Confidence predictive distributions: an application to temperature forecasting in Veneto

PROCEEDINGS

Edited By: Giada Adelfio and Antonino Abbruzzo

GRASPA 2023

GRASPA-SIS BIENNAL CONFERENCE

The Researcher Group for Environmental Statistics of The Italian Statistical Society

TIES EUROPEAN REGIONAL MEETING

The International Environmetrics Society

Palermo, 10-11 July, 2023

Dipartimento di Scienze Economiche Aziendali e Statistiche, Università degli Studi di Palermo

Proceedings of the GRASPA 2023 Conference
Palermo, 10-11 July 2023
Edited by: Giada Adelfio and Antonino Abbruzzo
-
Palermo: Università degli Studi di Palermo.

ISBN: 979-12-210-3389-2

Sponsored by:

Contents

Confidence predictive distributions: an application to temperature forecasting in Veneto

F. Giummolè [1] and V. Mameli[2,*]

[1] *Department of Environmental Sciences, Informatics and Statistics, Ca' Foscari University of Venice; giummole@unive.it,*
[2] *Department of Economics and Statistics, Udine; valentina.mameli@uniud.it*
[*] *Corresponding author*

Abstract. *Post-processing techniques are nowadays frequently used in order to reduce the impact of errors in ensemble forecasts of meteorological variables. Ensemble model output statistics (EMOS) are a widely spread post-processing approach built on a heteroscedastic linear regression model. After replacing unknown parameters with suitable estimates, an estimative EMOS distribution function for prediction is obtained. However, it is well known that forecasts based on estimative EMOS may lack calibration, particularly when the number of ensembles is large compared to the number of historical observations. Here, we suggest overcoming this drawback by applying in the EMOS context a predictive approach based on the concept of confidence distribution. The result is a new predictive distribution that takes the form of a variance correction of the classical estimative EMOS distribution. The performance of the confidence EMOS distribution is tested on a real-data application for temperature forecasting. It can be seen that our proposal performs better than the classical estimative EMOS, both in terms of coverage probabilities and log-score.*

Keywords. *Confidence distribution; Coverage probabilities; EMOS.*

1 Introduction

Forecasting the weather is an essential component of support to decision making in many different situations. Forecasts have gradually improved over the past few decades, in part as a consequence of advancements in numerical weather prediction (NWP) ([1]). The forecasts produced by physics-based models, which are typically provided as forecast ensembles, still exhibit systematic bias and are frequently under-dispersive despite these advances ([2]). Today, it is common practice to further refine, enhance, and calibrate NWPs using statistical post-processing techniques. One of the most widely used methods for post-processing ensemble forecasts is ensemble model output statistics (EMOS, [4]).

Classic EMOS is nothing but a linear heteroscedastic regression model with Gaussian errors, where the ensembles play the role of explanatory variables and their sample variance contributes to the variance component. Unknown parameters are estimated on the basis of past observations by minimising some suitable scoring rules. After substituting parameter estimates in the Gaussian distribution function of the variable of interest, a so-called estimative distribution is obtained. Estimative EMOS distribution functions are commonly used to forecast the variable of interest and, in particular, to obtain prediction intervals or limits with the desired level of confidence. Unfortunately, even if this procedure is able to correct for bias and under-dispersion of the ensemble, it is not fully calibrated, and the actual coverage of its quantiles may consistently differ from the nominal one.

Ad-hoc techniques, useful for modelling specific datasets, have been proposed in the literature to overcome this problem. A method that can be applied in the general context of EMOS, as well as with distributions that differ from the normal case, is considered in [5], where an easy bootstrap procedure is used for calibrating estimative EMOS distributions. Here, we consider the approach to prediction based on confidence distributions, presented by [6], and capable to include frequentist, Bayesian, and fiducial predictive inference within a unique framework. With this method, unknown parameters are eliminated by integration with respect to a so-called confidence distribution. Depending on the features of the model and of the chosen confidence distribution, the resulting confidence predictive distribution is, at least approximately, well-calibrated, giving quantiles with the correct coverage probability.

This work is a seminal attempt to use confidence predictive distributions in the context of EMOS. Using data about maximum daily temperatures in the Veneto region, north of Italy, we show the potentiality of this method for improving the usual estimative approach. In the next section, we present some basic concepts on EMOS and confidence-based prediction and obtain the confidence EMOS predictive distribution. In the last section, we apply the result to the problem of forecasting temperatures in the Veneto region. In particular, we use data regarding Cavallino-Treporti station, located in the Venice lagoon. We show the superiority of the new predictive distribution on the usual estimative distribution, both with respect to the log-score and the coverage of predictive quantiles.

