

Original Research

On Prediction of Air Pollution Using Piecewise Affine Models

Jianfeng Guo, Zhenxing Ren*

College of Computer Science and Technology & College of Data Science, Taiyuan University of Technology, Shanxi, China

*Received: 22 January 2024**Accepted: 5 March 2024*

Abstract

Since air pollution affects both public health and economic growth, the issue has received more attention recently. Model-based early warning systems or pollution management tactics can be used to assist in combating dangerous air pollutants if accurate prediction models are available. This paper presents an approach to forecasting air contaminants using a piecewise affine model, which has a high prediction power. To identify the piecewise affine model, this study adopts effective clustering to identify the model. The proposed hierarchical clustering method improves the widely used BIRCH (Balanced Iterative Reducing and Clustering using Hierarchies) by adding a refining step to handle clusters with arbitrary geometries. Additionally, an optimization strategy like GA (Genetic Algorithm) is used to jointly estimate the model order and parameters. Measurements of Shenyang's air quality are used to illustrate the proposed approach, and the outcomes reflect the method's good prediction ability.

Keywords: piecewise affine model, prediction of air pollutants, clustering-based identification, model structure selection

Introduction

Due to swiftly accelerating economic growth, air pollution is currently a global public health issue. The natural environment, public health, and economic development are all impacted by air pollution. According to a meta-analysis, cardiovascular disease can increase by 11% on average with long-term exposure to $PM_{2.5}$ [1]. Over a nine-year follow-up period, a 10% increase in cardiovascular mortality was linked to an increase in $PM_{2.5}$ exposure in a large cohort aged 50-71 from six U.S. states and two metropolitan regions [2].

As a result, techniques for decision-making and analysis based on models may be created to enhance the creation of air protection plans [3]. When developing air pollution models, two basic methods are often employed: data-driven or empirical modeling and theoretical or deterministic modeling. Theoretical models that are often employed include Community Multiscale Air Quality (CMAQ) [4, 5] and Gaussian diffusion [6]. Air pollution forecasting has made extensive use of classical data-driven models such as the Autoregressive Integrated Moving Average (ARIMA) model [7], Support Vector Machine (SVM) [8, 9], Long Short-Term Memory (LSTM) [10, 11], and Bidirectional Long Short-Term Memory (Bi-LSTM) [12, 13]. Due to its great approximation capability, the PWA (Piecewise Affine) model is a crucial data-driven technique that has

*e-mail: renzhenxing@tyut.edu.cn

attracted a lot of attention and is already being utilized to solve difficult issues in simulation, prediction, and system analysis [14]. Typical application fields of PWA models are providing solutions for vehicle powertrains [14-17], robotics [18], power and energy [19], industry production processes [20-22], and rainfall runoff [23]. Also, the PWA modeling method was adopted for motion segmentation in computer vision [24, 25]. More details about applying PWA models can be found in [14]. By dividing the scheduling space into polyhedral convex regions and creating local models, the PWA model is developed. Partitioning regression space, estimating partition borders, and parameterizing each affine sub-model are typical processes involved in the discovery of PWA models [26]. The challenge of identifying PWA models is often regarded as difficult and is NP-hard [14]. The connection between identification and the classification problem is a major source of difficulty [27]. Numerous techniques have been presented in the past for identifying PWA models. It can provide a model with similar training data, which shortens the training period and increases generalizability. It is possible to estimate the sub-model parameters concurrently or later.

This work focuses on the identification of the PWA model with optimum model order selection using hierarchical clustering. The BIRCH (Balanced Iterative Reducing and Clustering using Hierarchies) technique has garnered a lot of attention in the past several decades as an effective agglomerative hierarchical clustering algorithm. However, clusters with arbitrary forms or changeable volumes should not use this approach [28]. This study introduces an additional phase to BIRCH, whereby partitions with the closest distances are merged until the desired number is reached. Next, the suggested

modeling approach is used to forecast Shenyang, China's air pollution. The proposed method is broad and might be extended to include different kinds of models. The following are this paper's primary contributions:

- To forecast air pollutants, a data-driven PWA model is presented. This model is defined by dividing the scheduling space into local regions.
- A hierarchical clustering-based method is adopted for the identification of the PWA model, and the optimization methodology estimates the model order and parameters.
- The performance of the proposed model is compared with the baseline models, and the proposed method is used to forecast air pollution in Shenyang.

