Eta model forecasts

This section analyzes the results obtained by applying the MOC in the correction of temperature and relative humidity generated by the regional model Eta 15km from December 2015. The locations covered in the analysis are part of a set of 36 Brazilian locations, members of a project between the PPGCA/UPF and CPTEC/INPE allocated to the Brazilian Electric System National Operator (ONS). This project aims to improve the quality of energy load forecasting process in Brazil through these corrections.

METAR Station	Obs.	Eta	MOC	Eta			MOC		
				BIAS	MAE	RMSE	BIAS	MAE	RMSE
SBBE	27.24	26.68	27.24	-0.56	1.23	1.68	0.00	0.93	1.35
SBBH	24.55	22.15	24.66	-2.40	2.68	3.11	0.11	1.56	2.07
SBBR	23.22	21.90	23.47	-1.32	2.16	2.73	0.24	1.71	2.30
SBCG	25.21	27.47	25.04	2.26	2.67	3.52	-0.17	2.17	2.76
SBCH	23.17	21.51	22.65	-1.66	2.74	3.43	-0.52	2.30	2.88
SBCY	28.58	29.83	28.50	1.25	2.31	2.97	-0.08	1.87	2.47
SBCY	21.05	21.05	20.30	0.00	1.54	1.94	-0.75	1.73	2.24
SBFH	23.93	23.92	23.49	-0.00	1.56	2.02	-0.44	1.45	1.86
SBFZ	28.04	26.47	27.98	-1.57	1.72	1.91	-0.05	0.67	0.97
SBFH	24.69	26.71	24.07	2.01	2.53	3.21	-0.62	2.12	2.62
SBGO	25.37	24.45	25.55	-0.93	1.99	2.49	0.17	1.76	2.36
SBJP	26.66	26.08	26.77	-0.58	1.38	1.74	0.10	1.00	1.67
SBLO	24.55	23.87	24.04	-0.68	1.99	2.52	-0.51	1.87	2.38
SBMO	26.25	24.87	26.27	-1.38	1.68	2.00	0.02	0.92	1.60
SBMN	28.66	27.18	28.83	-1.48	2.06	2.51	0.17	1.41	1.84
SBMG	24.33	25.31	23.78	0.99	2.01	2.55	-0.54	2.07	2.59
SBNT	27.01	26.42	27.01	-0.59	1.15	1.42	-0.00	0.80	1.12
SBPJ	27.68	27.78	27.64	0.10	1.78	2.27	-0.04	1.73	2.23
SBPA	23.97	23.69	23.24	-0.29	1.84	2.30	-0.73	1.94	2.45
SBPV	26.55	26.74	27.18	0.19	1.82	2.35	0.62	1.39	1.97
SBDN	25.65	25.15	25.31	-0.50	2.23	2.83	-0.34	2.26	2.88
SBRF	27.50	26.76	27.62	-0.74	1.56	1.94	0.12	0.89	1.35
SBRP	25.41	24.25	25.32	-1.16	2.18	2.77	-0.09	1.85	2.51
SBRB	26.12	25.39	26.71	-0.73	1.64	2.08	0.60	1.51	2.08
SBAF	27.34	26.76	27.01	-0.59	1.68	2.17	-0.33	1.93	2.42
SBRJ	27.54	26.76	26.77	-0.78	1.68	2.09	-0.77	1.88	2.40
SBGL	27.57	26.76	26.95	-0.82	1.69	2.05	-0.62	1.99	2.44
SBJR	27.66	26.76	26.88	-0.90	1.67	2.10	-0.78	1.68	2.16
SBSV	27.73	26.21	27.65	-1.52	1.74	1.95	-0.08	0.76	1.02
SBSL	29.05	26.92	28.88	-2.13	2.14	2.32	-0.17	0.56	0.75
SBSP	23.26	23.15	22.92	-0.10	1.43	1.81	-0.33	1.72	2.25
SBGR	23.11	23.15	22.94	0.03	1.57	2.01	-0.18	1.60	2.15
SBMT	23.66	23.15	23.43	-0.51	1.56	1.96	-0.23	1.65	2.17
SBTE	31.79	29.46	31.62	-2.33	2.91	3.47	-0.17	1.28	1.92
SBUG	24.51	24.75	24.06	0.23	2.28	3.00	-0.46	2.32	2.96
SBVT	27.78	25.94	27.55	-1.84	2.26	2.57	-0.23	1.18	1.59

