

Particle Swarm Optimized–Support Vector Regression Hybrid Model for Daily Horizon Electricity Demand Forecasting Using Climate Dataset

Mohanad S AL-Musaylh^{1,2}, Ravinesh C Deo¹, and Yan Li¹

¹School of Agricultural, Computational and Environmental Sciences, University of Southern Queensland, QLD 4350, Australia

²Management Technical College, Southern Technical University, Basrah, Iraq

Abstract. This paper has adopted six daily climate variables for the eleven major locations, and heavily populated areas in Queensland, Australia obtained from Scientific Information for Land Owners (SILO) to forecast the daily electricity demand (G) obtained from the Australian Energy Market Operator (AEMO). Optimal data-driven technique based on a support vector regression (SVR) model was applied in this study for the G forecasting, where the model’s parameters were selected using a particle swarm optimization (PSO) algorithm. The performance of PSO–SVR was compared with multivariate adaptive regression spline (MARS) and the traditional model of SVR. The results showed that the PSO–SVR model outperformed MARS and SVR.

1 Introduction

Electricity demand (G) forecasting is a purely fundamental yet a challenging optimisation task for improving business efficiency of the electricity industry. A relationship between the G data and temperature is clearly evident in winter and summer [1]. Hence, it would be significant to develop a forecasting model employing both the G and related climate input datasets.

In recent years, support vector regression (SVR), PSO algorithm, and multivariate adaptive regression splines (MARS) have been widely adopted in energy demand forecasting [1]. Those methods have been used to forecast G in [1, 2], however, the influences of the climate datasets are not incorporated yet.

The main contribution of this research paper is to improve the G forecasting accuracy by involving climate datasets and integrating the merits of the PSO algorithm with the SVR model. To evaluate the PSO-SVR model, the traditional methods of the SVR and MARS algorithms are also developed.

2 Theoretical Background

2.1 Support vector regression

A nonlinear regression problem can be solved by a SVR model, which is a machine learning method and pioneered by [3], below:

$$y = f(X) = \omega \cdot \phi(X) + b \tag{1}$$

where $X = \{x_i\}_{i=1}^{i=n} \in \mathcal{R}^N, y = \{y_i\}_{i=1}^{i=N} \in \mathcal{R}$ are the predictors and target variables, respectively. b is a constant, ω is the weighted vector, and $\phi(X)$ represents the mapping function employed in the feature space. A minimisation technique is used to estimate the coefficients ω and b as follows [3]:

$$\text{Minimize } \frac{1}{2} \|w\|^2 + C \frac{1}{N} \sum_{i=1}^N (\xi_i + \xi_i^*) \tag{2}$$

$$\text{Subject to } \begin{cases} |y_i - (\langle w, x_i \rangle + b) \geq \varepsilon + \xi_i \\ \langle w, x_i \rangle + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases} \tag{3}$$

where the smoothness of the function is determined by $\frac{1}{2} \|w\|^2$, C and ε are the model’s parameters and the nonnegative slack variables (ξ and ξ^*) demonstrate the distance between actual and equivalent boundary values of a function approximation. A nonlinear regression function can be expressed by Eq. 4 after applying Lagrangian multipliers and optimising conditions [3]:

$$f(X) = \sum_{i=1}^{i=N} (\alpha_i - \alpha_i^*) K(x_i, x_j) + b \tag{4}$$

where x_i and $x_j \in X$, and the term $K(x_i, x_j)$ denotes the kernel function. α_i and α_i^* are Lagrangian multipliers [3].

In this study, the radial basis function (RBF) was used in the processing of the SVR model as follows [4]:

$$K(x_i, x_j) = \exp\left(\frac{-\|x_i - x_j\|^2}{2\sigma^2}\right) \quad (5)$$

where the kernel width and inputs are represented by σ and x_i, x_j , respectively. The critical task for developing the SVR model with a good accuracy is to determine the three parameters which are kernel width (σ), the loss function (ϵ) and regulation (C) during the training period [5]. This is achieved through a hybrid method called particle swarm optimization (PSO) in section 2.3 below.

