CALIBRATION INDONESIAN-NUMERICAL WEATHER PREDICTION USING GEOSTATISTICAL OUTPUT PERTURBATION

Sutikno[1]*, Fajar Dwi Cahyoko[1], Fernaldy Wananda Putra[1], Erwin Eka Syahputra Makmur[2], Wido Hanggoro[2], Muhamad Rifki Taufik[2] and Vestiana Aza[2]

[1]Department of Statistics, Faculty of Science and Data Analytics, Institut Teknologi Sepuluh Nopember, Sukolilo, Surabaya, 60115

[2] Meteorology, Climatology and Geophysics Agency, Jakarta Pusat, 10610

*E-mail: sutikno@statistika.its.ac.id

Article submitted: August 25, 2023 Article revised: January 26, 2024 Article accepted: February 23, 2024

ABSTRACT

Indonesian-Numerical Weather Prediction (INA-NWP) is a numerical-based weather forecast method that has been developed by the Meteorology, Climatology and Geophysics Agency. However, the forecast is still unable to produce accurate weather forecasts. Geostatistical Output Perturbation (GOP) is a weather forecast method derived from only one deterministic output. GOP takes into consideration the spatial correlation among multiple locations simultaneously. GOP is capable to identify spatial dependency patterns that are associated with error models. This study aims to obtain calibrated forecasts for daily maximum and minimum temperature variables using GOP at 10 meteorological stations in Surabaya and surrounding areas. The stages in performing temperature forecasts using GOP are obtaining regression coefficient estimators, then calculating empirical semivariograms and estimating spatial parameters. Based on several weather forecast indicators, such as RMSE and CRPS, GOP is better than INA-NWP in terms of precision and accuracy.

Keywords: calibration, GOP, INA-NWP, temperature

1. Introduction

Nowadays, accurate weather information encourages several countries, including Indonesia to develop a numerical-based weather prediction system called Numerical Weather Prediction. The basic concept of NWP is to solve a set of partial differential equations (PDEs) that govern the movement and evolution of the atmosphere [1]. Meteorology, Climatology and Geophysics Agency of Indonesia (BMKG) is developing the Indonesia Numerical Weather Prediction (INA-NWP) model that is suitable for Indonesia's geographical and climatic conditions. Indonesia has very special characteristics, namely consisting of vast land and oceans, so the INA-NWP model was developed to be able to produce more accurate weather forecasts [2], [3].

The NWP still produces bias [4], does not fully explain the meteorological stochastic process [5], and still needs to be calibrated. Statistical postprocessing eliminates systematic error present in the NWP models by recognizing and adjusting connections between past forecasts and the resulting observations [6]. Statistical postprocessing is a method that transforms ensemble forecasts into calibrated and sharp predictive distributions [7].

Over the past decade, various statistical post-processing methods have been proposed in the meteorological literature [8], for example, Ensemble Model Output Statistics (EMOS), Bayesian Model Averaging (BMA), Quantile Regression, etc. BMA is a method capable of predicting the best model based on the weighted average of all models. The mean BMA unifies the posterior distribution of all the final trained models. The purpose of BMA is to combine uncertain models to get the best model [9]. EMOS used a single normal distribution, where the mean is an affine function of the ensemble members, and the variance is an affine function of the ensemble variable [10]. Quantile regression approaches focus on the range of quantiles within the forecast distribution and put few or no constraints on the overall shape [11]. However, these methods ignore spatial correlation patterns that may affect the forecasting results.

Geostatistical Output Perturbation is a model that captures the spatial dependency patterns identified in the errors [12]. GOP is a weather forecasting method obtained from only one deterministic output, just like NWP. Estimating the parameters for GOP is a popular geostatistical model that involves using maximum likelihood estimation (MLE) to estimate regression parameters with empirical semivariograms to determine the spatial correlation [13]. From these semivariograms, the spatial parameters of the GOP model were estimated using the weighted least squares method [14]. Then, the weather forecasts from the simple linear regression are added to the simulated error based on the spatial estimator, to obtain the calibrated forecasts [15]. Additionally, Feldmann (2012) discovered that GOP can calibrate temperature forecasts more accurately than non-spatial methods, even though GOP only uses simulation and modifies a single deterministic forecast [16]. As a result, countries that have insufficient resources to build adequate NWPs, such as Indonesia, can only generate ensemble-based weather forecasts.

This research discusses the GOP modelling method in predicting short-term weather, where the predicted weather elements are air temperature which is closely related to rainfall events, air pressure, wind, and humidity [17]. The selection of the GOP method is based on previous research which can calibrate the ensemble forecast and can consider the spatial correlation between errors. Additionally, GOP can be classified as multivariate modelling since it considers the spatial effects of all locations simultaneously, although it only uses one predictor at the modelling stage.