Materials and Methods

Study Area

Shenyang, the largest city in northern China, serves as both the physical and transportation core for Northeast Asia and a key hub for the Belt and Road initiative, which connects the region to Southeast and Northeast Asia. Plains and mountains make up the majority of the landscape of this hilly area, which is situated in the southeast. Shenyang has four different seasons and a broad variety of temperatures due to its temperate, subhumid continental environment. The yearly average temperature ranges from 6.2 to 9.7°C, with considerable seasonal variations. Summer winds are generally southwest and northerly, winter winds are generally north, and spring winds are frequently southwest and northerly. Shenyang saw annual wind speeds ranging from 2.8 to 4.4 m/s, with springtime

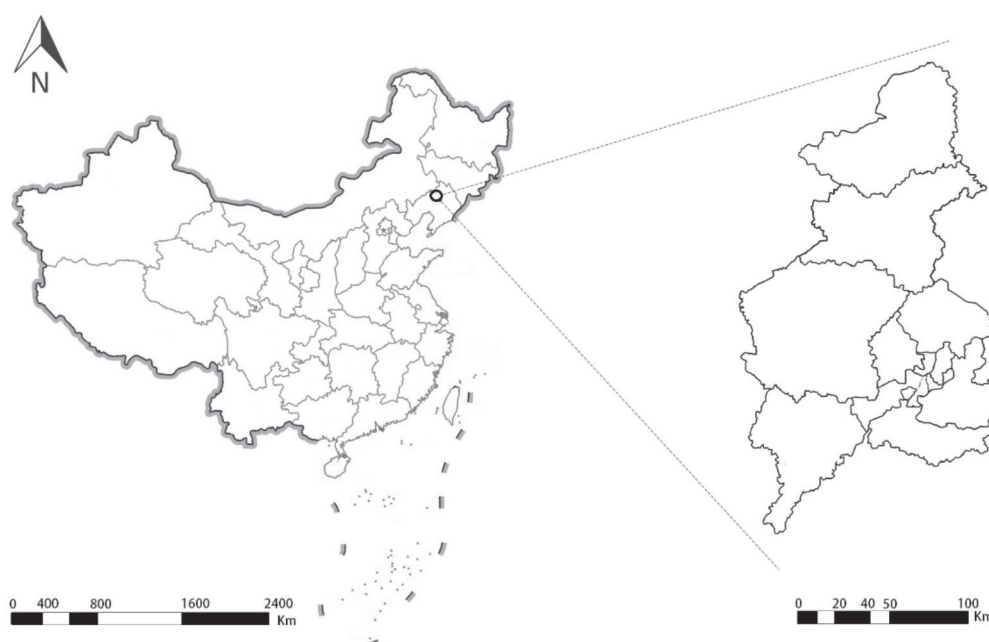


Fig. 1. Location of Shenyang in China.

increases and summertime decreases [29]. The location of Shenyang in China is shown in Fig. 1.

Transportation, central heating, and industrial activity are the key factors influencing Shenyang's air environment. Since the early days of the creation of New China, Shenyang, one of the first heavy industry bases in China, has established numerous heavy industry production businesses with a solid industrial basis. Its industrial structure is primarily responsible for the high coal use. For instance, in 2016, there was a significant increase in the amount of coal consumed overall – roughly 33 million tons – compared to the national average, which is much lower. Meanwhile, the share of renewable energy was less than 13%. Consequently, because of coal-based consumption, particulate matter (PM) is Shenyang's main source of air pollution; the city's annual $PM_{2.5}$ and PM_{10} daily exceed rates are 12.1% and 7.7%, respectively. Wintertime coal-burning heating raises $PM_{2.5}$ and PM_{10} levels, which leads to extreme pollution. The highest monthly average mass concentrations of $PM_{2.5}$ were seen in March and February, while the lowest were observed in August.

Data

As was already established, several variables affect air quality, including geography, environmental concerns, social or pollution emissions, and future air pollutants that can be predicted using previous variations. The main factor may be variations in climatic conditions, which have an impact on pollutant concentrations. It has been determined that meteorological data is a crucial input variable for air quality forecasting in statistical and mathematical models. As a result, the recommended forecasting models use meteorological and air pollution data as training sets. The two aspects that make up the datasets used in this study are meteorological variables and factors related to air quality. This contribution combines hourly data to train and assess the PWA air pollution forecasting model. Air pollutants, temperature, humidity, air pressure, dew point, and wind direction/speed are among the overserved data sets.