2.2 Multivariate adaptive regression splines

The relationship between X and y is demonstrated by the MARS model as follows [6]:

$$y = \hat{f}(X) = a_0 + \sum_{m=1}^M a_m BF_m(X) \quad (6)$$

where X and y are offered in Eq. 1, a_0 is a constant, $\{a_m\}_1^M$ are the model coefficients, M is the number of basis functions in MARS, and $BF_m(X)$ is a spline function defined as $C(X|s, t_1, t, t_2)$, where $t_1 < t < t_2$, and $s = \pm 1$.

The number of the basis functions for developing MARS model is determined through the Generalized Cross-Validation criterion (GCV) based on the mean square error (MSE) [6]:

$$GCV = MSE / \left[1 - \frac{\tilde{\sigma}(M)}{N}\right]^2 \quad (7)$$

where $MSE = \frac{1}{N} \sum_{i=1}^N [y_i - \hat{f}(X_i)]^2$. $\tilde{G}(M) = C(M) + v \cdot M$, where v is a penalty factor with a characteristic value of $v = 3$. $C(M)$ is the number of parameters being fitted. In the training dataset, the lowest value of the GCV refers to the optimal MARS model.

2.3 Particle swarm optimization (PSO)

To select the best parameters of the SVR model, which are the regulation function (C), kernel width (σ) and loss function (ϵ), the PSO algorithm, first introduced by Kennedy and Eberhart [7, 8], is employed in this paper using the mean square error (MSE) as the fitness function as follows [9-11]:

$$\begin{aligned} V_{i,j}(k+1) = & \omega * V_{i,j}(k) + c_1 * \text{rand}() \\ & * (P_{\text{best}_{i,j}}(k) - X_{i,j}(k)) + c_2 \\ & * \text{rand}() \\ & * (g_{\text{best}_j}(k) - X_{i,j}(k)) \end{aligned} \quad (8)$$

$$X_{i,j}(k+1) = X_{i,j}(k) + V_{i,j}(k+1) \quad (9)$$

where $X_i = (X_{i1}, X_{i2}, \dots, X_{iD})^T$ is the i^{th} particle from the initial population of the size of $i = 1, 2, \dots, N$ and a dimension of $j = 1, 2, \dots, D$. $V_i = (V_{i1}, V_{i2}, \dots, V_{iD})^T$ is the

velocity of each particle X_i in the population. According to [9, 11], $\text{rand}()$ represents a random number between zero and one while the individual and global extreme values are represented by $P_{\text{best}_{i,j}}$ and g_{best_j} , respectively. The two values of c_1 and c_2 are usually within [2, 2.05], whereas ω can be defined as follows [9, 10]:

$$\omega = \omega_{\min} + \frac{(T_{\max} - T) * (\omega_{\max} - \omega_{\min})}{T_{\max}} \quad (10)$$

where ω_{\min} and ω_{\max} usually equal to 0.4 and 0.9; T and T_{\max} are the current and maximum iteration numbers, respectively [9].

3 Materials and Methods

3.1 Electricity demand data (G)

In this study, the G data were recorded half-hourly (48 times per day) in Megawatts (MW) for the state of Queensland, and these data were acquired from the Australian Energy Market Operator (AEMO) [12] for the period of 01-01-2015 to 31-12-2016 (dd-mm-yyyy). The 30-minute data periods were converted to daily terms by obtaining total values for each day. Figure 1 showed the plots of the actual G data series.

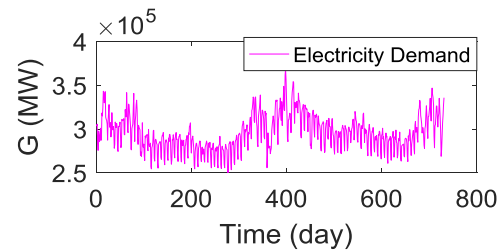


Fig. 1. Time-series of electricity demand (G MW) data.

3.2 Climate dataset

Historical climate datasets for the same period of the G data were obtained from Scientific Information for Land Owners (SILO) [13]. The data were collected for the main eleven stations, which contain the majority of the population of Queensland, that were shown in Fig. 2 and Table 1.