Section 2 of this research describes the variable used in this study, model evaluation, and methodology. The application of GOP for temperature forecast and the results of the analysis are explained in section 3. Finally, the conclusion of this research is presented in section 4.

2. Methods

The data are obtained from BMKG. The data consists of daily air temperature observation data (maximum and minimum temperature) for the period March, 1st 2021 to March, 22nd 2022 for 387 days. In formulating the GOP model, it is required to have daily NWP data for the Indonesia Numerical Weather Prediction (INA-NWP) model in the period corresponding to the observations. The observation stations that are the focus of this study are 10 meteorological stations in Surabaya and its surroundings as shown in Figure 1 and Table 1 as follows.

Table 1. Meteorological Stations of Interest Over Surabaya.

No	Meteorological Station	Latitude	Longitude
1	Juanda	-7.34869	112.7254
2	AWS Balong Bendo	-7.41332	112.5575
3	AWS Buduran	-7.42324	112.7209
4	AWS Sukolilo	-7.29018	112.7699
5	Perak I	-7.22694	112.7152
6	AWS Cerme	-7.18261	112.5780
7	AWS Karang Pilang	-7.32627	112.6165
8	AWS Sambikerep	-7.26305	112.6432
9	Perak II	-7.20524	112.7354
10	AWS Maritim II	-7.19885	112.7343

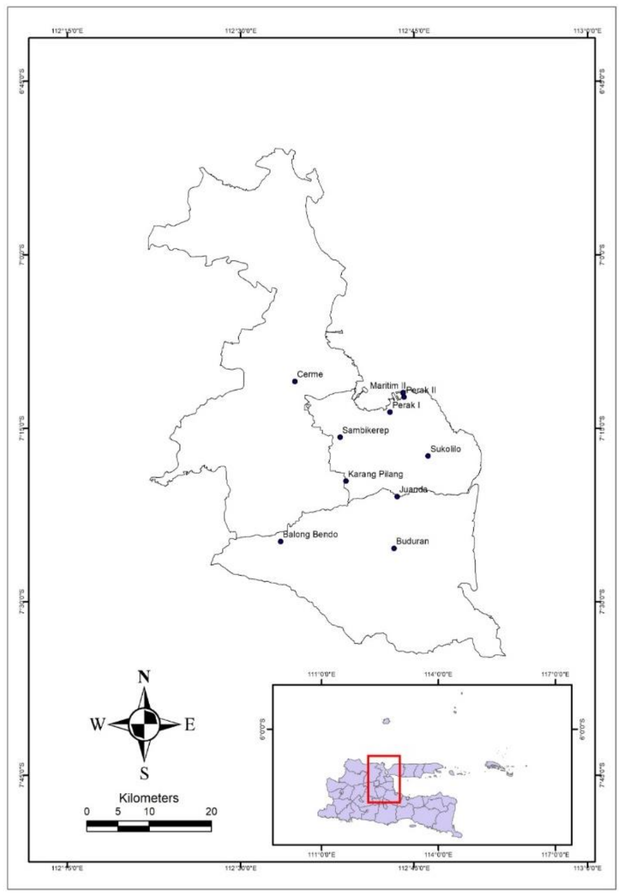


Figure 1. Meteorological Stations of Interest Over Surabaya

The predictand of the GOP is the observed temperature, which includes the maximum (Tmax) and minimum (Tmin) temperatures. Predictor variables are taken from the output of INA-NWP on the Weather Research and Forecasting Environmental Modelling System (WRF-EMS).

The steps to apply GOP for calibrated temperature forecast are as follows:

1. Perform data pre-processing. If there is a missing value in the observation data, then the input is done with the average value on 1 day before and after.
2. Modelling deterministic INA-NWP forecasts for daily temperature using the GOP method based on 30 days of training with considers 24-hour ahead forecast of NWP parameters. The stages is:
   i. Test the significance of spatial dependence with Moran's I
   ii. Regress the real observations against the NWP forecasts/outflows using Maximum Likelihood to obtain the regression bias coefficients, namely β_0 and β_1
   iii. Analyse the empirical semivariogram based on regression residuals. For this study, the empirical semivariogram selected is the exponential method.
   iv. Estimates spatial parameters on the variogram ρ^2, σ^2, and r based on the residuals that have been binned. The result of ρ^2, σ^2, and r are still not close form, so it is required to do numerical iteration using Limited-Memory BFGS (L-BFGS)
   v. Calculates the temperature forecast for a given day and the GOP model predictive interval for that day based on the 5th percentile and 95th percentile.
   vi. Evaluate and compare the accuracy of GOP forecasts with the accuracy of NWP parameters for temperature.
3. Evaluate the GOP forecast using:
   i. Root Mean Square Error (RMSE) use Eq. 1 [18]