Table 1

Daily average surface air temperature ($^{\circ}\text{C}$) at 2m observed, predicted by the Eta model, adjusted by the MOC and their corresponding error basic statistical measures (BIAS, MAE and RMSE) for the month of December 2015 to METAR stations located in Brazil

The locations covered in the project include meteorological stations located airports from capitals and in some cities in the interior of Sao Paulo State and in the South region from Brazil, regions where the demand for energy because of agriculture and irrigation process is more significant. The current computer system was developed for the Linux operating system, using Linux shell scripts and Fortran language. For the statistical analysis is used statistical language R [16].

During the data analysis period it was possible to identify possible problems in the forecast generated by Eta model, where the forecast results, in some cases, proved to be far from the observed data, as can be observed in respect of air tem-

METAR Station	Obs.	Eta	MOC	Eta			MOC		
				BIAS	MAE	RMSE	BIAS	MAE	RMSE
SBBE	78.84	85.04	78.79	6.21	11.18	13.74	-0.04	6.07	8.96
SBBH	66.77	77.52	66.92	10.75	14.68	17.58	0.15	7.90	10.03
SBBR	71.49	73.84	71.93	2.35	11.05	14.75	0.43	9.77	13.11
SBCG	74.65	74.89	74.40	0.25	12.66	15.41	-0.24	9.04	11.22
SBCH	76.78	76.48	76.53	-0.30	16.75	20.50	-0.25	10.47	13.19
SBCY	68.27	65.53	70.60	-2.74	12.59	15.57	2.33	8.37	10.53
SBCT	82.43	78.16	84.08	-4.27	13.67	17.39	1.65	7.73	11.00
SBFL	83.05	79.40	81.97	-3.65	13.64	16.88	-1.08	6.76	8.70
SBFZ	71.82	83.45	71.52	11.63	14.41	16.45	-0.29	4.35	6.24
SBFI	83.59	77.46	83.95	-6.14	11.01	14.34	0.35	7.26	9.22
SBGO	64.78	74.23	65.64	9.45	14.49	17.62	0.85	8.98	11.78
SBJP	73.98	83.99	72.40	10.01	12.01	14.13	-1.58	5.34	7.70
SBLO	83.06	77.11	83.09	-5.95	13.22	18.23	0.03	6.45	8.56
SBMO	81.20	83.54	79.73	2.34	7.74	9.70	-1.47	5.45	7.96
SBMN	71.78	75.92	71.56	4.15	11.80	14.73	-0.21	8.08	10.09
SBMG	83.51	77.99	83.39	-5.52	11.95	15.57	-0.117	7.53	9.35
SBNT	73.71	78.00	72.48	4.29	11.91	13.98	-1.24	5.03	6.95
SBPJ	75.23	74.68	76.96	-0.55	11.19	13.60	1.74	7.51	9.87
SBPA	78.02	77.09	76.65	-0.93	12.35	15.45	-1.37	7.92	10.56
SBPV	80.26	81.09	78.83	0.83	12.13	14.54	-1.44	5.76	7.98
SBDN	73.75	76.36	74.94	2.61	13.71	17.05	1.18	10.16	12.93
SBRF	73.64	71.76	72.09	-1.88	23.42	25.79	-1.55	5.26	7.70
SBRP	76.04	80.59	76.45	4.55	10.51	13.63	0.41	9.31	12.38
SBRB	83.70	82.63	80.86	-1.07	9.66	12.45	-2.84	6.50	8.78
SBAF	72.19	75.40	73.00	3.21	11.43	13.54	0.80	8.12	10.38
SBRJ	77.79	75.40	81.05	-2.39	10.03	12.52	3.26	8.05	10.40
SBGL	71.40	75.40	73.20	4.00	11.04	13.17	1.80	7.88	10.13
SBJR	75.78	75.40	79.52	-0.38	9.63	12.23	3.74	8.81	11.32
SBSV	71.48	80.30	72.74	8.82	16.17	18.71	1.26	5.41	7.03
SBSL	63.62	81.64	64.26	18.02	18.94	21.91	0.63	3.85	5.59
SBSP	73.71	75.14	74.88	1.43	12.57	15.40	1.17	7.32	9.61
SBGR	84.73	75.14	84.95	-9.59	12.56	17.46	0.22	7.41	9.50
SBMT	72.82	75.14	73.76	2.32	12.87	15.39	0.93	6.76	8.90
SBTE	50.37	53.72	51.23	3.35	19.52	22.27	0.86	6.91	9.93
SBUG	76.48	75.15	72.71	-1.33	13.91	17.51	-3.77	12.08	15.56
SBVT	75.28	79.38	77.26	4.11	12.57	14.98	1.98	7.42	9.49