Method

The PWA model with hierarchical clustering-based identification is presented in this section.

Model Structure

The PWA model is defined as superposed c local models that explain the MISO (Multiple-Input-Single-Output) system. The predictor of air pollution in this research is the PWARX (Piecewise AutoRegressive eXogenous) model, which has an interpretable nonlinear structure [30]. The model is defined as:

$$y(k) = f(\mathbf{x}(k)) = \begin{cases} \boldsymbol{\theta}_1^T \begin{bmatrix} \mathbf{x}(k) \\ 1 \end{bmatrix} & \text{if } \mathbf{x}(k) \in \mathcal{X}_1 \\ \vdots \\ \boldsymbol{\theta}_c^T \begin{bmatrix} \mathbf{x}(k) \\ 1 \end{bmatrix} & \text{if } \mathbf{x}(k) \in \mathcal{X}_c \end{cases} \quad (1)$$

with the regressor:

$$\mathbf{x}(k) = [y(k-1)y(k-2) \dots y(k-n_a) \\ u(k-1)u(k-2) \dots u(k-n_b)]^T \quad (2)$$

Where n_a and n_b are model orders, respectively, $\{\boldsymbol{\theta}_i\}_{i=1}^c$ are the parameter vectors. The regression space is split into c polyhedral partitions, $\mathcal{X}_i \in \mathfrak{R}^{n_a+n_b}$ and each local model is valid in each partition. The approach presented in this paper is used to identify the PWA model. Since the measurement cannot be used directly, data processing is necessary before training.

Data Processing

The dataset's outliers and missing data will be handled first. Subsequently, the correlation analysis is carried out to choose pertinent attributes for the prediction model. Pearson's correlation coefficient (PCC), a commonly used statistical indicator to assess and investigate the degree of correlation between variables in practice, is used to quantify the correlation among characteristics. In time series data, Pearson's correlation coefficient generally explains the degree to which the goal variable (y) and the input variable (x) correlate over a certain period. The target variable y in this study reflects the measured value of air pollution, whereas the input variables represent the inputs to the PWA model. The following formula may be used to compute Pearson's correlation coefficient mathematically:

$$PCC = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2} \cdot \sqrt{\sum_{i=1}^N (y_i - \bar{y})^2}} \quad (3)$$

where \bar{x} is the average of x and \bar{y} is the average value of y . With +1 denoting total positive correlation, 0 denoting no connection, and -1 denoting total negative linear correlation between x and y , Pearson's correlation coefficient is between +1 and -1. A greater correlation coefficient value, which denotes a closer relationship between the detected characteristics, will lead to the selection of additional processing.

Then, to solve the issue of the curse of dimensionality, Principal Components Analysis (PCA) will be utilized. By keeping low principal components and ignoring high-order principal components, PCA can minimize the number of dimensions. Then, low-order components

may typically preserve important information. The core principle of PCA is to keep the majority of the data while condensing numerous variables into new, uncorrelated ones. The chosen features should be normalized before PCA dimension reduction since the projection will attempt to mimic any substantial premium that a feature may have in the data after the projection, which may lead to a sizable quantity of missing data.

Clustering Method

Then, the scheduling space will be portioned into local regions using clustering. The ‘‘similarity’’ of the data must be assessed for techniques that rely on clustering. As a common machine-learning approach, clustering is used to categorize the data points of related objects. Hierarchical and partitional clustering are two categories of clustering methods. The goal of partitional clustering is to minimize predetermined criteria for dividing data points into c divisions. Hierarchical clustering applies agglomerative and divisive methodologies to progressively investigate a hierarchy of clusters. Agglomerative techniques combine the nearest clusters repeatedly after collecting clusters from the bottom up. Data points are iteratively split into smaller clusters from the top down. BIRCH performs clustering based on a tree structure, namely the Clustering Feature (CF) Tree, as a common agglomerative clustering technique.