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2} \qquad (1)$$

   ii. Continuous Rank Probability Score (CRPS) use Eq. 2 [19]–[21]

$$\text{crps}\left(F, y\right) = \int \left[F\left(x\right) - 1\left(x \geq y\right)\right]^2 \mathrm{d}x \qquad (2)$$

   iii. Percent Bias (PBIAS) use Eq. 3 [22]

$$\%\ Bias = \left[\frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)*100}{\sum_{i=1}^{n} y_i}\right] \qquad (3)$$

4. Calculate the calibrated temperature prediction value.
5. Draw Conclusion

3. Result and Discussion

Spatial Dependencies The significance of the spatial dependence of maximum and minimum temperatures among the ten meteorological stations was examined using Moran's I [23]. To be able to test for spatial dependence, a weighted distance matrix is required that only contains values of 0 and 1. Since all meteorological stations are not side by side, a cut-off distance is required as the maximum distance that still influences weather dynamics. In this research, it is assumed that the elevation (height) of meteorological stations is uniform, so a cut-off distance of 20 km is used.

Moran's I in this research are -0,252 and 0,404 for maximum temperature and minimum temperature, respectively. Moran's I in figure 2 state that there is a spatial dependence for maximum and minimum temperatures that are significant at the $\alpha = 0.05$ significance level. The A positive Moran's I value indicates that air temperatures in neighboring locations tend to have a stronger relationship than those in distant locations and vice versa. [24] [13].

Assessment Best Model. To determine the semivariogram model that is best for temperature forecasting in Surabaya and surrounding areas, three various models were used, namely Exponential, Gaussian, and Spherical**.** A comparison of the empirical semivariogram model is presented in Figure 3, then Table 2 reports the model evaluation and predictive checking by Verification Rank Histogram (VRH) that is shown in Figure 4.

By comparing and evaluating the sum square error of the three GOP models, Exponential model was identified as the best model for temperature forecasting. Figure 3 shows the development of the residual of the GOP model to generate the Spherical, Exponential, and Gaussian semivariogram. The semivariogram values in Figure 3 are constant at a distance r of 0.020 km or longer with a sill values (nugget + partial sill) are 0,43403683 and 0,05885673 for maximum and minimum temperature, respectively. A higher value of sill causes the estimated variance of forecasting increases. Forecasting variance affects the accuracy of temperature forecasts, accordingly affecting the interval of the GOP model that is then evaluated by CRPS.

Table 2. The Distance Matrix for 10 Meteorological Stations (km)

Stations	1	2	3	4	5	6	7	8	9	10
1	0,0	19,9	8,3	8,2	13,6	24,6	8,5	13,1	16,0	16,7
2	19,9	0,0	18,1	27,1	27,1	25,8	14,2	19,2	30,3	30,8
3	8,3	18,1	0,0	15,8	21,8	31,0	13,2	19,8	24,3	25,0
4	8,2	27,1	15,8	0,0	9,3	24,3	13,7	14,3	10,2	10,9
5	13,6	27,1	21,8	9,3	0,0	16,0	13,1	14,3	10,2	10,9
6	24,6	25,8	31,0	24,3	16,0	0,0	18,0	11,5	17,5	17,3
7	8,5	14,2	13,2	13,7	13,1	18,0	0,0	7,1	16,3	16,9
8	13,1	19,2	19,8	14,3	8,9	11,5	7,1	0,0	12,0	12,3
9	16,0	30,3	24,3	10,2	3,3	17,5	16,3	12,0	0,0	0,7
10	16,7	30,8	25,0	10,9	3,7	17,3	16,9	12,3	0,7	0,0

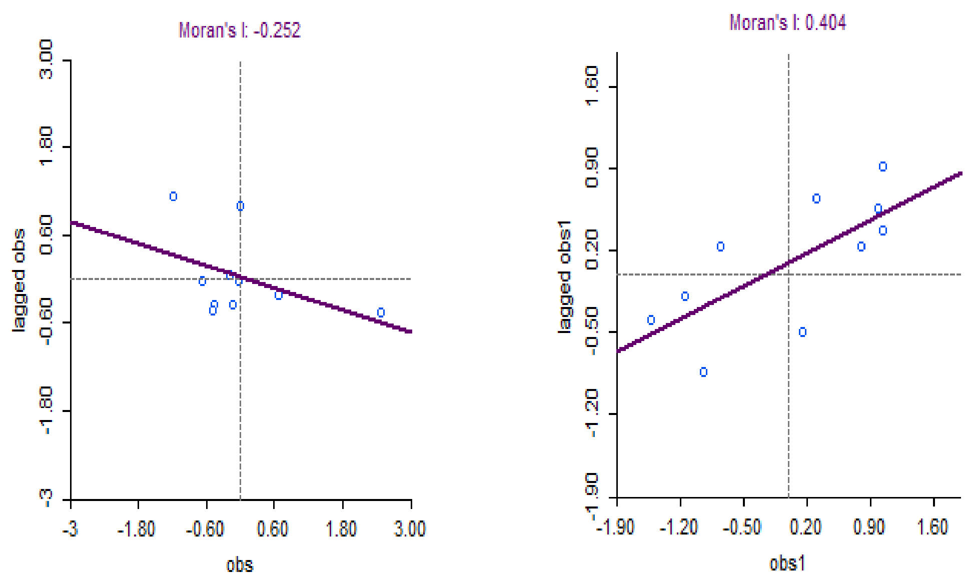
Note: 1. Juanda; 2. Balong Bendo; 3. Buduran; 4. Sukolilo; 5. Perak I; 6. Cerme; 7. Karang Pilang; 8. Sambikerep; 9. Perak II; 10. Maritim II