Table 2
Daily average relative humidity (%) observed, predicted by the Eta model, adjusted by the MOC and their corresponding error basic statistical measures (BIAS, MAE and RMSE) for the month of December 2015 to METAR stations located in Brazil

perature in some METAR stations, such as SBBH, SBCG, SBFI, SBMN, SBSL and SBTE. Complete data for all stations are shown in Table 1. This table shows the observed data, provided by the Eta model and corrected by MOC method. In Table 2 presents the complete data for relative humidity and, even in this case, the same potential problems reappear, as shown for example in METAR stations SBBH, SBFZ, SBJP, SBSV, SBSL and SBGR.

Refinement obtained by MOC suggests an improvement in the data, as can be verified, for example, in the stations in Belo Horizonte (SBBH), São Luís (SBSL) and Fortaleza (SBFZ) cities (Figures 2, 3 and 4, respectively). These blue curves in

METAR Station	SS MOC temperature			SS MOC relative humidity		
	Bias (%)	Mae (%)	Rmse (%)	Bias (%)	Mae (%)	Rmse (%)
SBBE	100.46	24.39	19.64	100.76	45.71	34.79
SBBH	104.58	41.79	33.44	98.56	46.19	42.95
SBBR	118.18	20.83	15.75	81.49	11.58	11.12
SBCG	107.52	18.73	21.59	196.80	28.59	27.19
SBCH	68.67	16.06	16.03	16.33	37.49	35.66
SBCY	106.56	19.05	16.84	185.04	33.52	32.37
SBCT	13031.03	-12.34	-15.46	138.64	43.45	36.75
SBFL	-11182.05	7.051	7.921	70.41	50.44	48.46
SBFZ	96.31	61.05	49.21	102.56	69.81	62.07
SBFI	130.85	16.21	18.38	105.81	34.06	35.70
SBGO	118.28	11.56	5.221	90.93	38.03	33.14
SBJP	117.24	27.54	4.023	115.78	55.54	45.51
SBLO	25.00	6.03	5.556	100.54	51.21	53.04
SBMO	101.45	45.24	20.00	162.82	29.59	17.94
SBMN	111.49	31.55	26.69	105.28	31.53	31.50
SBMG	154.55	-2.985	-1.569	97.88	36.99	39.95
SBNT	99.78	30.43	21.13	128.90	57.77	50.29
SBPJ	140.00	2.809	1.762	416.36	32.89	27.43
SBPA	-151.72	-5.435	-6.522	-47.31	35.87	31.65
SBPV	-226.32	23.63	16.17	273.49	52.51	45.12
SBDN	32.00	-1.345	-1.767	54.79	25.89	24.16
SBRF	116.22	42.95	30.41	17.55	77.54	70.14
SBRP	92.16	15.14	9.386	90.97	11.42	9.17
SBRB	182.19	7.927	0.00	-165.42	32.71	29.48
SBAF	44.07	-14.88	-11.52	74.83	28.96	23.34
SBRJ	1.282	-11.90	-14.83	236.40	19.74	16.93
SBGL	24.39	-17.75	-19.02	55.00	28.62	23.08
SBJR	13.33	-0.5988	-2.857	1084.21	8.515	7.44
SBSV	94.41	56.32	47.69	85.71	66.54	62.43
SBSL	92.02	73.83	67.67	96.45	79.67	74.49
SBSP	-230.00	-20.28	-24.31	18.18	41.77	37.60
SBGR	586.49	-1.911	-6.965	102.38	41.00	45.59
SBMT	54.90	-5.769	-10.71	59.74	47.47	42.17
SBTE	92.70	56.01	44.67	74.27	64.60	55.41
SBUG	300.00	-1.754	1.333	-183.46	13.16	11.14
SBVT	87.50	47.79	38.13	51.82	40.97	36.65