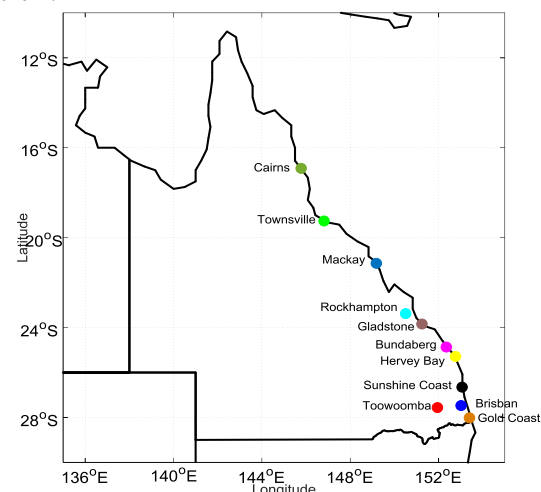


Fig. 2. Area map for the climate datasets.

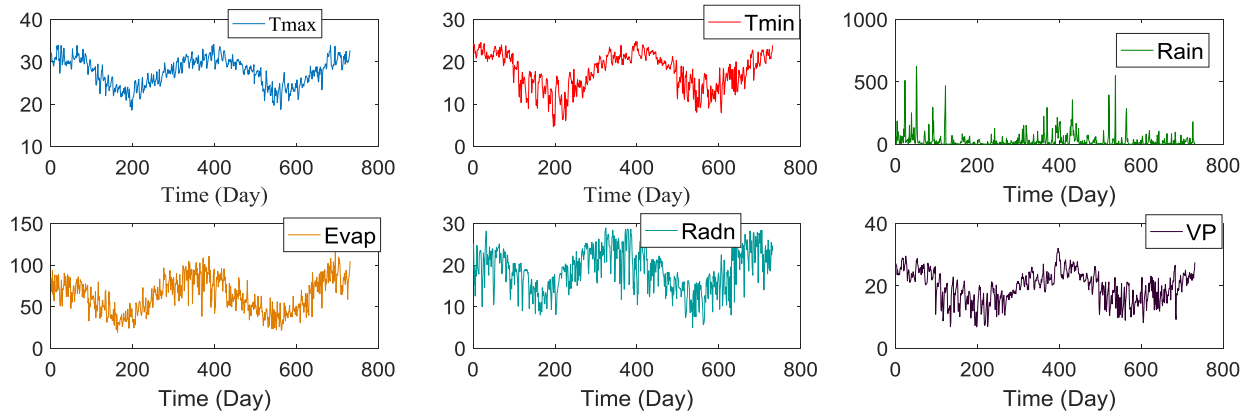


Fig. 3. Time-series of climate datasets used in this study.

Table 1. Population for each climate station

No.	Station	Population number
1	Cairns	240,190
2	Townsville	229,031
3	Mackay	43,364
4	Rockhampton	79,726
5	Gladstone	61,640
6	Bundaberg	92,897
7	Hervey Bay	53,035
8	Sunshine Coast	346,522
9	Brisbane	2,270,800
10	Gold Coast	555,721
11	Toowoomba	149,512
	Total	4,122,438

The population numbers were obtained from Australian Bureau of Statistics [14] where the total number of population resulted from the eleven stations in Table 1 is very close to the population of whole Queensland (4,883,739).

The input data were comprised of the time-series of maximum and minimum air temperature (T_{max} and T_{min}), rainfall (Rain) evaporation (Evap), solar radiation (Radn) and vapour pressure (VP). The datasets for whole Queensland were obtained by getting the average of T_{max} , T_{min} , Radn and VP and the total values of Rain and Evap of the eleven stations datasets. Those were used as the inputs of the models. Figure 3 showed the plots of those actual time series.

3.3 Forecast model development and validation

The climate variables in section 3.2 above were used to forecast the G data by developing three models: PSO-SVR, SVR and MARS. As there is no a single method for splitting data into training, validation and testing [5], the data were divided into subsets of 70% for training, 15% for validation and 15% for testing.

MATLAB-based Libsvm toolbox (version 3.22), developed by Chang and Lin [15], was used to build the SVR model. To develop a hybrid SVR model, the PSO algorithm (section 2.3) was used to select the optimal parameters based on the smallest value of MSE . To

evaluate the accuracy of the SVR model, the software packages version 1.13.0 was employed [16] for the MARS model.

The models were validated in Table 2 using the root-mean square error ($RMSE$, Eq. 11). The PSO-SVR model yielded the lowest $RMSE$, which indicated the best accuracy compared to the other models.