2 EMOS prediction with confidence

The simplest version of EMOS is just a regular linear regression model with heteroscedastic normal errors. The ensemble members are combined linearly to form the EMOS mean, with the contribution of ensemble members to the relevant weather variable represented by unknown coefficients. The EMOS variance, which considers the spread relationship, is a linear function of the ensemble variance. To set the notation, consider that $\{Y_i\}_{i\geq 1}$ is a sequence of independent continuous random variables. $Y^n = (Y_1,\dots,Y_n)$, $n>1$, is observable while $Z = Y_{n+1}$ is a future or not yet observed variable. In the classic EMOS all the Y_i's, $i \geq 1$, are normally distributed with mean and variance depending on m ensemble members. Let $X^{n+1} = (1, X_{1,n+1},\dots,X_{m,n+1})^T$ include 1 for the intercept of the model and the ensembles associated with the future variable Z, and let S^2_{n+1} be the variance of the ensembles $(X_{1,n+1},\dots,X_{m,n+1})$. The full distribution of Z is given by

$$Z \mid X_{1,n+1},\dots,X_{m,n+1} \sim \Phi\left(\frac{z-\mu_{n+1}}{\sigma_{n+1}}\right), \tag{1}$$

where $\Phi(\cdot)$ denotes the standard normal distribution function and $\mu_{n+1} = \beta_0 + \beta_1 X_{1,n+1} + \dots + \beta_m X_{m,n+1} = \beta X^{n+1}$ and $\sigma^2_{n+1} = \gamma_0 + \gamma_1 S^2_{n+1}$, see [4]. The parameters $\beta = (\beta_0,\dots,\beta_m)$, γ_0 and γ_1 are non-negative unknown coefficients. Log-score and CRPS are two appropriate scoring rules that are typically minimised in order to estimate unknown EMOS parameters. An estimative distribution for the future weather quantity Z is obtained by substituting the estimated parameters in the full distribution of Z:

$$\Phi\left(\frac{z-\widehat{\mu}_{n+1}}{\widehat{\sigma}_{n+1}}\right), \tag{2}$$

with $\widehat{\mu}_{n+1} = \widehat{\beta} X^{n+1}$, and $\widehat{\sigma}^2_{n+1} = \widehat{\gamma}_0 + \widehat{\gamma}_1 S^2_{n+1}$, where $\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_m)$, $\hat{\gamma}_0$, and $\hat{\gamma}_1$ are suitable estimates of β, γ_0, and γ_1, respectively, based on an observed sample from Y^n, $y^n = (y_1,\dots,y_n)$.

Unfortunately, estimative distributions occasionally exhibit poor performance, particularly when there are few historical observations in comparison to the size of the ensemble. In order to improve the estimative EMOS (2) we consider the method proposed in [6] based on the concept of confidence distribution. According to [6] a whole predictive distribution function for the variable Z can be obtained by integrating (1) with respect to a confidence distribution for the unknown parameters. The properties of the resulting predictive distribution for Z are strictly related to those of the used confidence distribution. Let X be the $n \times (m+1)$ design matrix, namely, $X = [X^1, \ldots, X^n]^T$ with $X^i = (1, X_{1,i}, \ldots, X_{m,i})^T$ for $i = 1, \ldots, n$. Following [6], we choose a confidence distribution for β derived from the asymptotic distribution of $\widehat{\beta}$:

$$\Phi_{m+1}\left(\left(X^T \Sigma^{-1} X\right)^{1/2} (\widehat{\beta} - \beta)\right), \tag{3}$$

where Φ_{m+1} denotes a $(m+1)-$dimensional normal distribution with zero mean vector and unit covariance matrix, $\Sigma = \text{diag}(\gamma_0 + \gamma_1 S^2)$ with $S^2 = (S_1^2, \ldots, S_n^2)$, $S_i^2 = \frac{1}{m-1} \sum_{j=1}^m (X_{ij} - \bar{X}_i)^2$ denoting the variance of the i-th observation of the ensemble and $\bar{X}_i = \frac{1}{m} \sum_{j=1}^m X_{ij}$ its mean, $i = 1, \ldots, n$.