Table 3
Skill score of measures BIAS, MAE and RMSE related to daily average air temperature and relative humidity for the month of December 2015 to METAR stations located in Brazil

the figures refer to the data observed in the meteorological station, the red curves refer to the measures provided by the Eta model and the green curves refer to corrected data by MOC.

The average daily cycle of air temperature and relative humidity may indicate periods of the day that errors often occur in the forecast at model Eta where it was observed that, in general, the temperature forecast correctly follows the amplitude variation, but however, the most of the time, underestimated the measures. In the relative humidity forecast, the largest errors occur between 00:00 and 9:00 am, when the measures forecast by Eta have been overestimated.

Table 3 shows the values of the Skill Score for the corrections performed. In

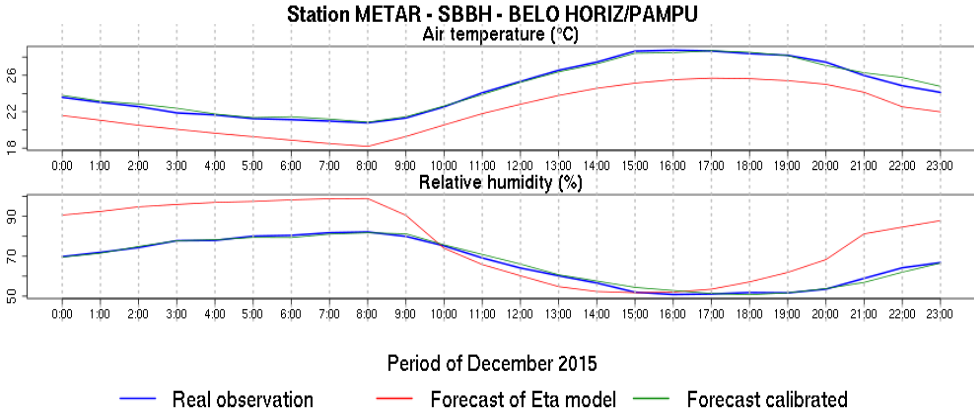


Fig. 2. Average daily cycle of the air temperature at 2m ($^{\circ}C$) and relative humidity (%) from December 2015 in METAR station SBBH located in the city of Belo Horizonte

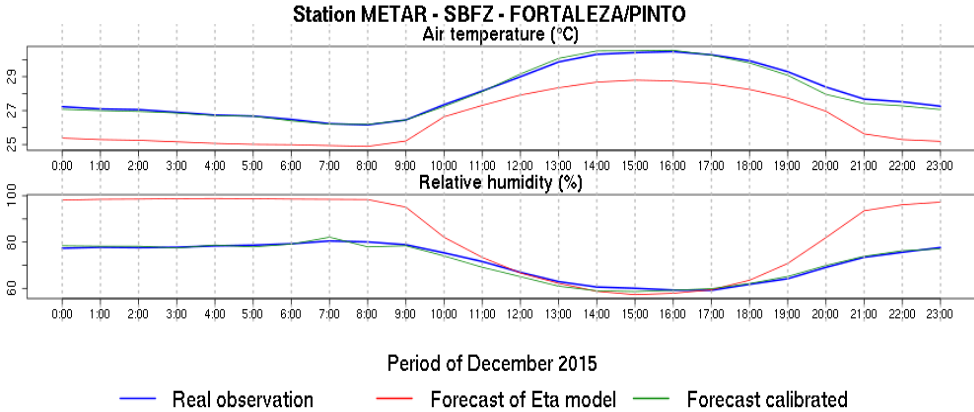


Fig. 3. Average daily cycle of the air temperature at 2m ($^{\circ}C$) and relative humidity (%) from December 2015 in METAR station SBFZ located in the city of Fortaleza

this table can be noted the gain obtained with the application of the MOC in most locations analyzed, which demonstrates the good potential for use of MOC in correction of weather forecast generated by Eta model.