It is observed that in the typical BIRCH algorithm, the shape and size of clusters are determined by the same hyperparameter. However, clusters with arbitrary forms or changeable volumes should not use this approach [28]. A modified version of the BIRCH approach is presented in this study to enhance the performance of irregularly shaped clusters. The second step of clustering refines the first stage using the average linkage, which finds the average distance d_{rs} between each cluster pair C_r and C_s :

$$d_{rs} = \frac{1}{n_r n_s} \sum_{i=1}^{n_r} \sum_{j=1}^{n_s} \text{dist}(x_{ri}, x_{sj}) \quad (4)$$

where n_r and n_s are the numbers of data points in cluster C_r and C_s , x_{ri} and x_{sj} are the i th and j th data points in cluster C_r and C_s , $\text{dist}(x_{ri}, x_{sj})$ is the distance between x_{ri} and x_{sj} . The second step is carried out following standard BIRCH until the required number of clusters is reached.

Estimation of the Order and Parameters of Local Models

Once data points have been clustered using the suggested approach, the values of model parameters are estimated using

$$\hat{\Theta} = \underset{\Theta}{\text{argmin}} \sum_{k=1}^N (y(k) - \hat{y}(k, \Theta))^2 \quad (5)$$

To ascertain the model order denoted by n_a and n_b in Eq. (2), the Lasso-regularization is introduced by adding a regularization L_1 -term in Eq. (5).

$$\hat{\Theta} = \underset{\Theta}{\text{argmin}} \sum_{k=1}^N (y(k) - \hat{y}(k, \Theta))^2 + \lambda \cdot \|\Theta\|_1 \quad (6)$$

where $\|\cdot\|_1$ is the L_1 -norm and $\lambda > 0$ is a tuning parameter, which is a trade-off between model-fit and parameter variations. Eq. (6) can be solved by using global optimization methods.

Experiment

Model Quality Assessment

Three criteria are adopted to assess the quantitative quality of the model in this work. They are the correlation coefficient (R), the mean absolute error (MAE), and the root mean square error (RMSE):

$$\text{MAE} = \sqrt{\frac{\sum_{k=1}^N (\hat{y}(k) - y(k))^2}{N}} \quad (7)$$

$$\text{RMSE} = \sqrt{\sum_{k=1}^N (\hat{y}(k) - y(k))^2} \quad (8)$$

$$R = \frac{\sum_{k=1}^N (y(k) - \hat{y}(k))(\hat{y}(k) - \bar{\hat{y}}(k))}{\sqrt{\sum_{k=1}^N (y(k) - \hat{y}(k))^2 (\hat{y}(k) - \bar{\hat{y}}(k))^2}} \quad (9)$$

where N is the number of test data, y is the true value, \hat{y} is the predicted value, \bar{y} is the mean of the true values, and $\bar{\hat{y}}$ is the mean of the predicted values. Higher R values and results with decreased MAE and RMSE suggest that the model can achieve better prediction performance.

Data Processing

The proposed approach is used in the experiment to forecast the level of air pollution in Shenyang, based on the previously described technique. In the case study, the period of the model construction is from 2016 to 2018. Correlation analysis is conducted to choose pertinent characteristics based on PCC before determining the suggested PWA model. The feature correlation

analysis is displayed in Fig. 2. It is shown that the majority of air pollution features have positive correlations, except O_3 , which is negatively correlated with other features. In Fig. 2 there is a weaker link between meteorological and air pollution characteristics; all parameters except wind and atmospheric pressure have a negative correlation with air pollutants. As previously mentioned, accuracy may be increased by excluding elements that are uncorrelated or weakly linked. In this work, a threshold of 0.2 is specified, and features with correlation coefficients greater than 0.2 are deemed to be strongly associated.

To get over the dimensionality problem, the PCA approach helped to lower the dimension of the number of selected features. The PCA method result shows that over 90% of the information supplied by the complete features is contained in the first three principal

components. Therefore, in addition to the anticipated values, these three principal components, which contain important information from the data, are employed for the prediction of air pollution. PCA has been used to extract.

Results and Discussion

Simulation Results

The proposed model will be used to forecast air pollution. In Fig. 3, the observed data is represented by the blue line, while the $PM_{2.5}$ forecast made by the PWA model is shown by the red line. The estimate accurately depicts the hourly $PM_{2.5}$. The model's forecast and the measurement are rather close, which means that