Table 2. Root-mean square error ($RMSE$; MW) in the validation dataset

Model	$RMSE$ (MW)
PSO-SVR	12730.52
SVR	12930.53
MARS	16701.61

3.4 Model performance evaluation

This study adopted a wide range of statistical error criteria in the testing period based on statistical indicators. Those are mean absolute error (MAE), root-mean square error ($RMSE$), relative error (%) based on MAE and $RMSE$ values ($MAPE$ and $RRMSE$), Willmott's Index (WI), the Nash-Sutcliffe coefficient (E_{NS}), and the Legates and McCabe Index (E_{LM}) [17-26], represented below:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{i=n} (G_i^{for} - G_i^{obs})^2} \quad (11)$$

$$MAE = \frac{1}{n} \sum_{i=1}^{i=n} |G_i^{for} - G_i^{obs}| \quad (12)$$

$$RRMSE = 100 \times \frac{\sqrt{\frac{1}{n} \sum_{i=1}^{i=n} (G_i^{for} - G_i^{obs})^2}}{G^{obs}} \quad (13)$$

$$MAPE = 100 \times \frac{1}{n} \sum_{i=1}^{i=n} \left| \frac{G_i^{for} - G_i^{obs}}{G_i^{obs}} \right| \quad (14)$$

$$WI = 1 - \left[\frac{\sum_{i=1}^{i=n} (G_i^{for} - G_i^{obs})^2}{\sum_{i=1}^{i=n} (|G_i^{for} - G^{obs}| + |G_i^{obs} - G^{obs}|)^2} \right], \quad 0 \leq \quad (15)$$

$$WI \leq 1$$

$$E_{NS} = 1 - \left[\frac{\sum_{i=1}^{i=n} (G_i^{for} - G_i^{obs})^2}{\sum_{i=1}^{i=n} (G_i^{obs} - G^{obs})^2} \right], \quad -\infty \leq E_{NS} \leq 1 \quad (16)$$

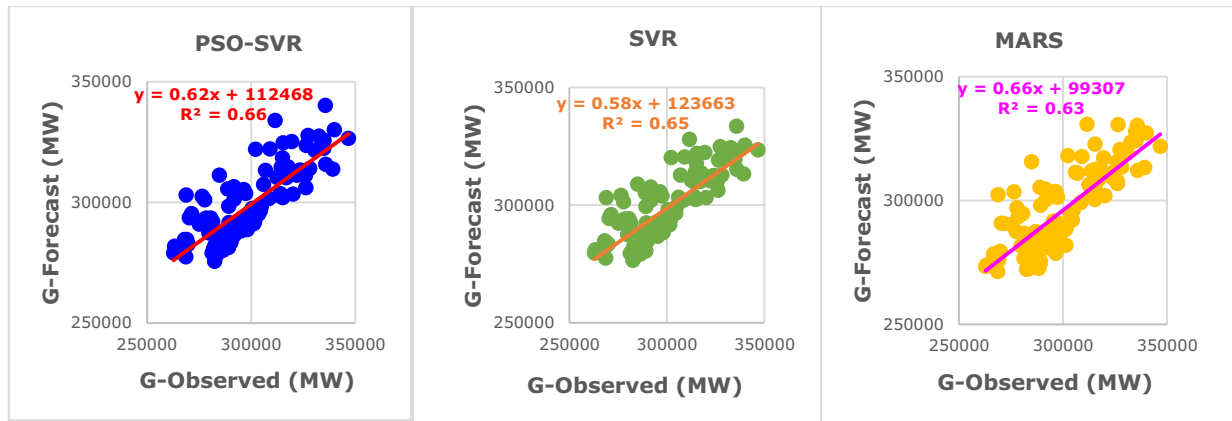


Fig. 4. Scatterplot of the G -forecasted vs. G -observed of electricity demand data in the testing period using the three models. The equation of linear regression line and the coefficient of determination are incorporated.

$$E_{LM} = 1 - \left[\frac{\sum_{i=1}^{i=n} |G_i^{obs} - G_i^{for}|}{\sum_{i=1}^{i=n} |G_i^{obs} - \bar{G}^{obs}|} \right], -\infty \leq E_{LM} \leq 1 \quad (17)$$

where G_i^{for} and G_i^{obs} are the i^{th} forecasted and observed values of G in the testing period, respectively; n is the total number of G_i^{for} or G_i^{obs} values, \bar{G}^{for} and \bar{G}^{obs} are the means of forecasted and observed values, respectively.