5 Conclusions and future works

This paper presents the use of the method Model Output Calibration for statistical refinement in a numerical weather prediction model, verifying its effectiveness in the use over the Eta model. The application of refinement showed a positive improvement over the forecast uncorrected from Eta model, enabling an improvement in the quality of the result of the forecast. The short training period required for carrying out the calibration and therefore its easy adaptation to changes in the model to be corrected is one of its advantages over other calibration methods.

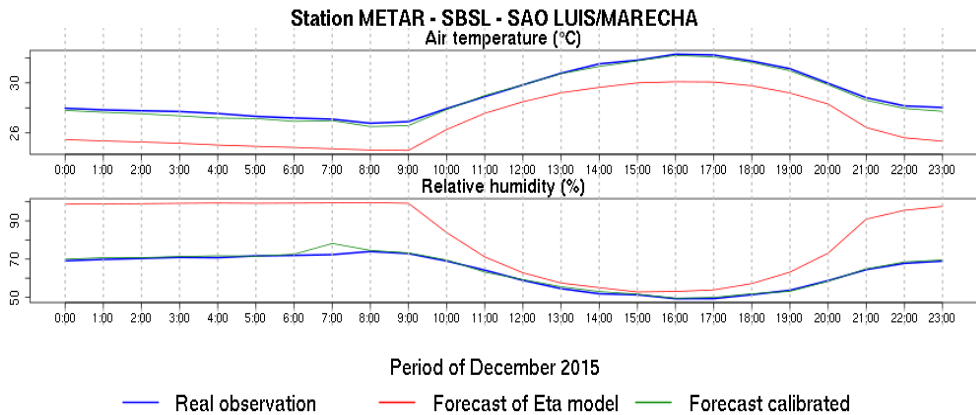


Fig. 4. Average daily cycle of the air temperature at 2m ($^{\circ}\text{C}$) and relative humidity (%) from December 2015 in METAR station SBSL located in the city of São Luís

On average, the errors of the forecast temperature of Eta model did not exceed the magnitude of 2°C , but generally underestimated the observations, mainly during the morning. On the other hand, the relative humidity was generally overestimated during the night period and underestimated during the day. The tendency to produce more saturated air during the night can be a result of underestimation of temperature forecasts at some stations. While the tendency to produce excessively dry air during the day seems to be derived from error in the estimate of specific humidity (amount of steam per total air volume) by the model. In stations in which prediction errors were considered relatively small, generally the correction scheme eventually produced an error.

The largest temperature errors occurred in stations where the model grid box used for comparison was far from the point of observation. Numerical forecasts of dynamic models always contain errors. The extracted information model grid represents the average inside that grid, and the average value of the time step of the integration model. The comparison of average information to the specific measure provided by observation is already an inconsistency and an error factor.

As future works, this research enables you to check the validity of using the Eta model forecast data corrected by MOC and the gains that can be achieved if when they are used in simulation models of crops and diseases that are used in research developed at PPGCA/UPF in partnership with Embrapa Trigo, the University of Florida and the CPTEC/INPE.

References

- [1] Black, T.L., The new NMC mesoscale Eta model: Description and forecast examples, *Weather and Forecasting*, **9**(2) (1994), 265–278, URL: [http://journals.ametsoc.org/doi/abs/10.1175/1520-0434\(1994\)009%3C0265%3ATNNMEM%3E2.0.CO%3B2](http://journals.ametsoc.org/doi/abs/10.1175/1520-0434(1994)009%3C0265%3ATNNMEM%3E2.0.CO%3B2).
- [2] Chou, S.C., *Modelo Regional Eta. Climanálise*, Especial Edition, Instituto Nacional de Pesquisas Espaciais, vol.1, 1996.