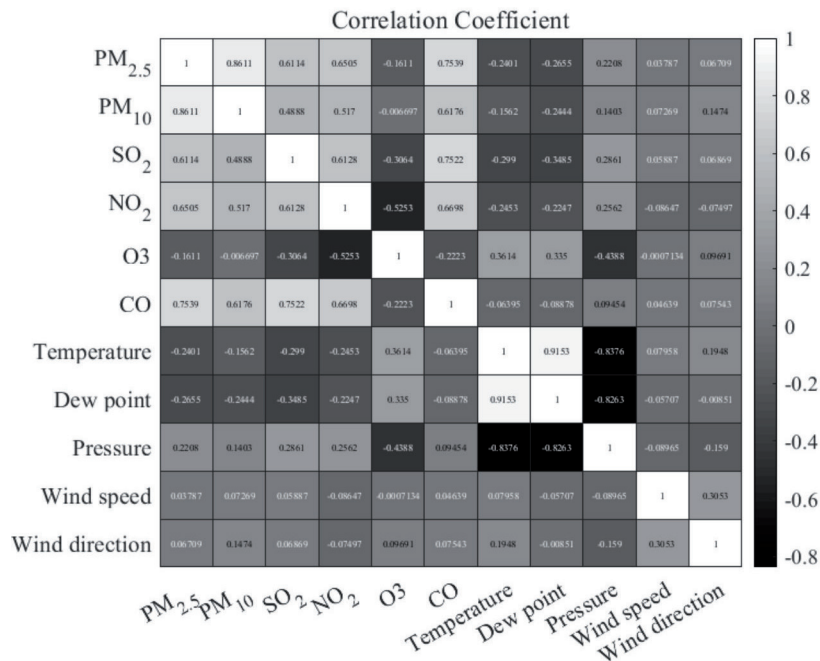


Fig. 2. Correlation analysis between various features.

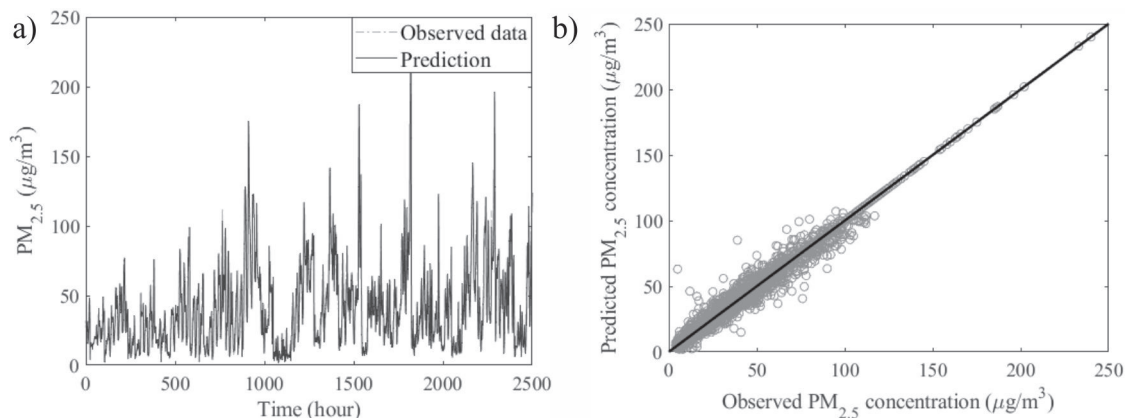


Fig. 3. PWA model for predicting $PM_{2.5}$ a) Comparison between observed data and prediction b) fit comparison between observed data and prediction.

the model can precisely predict $PM_{2.5}$, according to the testing samples. The correlation coefficient (R) between the measurement and prediction of $PM_{2.5}$ is 0.9849, indicating that almost 98% of the variation was explained by the model. It is also possible to accurately

forecast the majority of peak positions, and the prediction curve closely resembles the observed curve. It shows that significant state changes can be handled by the suggested paradigm. However, as Fig. 3 illustrates, the prediction could not consistently fit the actual