By integrating (1) with respect to (3) and replacing Σ with $\widehat{\Sigma} = \text{diag}(\widehat{\gamma}_0 + \widehat{\gamma}_1 S^2)$ and σ_{n+1}^2 with $\widehat{\sigma}_{n+1}^2 = \widehat{\gamma}_0 + \widehat{\gamma}_1 S_{n+1}^2$ we obtain the corresponding confidence predictive distribution:

$$Q(z; y^n) = \Phi\left(\frac{z - \widehat{\mu}_{n+1}}{\sqrt{\widehat{\sigma}_{n+1}^2 + (X^{n+1})^T \left(X^T \widehat{\Sigma}^{-1} X\right)^{-1} X^{n+1}}}\right). \tag{4}$$

The new confidence predictive distribution (4) is nothing but a variance adjustment of the estimative EMOS distribution (2). It corresponds to the exact predictive distribution in heteroscedastic regression models, for the case when the variance-covariance matrix of the errors is known, see for instance [3]. Indeed, our procedure accounts for the additional uncertainty in the estimates of the regression parameters but not in the estimate of the variance component. Thus, the resulting predictive distribution is not fully calibrated. Nonetheless, as shown in the next section, it provides prediction limits with a coverage probability very close to the nominal one.

3 Real case study

In this study, we focus on forecasts for the maximum daily temperatures at stations located in the Veneto region, northern Italy. We have a three-year period of interest that runs from August 16, 2009, to August 17, 2012. The dataset was previously analysed in [5] by using the classical EMOS and a bootstrap calibrated modification of it. Historical maximum daily temperature forecasts are available from http://www.scia.isprambiente.it/. While the ensemble forecasts are provided by the World Climate Research Programme. We examine the maximum daily temperatures for the Venetian lagoon station Cavallino-Treporti (Longitude: 12.48642, Latitude: 45.45805). In accordance with [4], we use a sliding window of 50 observations as our training set and the remaining 1039 days as our test set. Estimates of the EMOS parameters are obtained by optimizing the log-score (namely, minus the log-likelihood) over the sliding training period. Then, we consider the estimative EMOS in (2), the bootstrap predictive distribution obtained in [5], as well as the predictive distribution (4) produced by the suggested methodology. We compare the performances of the three distributions for each of the 1039 observations in terms of coverage probability and log-score. The mean and standard deviation of the log-score for the three predictive models are summarised in Table 1. Note that the confidence predictive distribution improves on the estimative distribution, since the lower the score the better the method. We have also obtained coverage probabilities and mean lengths of level 66.7% central intervals to properly assess the calibration and concentration of the various predictive models (Table 2). It can be seen that the confidence predictive distribution performs much better than the bootstrap predictive distribution in terms of coverage probability of the central prediction interval. The increased length of the interval is in fact justified by the higher and more precise coverage. Additionally, the coverage probabilities of the upper prediction limits of 90%, 95%, and 99% are derived and represented in Figure 1. In relation to the estimative distributions, it should be noted that the coverage probabilities obtained from the proposed predictive distribution are closer to nominal levels, and to those of the Bootstrap predictive distribution. The PIT histograms for the three considered predictive models are finally shown in Figure 2. We can see that the histogram produced by the suggested methodology is very similar to the uniform one, indicating good calibration. The excessive under-dispersion of the estimative EMOS distribution gives the PIT histogram its particular U shape.

As we have seen, in terms of all the considered measures, the confidence predictive distribution performs similar to the Bootstrap predictive distribution discussed in [5], but it has the advantage of not requiring any time-consuming computational process like the bootstrap itself.

	Est	Boot	Conf
Log-score	2.98	2.51	2.55
	(0.08)	(0.03)	(0.03)

Table 1: Log-score values of the three predictive distributions. Est denotes the estimative EMOS, Boot the Bootstrap predictive distribution obtained in [5], while Conf denotes the confidence predictive distribution in (4).

	Est	Boot	Conf
mean length	3.53	5.42	6.18
	(0.02)	(0.04)	(0.05)
coverage	0.451	0.641	0.692
	(0.015)	(0.015)	(0.014)

Table 2: Coverage probabilities and mean lengths of the central prediction interval of level 0.67 for the three predictive distributions. Est denotes the estimative EMOS, Boot the Bootstrap predictive distribution obtained in [5], while Conf denotes the confidence predictive distribution in (4). Standard errors in brackets.

References

[1] Bauer, P., Thorpe, A., Brunet, G. (2015). The quiet revolution of numerical weather prediction. *Nature*,**525**, 47–55.