- [3] Chou, S.C., and Bustamante, J.F. and Gomes, J.L., *Evaluation of Eta Model Seasonal Precipitation Forecasts over South America*, *Nonlinear Processes in Geophysics*, **12** (2005), 537–555, DOI: <http://dx.doi.org/10.5194/npg-12-537-2005>.
- [4] Chou, S. C., Souza, C. R., Gomes, J. L., Evangelista, E. F. D., Osório, C. and Cataldi, M., Refinamento estatístico das previsões horárias de temperatura a 2m do modelo Eta em estações do nordeste do Brasil, *Revista Brasileira de Meteorologia*, **22**(3) (2007), 287–296.
- [5] Chou, S. C., Lyra, A., Mourão, C., Dereczynski, C., Pilotto, I., Gomes, J.L., Bustamante, J., Tavares, P., Silva, A., Rodrigues, D., Campos, D., Chagas, D., Sueiro, G., Siqueira, G., Nobre, P. and Marengo, J., Evaluation of the Eta Simulations Nested in Three Global Climate Models, *American Journal of Climate Change*, **3**(05) (2014), 438–454, URL: <http://www.scirp.org/journal/PaperInformation.aspx?PaperID=52877>.
- [6] INPE. *Eta Model*, Instituto Nacional de Pesquisas Espaciais, 2016, URL: <http://etamodel.cptec.inpe.br>.
- [7] Ferreira, S. P., Varone, F. and Alvez, R. C. M., Uso do TRMM para a correção estatística da chuva prevista pelo modelo BRAMS, *IV Simpósio Brasileiro de Ciências Geodésicas e Tecnologias da Geoinformação* (2012), 01–09.
- [8] Glahn, H. R. and Lowry, D. A., The use of model output statistics (mos) in objective weather forecasting, *J. Appl. Meteor.*, **11** (1972), 1203–1211.
- [9] Jacks, E. J. B., Bower, V. J., Dagostro, J. P., Dallavalle, M. C. and Erickson, J. C. S., New ngm-based mos guidance for maximum/minimum temperature, probability of precipitation, cloud amount, and sea surface wind, *Wea. Forecasting*, **5** (1990), 128–138.
- [10] Klein, B., Lewis, M. and Enger, I., Objective prediction of 5 day mean temperature during winter, *Journal of Meteorology*, **16** (1959), 672–682.
- [11] Laing, A. and Jenni-Louise Evans, J., “Introduction to Tropical Meteorology,” 2nd Ed., University Corporation for Atmospheric Research, Boulder (2011), URL: <http://www.goes-r.gov/users/comet/tropical/textbook/2nd.edition/index.htm>.
- [12] Maini, P., Kumar, A., Rathore, L.S. and SINGH, S.V., Forecasting Maximum and Minimum Temperatures by Statistical Interpretation of Numerical Weather Prediction Model Output, *Weather and Forecasting*, **18**, (2003), 938–952, URL: [http://journals.ametsoc.org/doi/pdf/10.1175/1520-0434\(2003\)018%3C0938%3AFMAMTB%3E2.0.CO%3B2](http://journals.ametsoc.org/doi/pdf/10.1175/1520-0434(2003)018%3C0938%3AFMAMTB%3E2.0.CO%3B2).
- [13] Marzban, C., Sandgathe, S. and Kalnay, E., MOS, Perfect Prog, and Reanalysis, *Monthly Weather Review*, **134**, (2006), 657–663, URL: <http://journals.ametsoc.org/doi/pdf/10.1175/MWR3088.1>.
- [14] Q. Mao, R. T. McNider, S. F. Muellere, and H. Juang, An optimal model output calibration algorithm suitable for objective temperature forecasting, *American Meteorological Society*, **14** (1998), 190–202.
- [15] Mesinger, F., Janjić, D. A., Ničković, S., Gavrilo, D. and Deaven, D. A., 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*, **7**(116) (1988), 1493–1518.
- [16] R Core Team, “R: A language and environment for statistical computing”, R Foundation for Statistical Computing: Vienna, 2016, URL: <https://www.r-project.org/>.
- [17] Rajeevan, M., Madhulatha, A., Rajasekhar, M., Bhate, J., Kesarkar, A. and Appa Rao, B.V., Development of a perfect prognosis probabilistic model for prediction of lightning over south-east India, *Journal of Earth System Science*, **121**(2) (2012), 355–371, DOI: <http://dx.doi.org/10.1007/s12040-012-0173-y>.