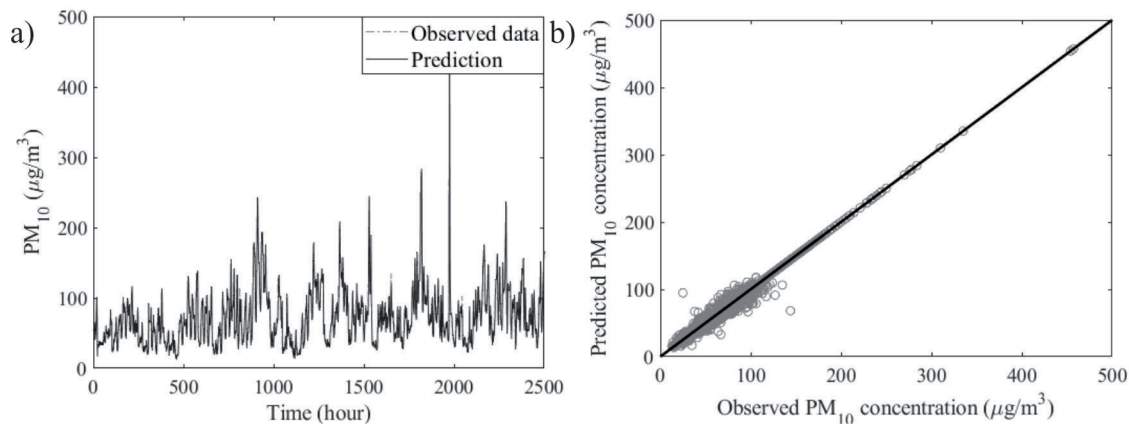


Fig. 4. Comparison between observed data and prediction with the PWA model for predicting PM_{10} .

Table 1. Performance of the PWA model.

Assessment	Air pollutant				
	$PM_{2.5}$	PM_{10}	SO_2	NO_2	CO
MAE	3.1432	2.8859	2.5238	3.1185	0.0770
RMSE	5.1462	5.2701	4.2518	4.6505	0.1211
R	0.9856	0.9914	0.9544	0.9654	0.9708

Table 2. Comparison of different models.

Pollutant	Assessment	Model type					
		ARIMA	SVM	MLP	LSTM	Bi-LSTM	PWA
$PM_{2.5}$	MAE	4.8397	4.9947	5.4228	4.1131	3.9283	3.1432
	RMSE	7.9783	8.0555	9.1091	6.9221	6.7698	5.1462
	R	0.9766	0.9730	0.9678	0.9815	0.9842	0.9856
PM_{10}	MAE	4.4571	4.6712	4.9376	3.9819	3.7758	2.8859
	RMSE	8.1971	8.2387	9.9853	7.3213	6.8945	5.2701
	R	0.9801	0.9758	0.9723	0.9865	0.9889	0.9914
SO_2	MAE	3.5284	3.8103	4.2158	3.2901	3.0921	2.5238
	RMSE	6.8816	7.2311	7.9701	6.3551	5.8911	4.2518
	R	0.9412	0.9378	0.9328	0.9445	0.9476	0.9544
NO_2	MAE	4.2924	4.7609	5.1291	3.8811	3.6354	3.1185
	RMSE	5.6025	6.2190	6.5452	5.3019	5.1922	4.6505
	R	0.9551	0.9512	0.9489	0.9577	0.9593	0.9654
CO	MAE	0.1401	0.1468	0.1569	0.1380	0.1299	0.0770
	RMSE	0.1858	0.1922	0.2021	0.1740	0.1682	0.1211
	R	0.9629	0.9577	0.9442	0.9631	0.9688	0.9708

data with lower concentration values when $PM_{2.5}$ is less than $110 \mu\text{g}/\text{m}^3$. On the other hand, greater $PM_{2.5}$ values may be predicted with accuracy by the PWA model.

When other air pollution, like PM_{10} , is predicted using the proposed approach, the prediction values generally still match the observed data. Fig. 4 results show that the proposed approach may be effectively used for PM_{10} prediction with a similar outcome to $PM_{2.5}$. As all of the pollutants in Table 1 have R values for the prediction over 0.95, the model that was provided can accurately forecast the air pollution.

Comparison with Other Methods

This work compares the proposed PWA model with several baseline techniques, like SVM, MLP, ARIMA, LSTM, and the Bi-LSTM model, to evaluate the effectiveness of the proposed approach. The outcomes of the suggested PWA model are shown in bold. In this study, several models are built using the same data sets, but the input sequences are altered by the structural differences between the models. The proposed model is contrasted with alternative models for $PM_{2.5}$ prediction in Table 2. Due to the constraints of the model architectures, MLP, ARIMA, and SVM display significantly more significant mistakes, although both LSTM and Bi-LSTM, two deep learning techniques, performed better. Table 2's MAE, RMSE, and R values show that the model that was proposed for this study performed better in terms of prediction than the other baselines. Additionally, it is better able to depict the features of pollutant concentrations. Overall, the performance of the suggested model was enough for the prompt implementation of additional safety measures using realistic prediction tasks.