Additionally, Figure 3 shows spatial irregularity on both maximum and minimum temperatures. It can be seen that at certain distances, there are several bins or groups of meteorological stations whose semivariogram values are much higher than other groups. Despite the spatial inconsistencies, the simulation process run by the GOP model by modifying the residuals to obtain calibrated weather forecasts is expected to overcome this phenomenon.

In Figure 4, the horizontal temperature is the i-th rank where $i = 1,2,..., M$ or corresponds to many ensemble members plus one observation, while the vertical axis is the cumulative Binomial probability with parameters n is the total observations and p is $1/k$ where k is the frequency of observations that are at the i-th rank. For this case, a simulation was carried out to generate 99 realizations of ensemble members. From these 99 realizations, the lower and upper limits that are considered to represent all realizations are selected, namely the 5^{th} (P5) and 95^{th} (P95) percentiles.

Table 3. Comparison of Evaluating Empirical Variogram

	Sum of Squared Error		
	Spherical	Exponential	Gaussian
Tmax	5,71	4,06	5,30
Tmin	2,45	1,74	2,27



(a) (b)

Figure 2. Moran's I plot among 10 stations: (a) maximum temperature and (b) minimum temperature.

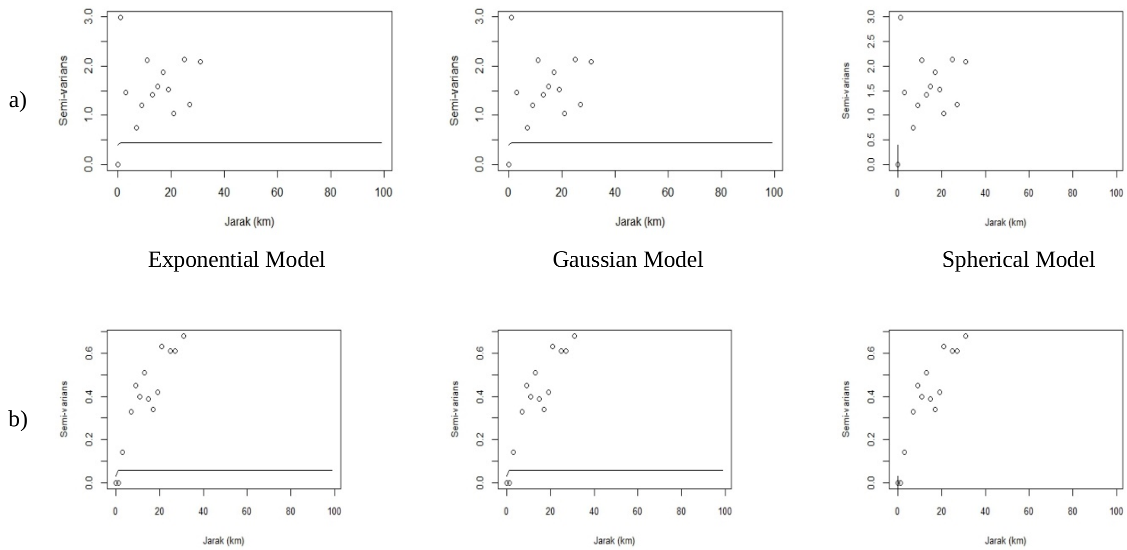


Figure 3. Empirical Semivariogram of Maximum Temperature (a) Minimum Temperature (b)