4 Results and Discussions

The performance of the PSO–SVR model for the daily forecast horizon was compared with traditional SVR and MARS models in the testing period. The results of the comparison indicated that the PSO–SVR yielded better performances (lowest $RMSE$, and MAE , as well as the largest WI , E_{NS} , and E_{LM}) than SVR and MARS models. Those values were summarized in Table 3.

Table 3. The performance evaluation of the models in the test period

Model	WI	E_{NS}	$RMSE$ (MW)	MAE (MW)	$MAPE$ (%)	$RRMSE$ (%)	E_{LM}
PSO–SVR	0.88	0.66	11716.97	9668.15	3.27	3.92	0.41
SVR	0.87	0.64	11909.00	9895.09	3.34	3.99	0.40
MARS	0.87	0.60	12612.22	11003.20	3.69	4.22	0.33

The scatterplots of G_i^{for} vs. G_i^{obs} and the model forecasted errors, $|FE| = |G_{FOR,i} - G_{OBS,i}|$ in the testing period for the three models were shown in Figs. 4 and 5, respectively. The lowest forecasted errors ($|FE|$) were shown by the PSO-SVR model in this study (Fig. 5). On the other hand, the highest correlation of determination (R^2) was achieved by the PSO-SVR model (Fig. 4). Overall, a significantly greater accuracy was attained by the PSO–SVR model than the other models.

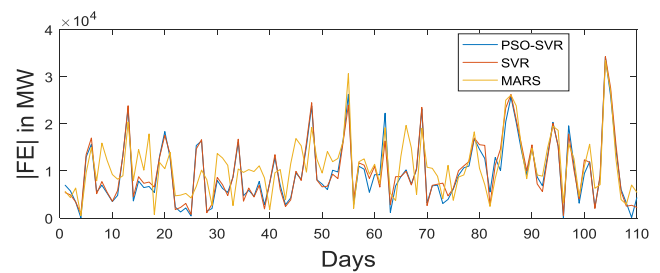


Fig. 5. Model forecasted errors, $|FE| = |G_{FOR,i} - G_{OBS,i}|$ in the testing period using the three models.

5 Concluding Remark

In this paper, a hybrid PSO–SVR model was proposed for daily G forecasting horizon in Queensland, Australia, where the model used data from Australian Energy Market Operator (AEMO) and Scientific Information for Land Owners (SILO). The MARS and the traditional method of SVR were also used in this research study to evaluate the main model. The results showed that the PSO–SVR outperformed the MARS and SVR models. As a result, the data-driven tool constructed by the PSO–SVR model is a powerful forecasting framework which can support the National Electricity Market (e.g., AEMO). Although the PSO–SVR model performed well in this paper, some challenges in model development section could be appeared. As the PSO algorithm needs a longer time to produce the SVR parameters, alternative methods, such as multi-swarm PSO and sine cosine algorithm may need to be used. In addition, the model could be improved using ensemble-based uncertainty testing by a bootstrapping technique. Those should be addressed in future studies.

Acknowledgments

Firstly, we would like to offer our thanks to the Australian Energy Market Operator (AEMO), Australian Bureau of Statistics, and Scientific Information for Land Owners (SILO) which provided all the required data. We are very grateful to the Ministry of Higher Education and Scientific Research in the Government of Iraq for funding the first author’s PhD project.