Conclusions

The data-driven PWA method was proposed in this work to predict air pollution. Initially, a new clustering-based technique was used to find a data-driven PWA model for air pollution prediction. The widely used BIRCH algorithm was enhanced by the suggested clustering approach, which added a refinement step to handle clusters of any shape. Next, global optimization is used to estimate both the model order and parameters at the same time.

This study employed this approach to forecast Shenyang, China's air pollution, and it compares the performance of the suggested model to the baseline models. The proposed models for predicting air pollution were compared against several baseline models. The proposed approach, according to the results, might successfully generate models with higher quality that are more suitable for early warning of higher air pollution concentrations and for developing trustworthy management measures for effective environmental

protection. The proposed approach may be used in future research to forecast more air contaminants in different regions. The second research would target the enhancement of model quality by taking into account other potential affecting factors.

Acknowledgments

The Natural Science Foundation of Shanxi Province, China (Grant No. 20210302123188), the National Natural Science Foundation of China (Grant No. 12204345), and the Natural Science Foundation for Young Scientists of Shanxi Province, China (Grant No. 202103021223107) all provide funding for this work.

Conflict of Interest

The authors declare no conflict of interest.

References

1. HOEK G., KRISHNAN R.M., BEELEN R.M.J., PETERS A., OSTRO B., BRUNEKREEF B., KAUFMAN J.D. Long-term air pollution exposure and cardio- respiratory mortality: a review. *Environmental Health*. **12**, 43, **2013**.
2. THURSTON G.D., AHN J., CROMAR K.R., SHAO Y., REYNOLDS H.R., JERRETT M., LIM C.C., SHANLEY R.P., PARK Y., HAYES R. B. Ambient Particulate Matter Air Pollution Exposure and Mortality in the NIH-AARP Diet and Health Cohort. *Environmental Health Perspectives*. **124**, (4), 484 **2015**.
3. YANG H., ZHU Z., LI C., LI R. A novel combined forecasting system for air pollutants concentration based on fuzzy theory and optimization of aggregation weight. *Applied Soft Computing*. **87**, 105972, **2020**.
4. THONGTHAMMACHART T., ARAKI S., SHIMADERA H., ETO S., MATSUO T., KONDO A. An integrated model combining random forests and WRF/CMAQ model for high accuracy spatiotemporal PM2.5 predictions in the Kansai region of Japan. *Atmospheric Environment*. **262**, 118620, **2021**.
5. LAGO KITAGAWA Y.K., PEDRUZZI R., GALVÃO E.S., BAPTISTA DE ARAÚJO I., TOLEDO DE ALMEIDA ALBURQUERQUE T., KUMAR P., SPERANDIO NASCIMENTO E.G., MOREIRA D.M. Source apportionment modelling of PM2.5 using CMAQ-ISAM over a tropical coastal-urban area. *Atmospheric Pollution Research*. **12** (12), 101250, **2021**.
6. YANG Z., YAO Q., BUSER M.D., ALFIERI J.G., LI H., TORRENTS A., MCCONNELL L.L., DOWNEY P.M., HAPEMAN C.J. Modification and validation of the Gaussian plume model (GPM) to predict ammonia and particulate matter dispersion. *Atmospheric Pollution Research*. **11** (7), 1063, **2020**.
7. GARCÍA NIETO P.J., SÁNCHEZ LASHERAS F., GARCÍA-GONZALO E., DE COS JUEZ F.J. PM10 concentration forecasting in the metropolitan area of Oviedo (Northern Spain) using models based on SVM, MLP, VARMA and ARIMA: A case study. *The Science of The Total Environment*. **621**, 753, **2018**.