[2] Buizza, R. (1997). Potential forecast skill of ensemble prediction and spread and skill distributions of the ECMWF ensemble prediction system. *Montly Weather Review*, **125**, 99–119.

[3] Faraway, J.J. (2005). Linear Models with R. Chapman & Hall.

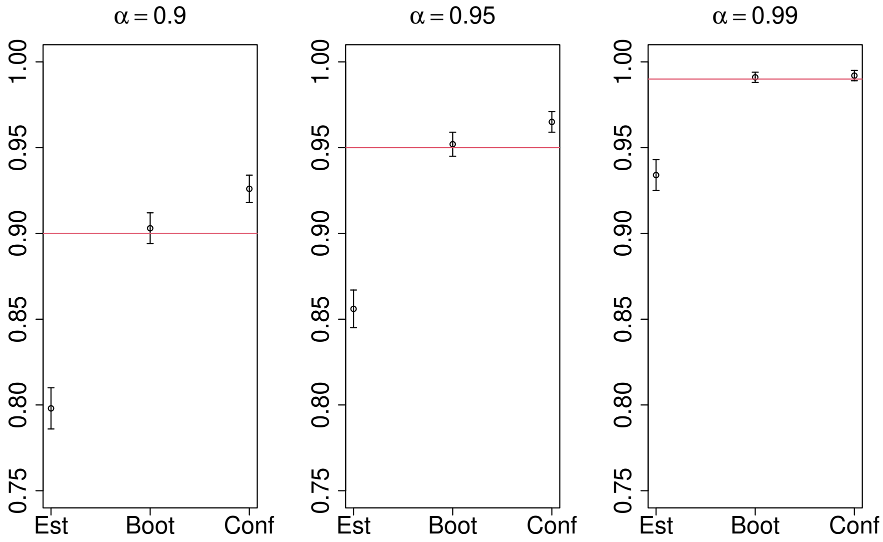

Figure 1: Coverage probabilities of upper prediction limits for the three predictive distributions. Left panel $\alpha = 0.90$, middle panel $\alpha = 0.95$, right panel $\alpha = 0.99$. Est denotes the estimative EMOS, Boot the Bootstrap predictive distribution obtained in [5], while Conf denotes the confidence predictive distribution in (4).

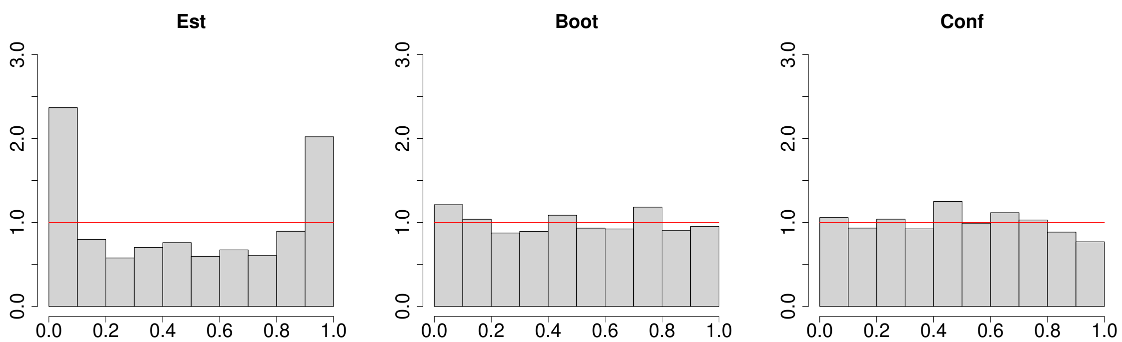

Figure 2: PIT histograms of the three predictive distributions. Est denotes the estimative EMOS, Boot the Bootstrap predictive distribution obtained in [5], while Conf denotes the confidence predictive distribution in (4).

[4] Gneiting, T., Raftery, A.E., Westveld III, A.H., Goldman , T. (2005). Calibrated Probabilistic Forecasting Using Ensemble Model Output Statistics and Minimum CPRS Estimation. *Monthly Weather Review*, **133(5)**, 1098–1118.

[5] Gaetan, C., Giummolè, F., Mameli, V., Siad, S. (2022). Ensemble model output statistics for temperature forecasts in Veneto. Book of short papers of the 51st scientific meeting of the Italian Statistical Society, 1253–1258.

[6] Shen, J., Liu, R.Y., Xie, M. (2018). Prediction with confidence-A general framework for predictive inference. *Journal of Statistical Planning and Inference*, **195**, 126–140.