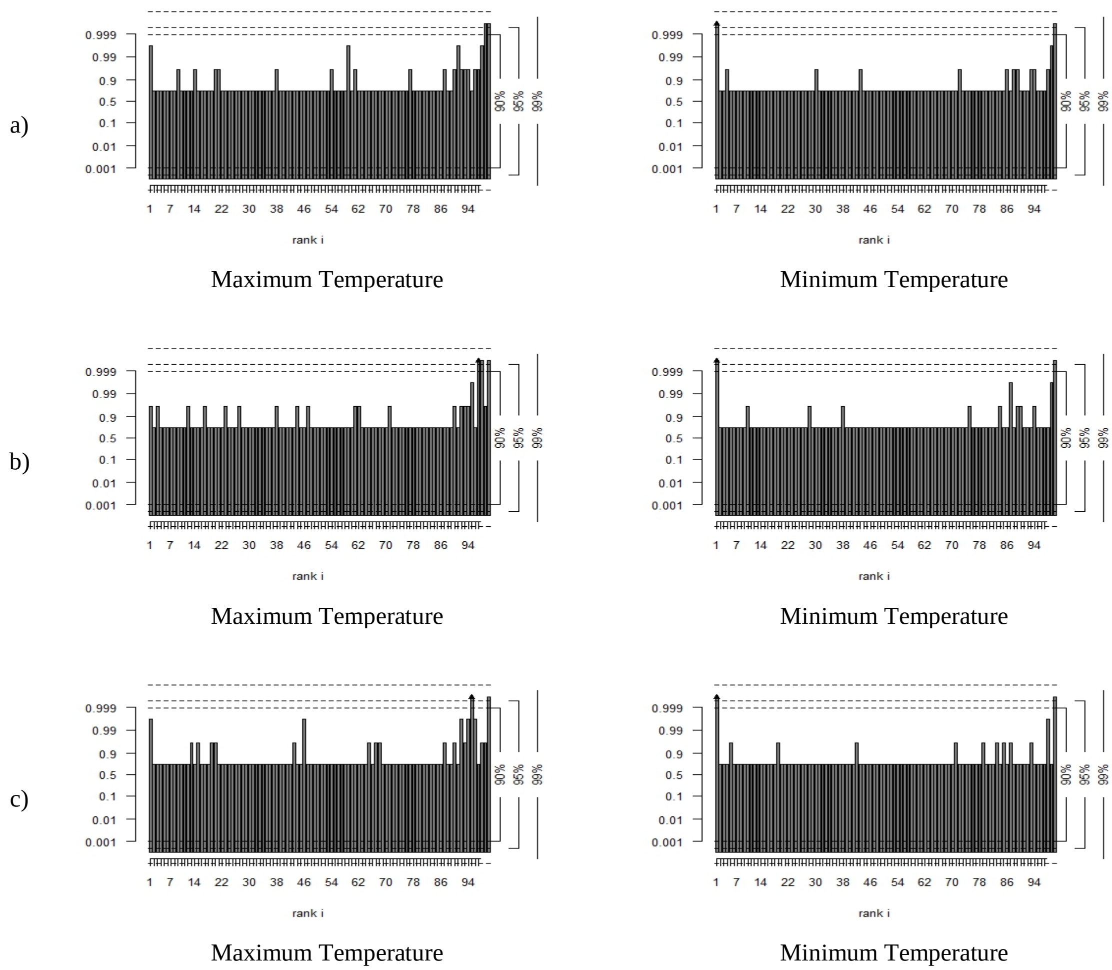


Figure 4. Verification Rank Histogram of GOP forecast, 31/3/2021 – 22/3/2022 (a) Exponential, (b) Gaussian, (c) Spherical.

Forecast Model. In the GOP model with a 30-day training period for both maximum and minimum temperature, the following equation is obtained as follows:

$$T_{MAX_{s,t}} = 28,006 + 0,136 \text{tmax}_{s,t}$$

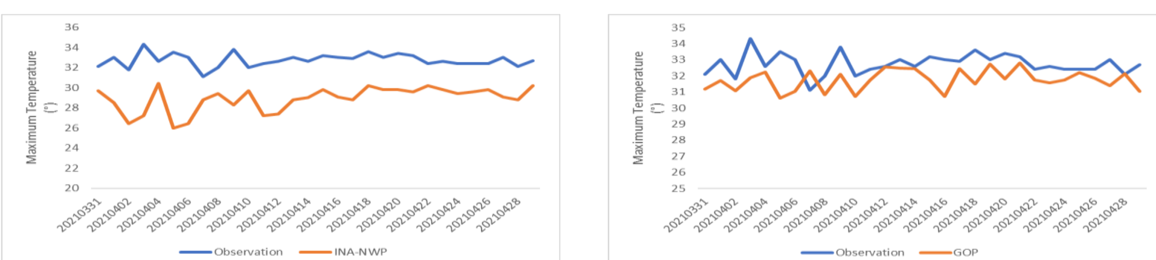
$$T_{MIN_{s,t}} = 28,751 - 0,168 \text{tmin}_{s,t}$$

The GOP model can improve the bias correction rate as presented in Table 4, which is indicated by the RMSE of GOP for maximum and minimum temperature being lower (i.e. the GOP model is better) than the RMSE of INA-NWP. The selection of March 31, 2021 is merely an illustration of the comparison between INA-NWP and GOP in temperature prediction. Compare the results before and after using the GOP method is represented by the maximum temperature comparison graph in Juanda in Figure 5.