8. SU X., AN J., ZHANG Y., ZHU P., ZHU B. Prediction of ozone hourly concentrations by support vector machine and kernel extreme learning machine using wavelet transformation and partial least squares methods. *Atmospheric Pollution Research*. **11** (6), 51, **2020**.
9. LEONG W.C., KELANI R.O., AHMAD Z.A. Prediction of air pollution index (API) using support vector machine (SVM). *Journal of Environmental Chemical Engineering*. **8** (3), 103208, **2020**.
10. LU X., SHA Y., LI Z., HUANG Y., CHEN W., CHEN D.H., SHEN J., CHEN Y., FUNG J.C.H. Development and application of a hybrid long-short term memory - three dimensional variational technique for the improvement of PM2.5 forecasting. *The Science of The Total Environment*. **770**, 144221, **2021**.
11. WU C.L., HE H.D., SONG R.F., PENG Z.R. Prediction of air pollutants on roadside of the elevated roads with combination of pollutants periodicity and deep learning method. *Building and Environment*. **207**, 108436, **2021**.
12. MA J., LI Z., CHENG J.C.P., DING Y., LIN C., XU Z. Air quality prediction at new stations using spatially transferred bi-directional long short-term memory network. *The Science of The Total Environment*. **705**, 135771, **2020**.
13. ZHANG B., ZHANG H., ZHAO G., LIAN J. Constructing a PM2.5 concentration prediction model by combining auto-encoder with Bi-LSTM neural networks. *Environmental Modelling & Software*. **124**, 104600, **2020**.
14. REN Z. An optimized excitation signal design for identification of PWA model and application to automotive throttles. *Measurement and Control*. **56** (3-4), 844, **2023**.
15. REN Z., KROLL A., SOFSKY M., LAUBENSTEIN F. On Physical and Data-Driven Modeling of Systems with Friction: Methods and Application to Automotive Throttles. *At-Automatisierungstechnik*. **61** (3), 155, **2013**.
16. SUN X., WU P., CAI Y., WANG S., CHEN L. Piecewise affine modeling and hybrid optimal control of intelligent vehicle longitudinal dynamics for velocity regulation. *Mechanical Systems and Signal Processing*. **162**, 108089, **2022**.
17. ZHANG Q., JING H., LIU Z., JIANG Y., GU M. A Novel PWA Lateral Dynamics Modeling Method and Switched T-S Observer Design for Vehicle Sideslip Angle Estimation. *IEEE Transactions on Industrial Electronics*. **69** (2), 1847, **2022**.
18. MOUSTAKIS N., ZHOU B., LE QUANG T., BALDI S. Fault detection and identification for a class of continuous piecewise affine systems with unknown subsystems and partitions. *International Journal of Adaptive Control and Signal Processing*. **32** (7), 980, **2018**.
19. SINDAREH-ESFAHANI P., PIEPER J. Machine learning-based piecewise affine model of wind turbines during maximum power point tracking. *Wind Energy*. **23** (2), 404, **2020**.
20. WANG J., SONG C., ZHAO J., XU Z. A PWA model identification method for nonlinear systems using hierarchical clustering based on the gap metric. *Computers & Chemical Engineering*. **138**, 106838, **2020**.
21. XIN J., NEGENBORN R.R., LIN X. Piecewise affine approximations for quality modeling and control of perishable foods. *Optimal Control Applications and Methods*. **39** (2), 860, **2018**.
22. BEST D., BUKKEMS B.B., MOLENGRAFT V.D.R.R., HEEMELS W.P.M. H., STEINBUCH M. Robust control of piecewise linear systems: A case study in sheet flow control. *Control Engineering Practice*. **16**, 991, **2008**.
23. HADID B., DUVIELLA E., LECOEUICHE S. Data-driven modeling for river flood forecasting based on a piecewise linear ARX system identification. *Journal of Process Control*. **86**, 44, **2020**.
24. VIDAL R., SOATTO S., CHIUSO A. Applications of hybrid system identification in computer vision. *2007 European Control Conference (ECC)*. 4853, **2007**.
25. VIDAL R., MA Y. A Unified Algebraic Approach to 2-D and 3-D Motion Segmentation and Estimation. *Journal of Mathematical Imaging and Vision*. **25**, 403, **2006**.
26. SINDAREH-ESFAHANI P., PIEPER J.K. Machine learning-based piecewise affine model of wind turbines during maximum power point tracking. *Wind Energy*. **23**, 404, **2020**.
27. YANG Y., XIANG C., GAO S., LEE T.H. Data-driven identification and control of nonlinear systems using multiple NARMA-L2 models. *International Journal of Robust and Nonlinear Control*. **28** (12), 3806, **2018**.
28. WAN L., YANG J. Advanced Split BIRCH Algorithm in Reconfigurable Network. *Journal of Networks*. **8** (9), 2050, **2013**.
29. MA Y., WANG M., WANG S., WANG Y., FENG L., WU K. Air pollutant emission characteristics and HYSPLIT model analysis during heating period in Shenyang, China. *Environmental Monitoring and Assessment*. **193**, 9, **2020**.
30. GU Y., LI B., MENG Q. Hybrid Interpretable Predictive Machine Learning Model for Air Pollution Prediction. *Neurocomputing*. **468**, 123, **2021**.