References

1. M. S. Al-Musaylh, R. C. Deo, J. F. Adamowski, and Y. Li, "Short-term electricity demand forecasting with MARS, SVR and ARIMA models using aggregated demand data in Queensland, Australia," *Advanced Engineering Informatics*, vol. 35, pp. 1-16, 2018.
2. M. S. Al-Musaylh, R. C. Deo, Y. Li, and J. F. Adamowski, "Two-phase particle swarm optimized-support vector regression hybrid model integrated with improved empirical mode decomposition with adaptive noise for multiple-horizon electricity demand forecasting," *Applied Energy*, vol. 217, pp. 422-439, 5/1/ 2018.
3. V. N. Vapnik, *Statistical learning theory* (no. Book, Whole). New York: Wiley, 1998.
4. J. A. K. Suykens, J. De Brabanter, L. Lukas, and J. Vandewalle, "Weighted least squares support vector machines: robustness and sparse approximation," *Neurocomputing*, vol. 48, no. 1-4, pp. 85-105, 10// 2002.
5. R. C. Deo, X. Wen, and F. Qi, "A wavelet-coupled support vector machine model for forecasting global incident solar radiation using limited meteorological dataset," *Applied Energy*, vol. 168, pp. 568-593, 2016.
6. J. H. Friedman, "Multivariate adaptive regression splines," *The annals of statistics*, pp. 1-67, 1991.
7. J. Kenny, "Particle swarm optimization," in *Proc. 1995 IEEE Int. Conf. Neural Networks*, 1995, pp. 1942-1948.
8. R. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," in *Micro Machine and Human Science, 1995. MHS'95., Proceedings of the Sixth International Symposium on*, 1995, pp. 39-43: IEEE.
9. X. Li and C. Li, "Improved CEEMDAN and PSO-SVR Modeling for Near-Infrared Noninvasive Glucose Detection," *Computational and mathematical methods in medicine*, vol. 2016, 2016.
10. M. N. Alam, "Particle Swarm Optimization: Algorithm and its Codes in MATLAB," ed: ResearchGate, 2016.
11. M. Hasanipanah, A. Shahnazar, H. B. Amnieh, and D. J. Armaghani, "Prediction of air-overpressure caused by mine blasting using a new hybrid PSO-SVR model," *Engineering with Computers*, vol. 33, no. 1, pp. 23-31, 2017.
12. AEMO. (2016). *Aggregated Price and Demand Data - Historical*. Available: <http://www.aemo.com.au/>
13. S. J. Jeffrey, J. O. Carter, K. B. Moodie, and A. R. Beswick, "Using spatial interpolation to construct a comprehensive archive of Australian climate data," *Environmental Modelling & Software*, vol. 16, no. 4, pp. 309-330, 2001.
14. A. B. o. Statistics. (2018). *Census*. Available: <http://www.abs.gov.au/>
15. C.-C. Chang and C.-J. Lin, "LIBSVM: a library for support vector machines," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 2, no. 3, p. 27, 2011.
16. G. Jekabsons, "Adaptive Regression Splines toolbox for Matlab/Octave," *Version*, vol. 1, p. 72, 2013.
17. K. Mohammadi, S. Shamsirband, M. H. Anisi, K. A. Alam, and D. Petković, "Support vector regression based prediction of global solar radiation on a horizontal surface," *Energy Conversion and Management*, vol. 91, pp. 433-441, 2// 2015.
18. C. J. Willmott, S. M. Robeson, and K. Matsuura, "A refined index of model performance," *International Journal of Climatology*, vol. 32, no. 13, pp. 2088-2094, 2012.
19. C. J. Willmott, "On the evaluation of model performance in physical geography," in *Spatial statistics and models*: Springer, 1984, pp. 443-460.
20. C. J. Willmott, "Some comments on the evaluation of model performance," *Bulletin of the American Meteorological Society*, vol. 63, no. 11, pp. 1309-1313, 1982.
21. C. J. Willmott, "On the validation of models," *Physical geography*, vol. 2, no. 2, pp. 184-194, 1981.
22. C. W. Dawson, R. J. Abraham, and L. M. See, "HydroTest: a web-based toolbox of evaluation metrics for the standardised assessment of hydrological forecasts," *Environmental Modelling & Software*, vol. 22, no. 7, pp. 1034-1052, 2007.
23. D. R. Legates and G. J. McCabe, "Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation," *Water resources research*, vol. 35, no. 1, pp. 233-241, 1999.
24. P. Krause, D. Boyle, and F. Bäse, "Comparison of different efficiency criteria for hydrological model assessment," *Advances in Geosciences*, vol. 5, pp. 89-97, 2005.
25. T. Chai and R. R. Draxler, "Root mean square error (RMSE) or mean absolute error (MAE)?—Arguments against avoiding RMSE in the literature," *Geoscientific Model Development*, vol. 7, no. 3, pp. 1247-1250, 2014.
26. J. E. Nash and J. V. Sutcliffe, "River flow forecasting through conceptual models part I—A discussion of principles," *Journal of hydrology*, vol. 10, no. 3, pp. 282-290, 1970.