Based on Figure 5, it can be seen that before using the GOP method, the NWP results are still biased towards the maximum observed temperature which is denoted by graphs that are far apart from each other. Whereas after using the GOP method of prediction results from NWP, the existing bias has been corrected from the maximum observed temperature which is denoted by the distance between adjacent graphs.

Table 4. RMSE of Maximum and Minimum Temperature Forecasts Using GOP and INA-NWP, 31st March 2021

Temperature	Station	Obs (°C)	INA-NWP (°C)	GOP (°C)	P_5 (°C)	P_{95} (°C)	RMSE of INA-NWP	RMSE of GOP
Maximum	Juanda	32,1	29,7	31,2	31,0	32,9	4,897	1,489
	Balong Bendo	33,9	29,7	31,8	30,8	33,0	4,388	1,502
	Buduran	33,4	29,7	33,4	30,8	32,9	4,545	1,811
	Sukolilo	35,1	29,7	31,8	30,9	33,2	4,426	1,532
	Perak I	31,4	26,2	30,6	30,6	32,5	5,019	1,446
	Cerme	32,4	26,2	31,5	30,6	32,5	4,947	1,681
	Karang Pilang	32,4	26,2	32,6	30,4	32,6	4,771	1,573
	Sambikerep	33,1	29,7	31,5	30,8	33,1	4,084	1,443
	Perak II	32,8	26,9	30,5	30,8	32,8	5,568	2,528
	Maritim II	32,8	26,9	32,0	30,5	32,5	3,089	1,658
Temperature	Station	Obs (°C)	INA-NWP (°C)	GOP (°C)	P_5 (°C)	P_{95} (°C)	RMSE of INA-NWP	RMSE of GOP
Minimum	Juanda	24,2	22,7	24,5	24,4	25,3	2,569	1,056
	Balong Bendo	24,2	22,7	25,2	24,6	25,3	2,462	1,042
	Buduran	24,9	22,7	24,8	24,6	25,4	2,607	1,043
	Sukolilo	25,7	22,7	25,0	24,6	25,3	2,673	1,062
	Perak I	25,3	23,8	24,8	24,4	25,1	2,825	1,148
	Cerme	25,5	23,8	24,9	24,3	25,1	2,896	1,121
	Karang Pilang	23,7	23,8	25,3	24,5	25,2	2,664	1,077
	Sambikerep	25,3	23,8	24,9	24,4	25,2	2,717	1,081
	Perak II	25,2	22,9	24,8	24,5	25,2	3,493	1,132
	Maritim II	25,2	22,9	24,7	24,6	25,3	3,492	1,106



(a) (b)

Figure 5. Comparison of Forecast Results (a) Before Using GOP and (b) After Using GOP

Table 5. Comparison of Evaluating GOP and INA NWP

Temperature	Coverage	CRPS
Maximum	51,76	0,963
Minimum	28,32	0,735

In addition to the RMSE value, there is a CRPS value that is also able to verify the forecast model in terms of the level of bias correction and sharpness achieved. Sharpness indicates that the predictive PDF of the calibrated model can improve the predictive PDF of the model before calibration, i.e., the PDF interval of the calibrated model is narrower than that before it was calibrated. The lower the CRPS value, which in this case is close to the value of 0, the better and more reliable the resulting model is in performing weather forecasts. In Table 5, it is known that the CRPS for maximum temperature is 0,963, which means that the model is not significantly good and less reliable in conducting weather forecasts, but at the minimum temperature has a CRPS value of 0,735 which means the model is quite good at making forecasts.

The forecast results are considered calibrated if the coverage value is close to the standard coverage value. Coverage is defined as the proportion of true observations between the lower and upper quartiles of the predicted distribution $\alpha/2$, which for a corrected probability forecast should be approximately $(1-\alpha)100\%$. In order to enable direct comparison with raw ensembles, level α is usually chosen to correspond to the nominal coverage of the ensemble range, which for a group of M members equals $(M - 1)/(M + 1)100\%$ [25]. For this study, the number of ensembles used is 99, following the number of ensemble realization simulations on the GOP model, so the standard coverage used is 98%.

Table 6. Percent Bias in 10 Meteorological Stations

Stations	Maximum		Minimum	
	PBIAS INA-NWP	PBIAS GOP	PBIAS INA-NWP	PBIAS GOP
Juanda	7,48	2,85	6,20	-1,45
Balong Bendo	12,39	6,24	6,08	-4,37
Buduran	11,05	0,06	8,72	0,36
Sukolilo	15,46	9,51	11,84	3,06
Perak I	16,56	2,62	5,93	2,05
Cerme	19,23	2,96	6,81	2,37
Karang Pilang	19,16	-0,53	-0,55	-7,05
Sambikerep	10,19	4,62	6,08	1,73
Perak II	17,99	7,01	9,13	1,37
Maritim II	17,99	2,55	9,13	2,00

PBIAS calculate the average trend of the INA-NWP and GOP model relative to their observed counterparts. A negative value indicates overestimated (i.e., the GOP model and INA-NWP is higher than the observed data), meanwhile, a positive value indicates underestimated (the GOP model and INA-NWP are lower than the observed data). The optimal value of PBIAS is 0.0, with low-magnitude values in both directions, possibly indicating accurate model simulations. Based on Table 6, generally, the INA NWP percent bias has a larger amplitude than the GOP percent bias, which affects the RMSE calculation.

4. Conclusion

This study proposes a GOP model to analyze the spatial relationship in ten meteorological stations to achieve higher accuracy and precision in weather forecasting. For a training period of 30 days, the GOP model performs well and improved, with an RMSE value of GOP lower than INA-NWP's. The best model for maximum and minimum temperature is selected based on the smallest RMSE value.

Based on the standard coverage criteria, the GOP model has not been able to improve the sharpness of ensemble forecasts. GOP forecast coverage is still far from the specified 98% coverage standard. The selected empirical semivariogram is the exponential model based on the smallest sum squared error criterion. In practice, we estimate the empirical variogram by a theoretical model chose from the group of valid variograms. Statistical postprocessing is still required to improve the accuracy of INA-NWP and dynamic model development for weather prediction in tropical climate regions. As a future direction, this model can be attempted on other INA NWP output variables, such as precipitation, humidity, wind speed, etc. GOP should be carried out by involving a much larger number of meteorological stations, at least dozens of stations in a fairly close location. This is to minimize the potential for inaccurate forecasts due to one or more stations that are inaccurate forecasts due to the existence of one or more stations that are geographically more distant than a set of other stations. In addition, validation and various lengths of training windows can also be applied so that the model used is not constant / static and always adaptive to the characteristics of the weather.

Acknowledgment

The author gratefully acknowledges BMKG for data assistance for this study

References

[1] Z. Pu and E. Kalnay, “Numerical Weather Prediction Basics: Models, Numerical Methods, and Data Assimilation,” in *Handbook of Hydrometeorological Ensemble Forecasting*, 1st ed., Q. Duan, F. Pappenberger, A. Wood, H. Cloke, and J. Schaake, Eds. Heidelberg: Springer Berlin Heidelberg, 2019, pp. 67–97.

[2] H. Thifari, “Seminar Ilmiah dan General Discussion ‘Membangun Kemandirian Permodelan Cuaca Numerik Nasional (Ina-NWP),’” *Badan Meteorologi, Klimatologi dan Geofisika*, 2019. https://www.bmkg.go.id/berita/?p=seminar-ilmiah-dan-general-discussion-membangun-kemandirian-permodelan-cuaca-numerik-nasional-ina-nwp&lang=ID&tag=berita-utama (accessed Jan. 25, 2024).

[3] F. D. Cahyoko, “Spatial Ensemble Model Output Statistics Untuk Kalibrasi Dan Prakiraan Cuaca Jangka Pendek,” repository.its.ac.id, 2023.

[4] J. Dong, W. Zeng, L. Wu, J. Huang, T. Gaiser, and A. K. Srivastava, “Enhancing short-term forecasting of daily precipitation using numerical weather prediction bias correcting with XGBoost in different regions of China,” *Eng. Appl. Artif. Intell.*, vol. 117, no. November 2022, p. 105579, 2023, doi: 10.1016/j.engappai.2022.105579.

[5] J. A. Brotzge *et al.*, “Challenges and Opportunities in Numerical Weather Prediction,” *Bulletin of the American Meteorological Society*. journals.ametsoc.org, 2023, [Online]. Available: https://journals.ametsoc.org/view/journals/bams/104/3/BAMS-D-22-0172.1.xml.

[6] S. Allen, C. A. T. Ferro, and F. Kwasniok, “Recalibrating wind-speed forecasts using regime-dependent ensemble model output statistics,” *Q. J. R. Meteorol. Soc.*, vol. 146, no. 731, pp. 2576–2596, 2020, doi: 10.1002/qj.3806.

[7] V. A. Siqueira, A. Weerts, B. Klein, F. M. Fan, R. C. D. de Paiva, and W. Collischonn, “Postprocessing continental-scale, medium-range ensemble streamflow forecasts in South America using Ensemble Model Output Statistics and Ensemble Copula Coupling,” *J. Hydrol.*, vol. 600, no. 126520, pp. 1–20, 2021, doi: 10.1016/j.jhydrol.2021.126520.

[8] S. Vannitsem *et al.*, “Statistical Postprocessing for Weather Forecasts,” *Bull. Am. Meteorol. Soc.*, pp. E681–E699, 2021.

[9] A. M. Basher, A. K. M. S. Islam, M. A. Stiller-Reeve, and P. Chu, “Changes in future rainfall extremes over Northeast Bangladesh: A Bayesian model averaging approach,” *Int. J. Climatol.*, vol. 40, pp. 3232–3249, 2019, [Online]. Available: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/joc.6394.

[10] H. W. Choi, Y. H. Kim, K. Han, and C. Kim, “Probabilistic Forecast of Low Level Wind Shear of Gimpo, Gimhae, Incheon and Jeju International Airports Using Ensemble Model Output Statistics,” *Atmosphere (Basel).*, vol. 12, 2021, [Online]. Available: https://www.mdpi.com/1395904.

[11] J. B. Bremnes, “Ensemble postprocessing using quantile function regression based on neural networks and bernstein polynomials,” *Mon. Weather Rev.*, vol. 148, no. 1, pp. 403–414, 2020, doi: 10.1175/MWR-D-19-0227.1.

[12] Y. Gel, A. E. Raftery, and T. Gneiting, “Calibrated probabilistic mesoscale weather field forecasting: The geostatistical output perturbation method,” *J. Am. Stat. Assoc.*, vol. 99, no. 467, pp. 575–583, 2004, doi: 10.1198/016214504000000872.

[13] Sutikno, Purhadi, I. Mukhlash, K. N. Anisa, U. Haryoko, and H. Harsa, “Calibrating weather forecasting in Indonesia: The geostatistical output perturbation method,” *Malaysian J. Sci.*, vol. 38, no. 2, pp. 100–112, 2019, doi: 10.22452/mjs.sp2019no2.9.

[14] N. Cressie, “Fitting variogram models by weighted least squares,” *J. Int. Assoc. Math. Geol.*, vol. 17, no. 5, pp. 563–586, 1985, doi: 10.1007/BF01032109.

[15] V. J. Berrocal, A. E. Raftery, and T. Gneiting, “Combining spatial statistical and ensemble information in probabilistic weather forecasts,” *Mon. Weather Rev.*, vol. 135, no. 4, pp. 1386–1402, 2007, doi: 10.1175/MWR3341.1.

[16] K. Feldmann and T. Thorarinsdottir, “Statistical postprocessing of ensemble forecasts for temperature: The importance of spatial modeling,” *Ruperto-Carola University of Heidelberg, Germany*. 2012.

[17] C. H. Huang, H. H. Tsai, and H. C. Chen, “Influence of weather factors on thermal comfort in subtropical urban environments,” *Sustain.*, vol. 12, no. 5, 2020, doi: 10.3390/su12052001.

[18] T. Gneiting, “Making and Evaluating Point Forecasts,” *J. Am. Stat. Assoc.*, vol. 106, no. 494, pp. 746–762, 2011, doi: 10.1198/jasa.2011.r10138.

[19] A. Jordan, F. Krüger, and S. Lerch, “Evaluating probabilistic forecasts with scoringRules,” *J. Stat. Softw.*, vol. 90, no. 12, pp. 1–37, 2019, doi: 10.18637/jss.v090.i12.

[20] V. V’yugin and V. Trunov, “Online Learning with Continuous Ranked Probability Score,” in *Proceedings of Machine Learning Research 105*, 2019, pp. 1–15, [Online]. Available: http://arxiv.org/abs/1902.10173.

[21] M. Zamo and P. Naveau, “Estimation of the Continuous Ranked Probability Score with Limited Information and Applications to Ensemble Weather Forecasts,” *Math. Geosci.*, vol. 50, no. 2, pp. 209–234, 2018, doi: 10.1007/s11004-017-9709-7.

[22] B. Berhanu, Y. Seleshi, S. S. Demisse, and A. M. Melesse, “Bias correction and characterization of climate forecast system re-analysis daily precipitation in Ethiopia using fuzzy overlay,” *Meteorol. Appl.*, vol. 23, no. 2, pp. 230–243, 2016, doi: 10.1002/met.1549.

[23] O. Raza, M. A. Mansournia, A. R. Foroushani, and K. Holakouie-Naieni, “Exploring spatial dependencies in the prevalence of childhood diarrhea in Mozambique using global and local measures of spatial autocorrelation,” *Med. J. Islam. Repub. Iran*, vol. 34, no. 1, 2020, doi: 10.34171/mjiri.34.59.

[24] C. Feng *et al.*, “Spatial and temporal analysis of liver cancer mortality in Yunnan province, China, 2015–2019,” *Front. Public Heal.*, vol. 10, no. 6, pp. 1–10, 2022, doi: 10.3389/fpubh.2022.1010752.

[25] B. Schulz, M. El Ayari, S. Lerch, and S. Baran, “Post-processing numerical weather prediction ensembles for probabilistic solar irradiance forecasting,” *Solar Energy*. Elsevier, 2021, [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0038092X21002097.