J Inf Process Syst, Vol.21, No.3, pp.296~307, June 2025
https://doi.org/10.3745/JIPS.04.0350
ISSN 2092-805X (Electronic)

JOURNAL OF INFORMATION PROCESSING SYSTEMS JIPS

Research on Cooperative Early Warning Coastal Air Environment of Liaoning Based on Multi-Feature and Multi-Model Fusion

Xiuli Yu and Peng Du*

Abstract

Enterprises in the coastal regions of China release abundant pollutants that have considerably deteriorated the air quality. To address this issue, an information fusion technology has been proposed herein for predicting coastal air quality in Liaoning Province. To this end, real-time data analysis of water, air, and soil pollutants emitted from diverse coastal enterprises was performed using a multi-model selection strategy for ensemble learning. This approach integrated meteorological information and considered the unique learning principles and observational disparities among various algorithms. The proposed approach explored the influence of collaborative early warning of multi-feature pollution source emissions on the surrounding environment. By combining the base learner and meta-learner in the multi-model fusion strategy, the ensemble model yielded better prediction results, particularly using strong learners at the primary level and linear mode at the secondary level. This optimal combination strategy was used to develop a collaborative monitoring and early warning model, which incorporated multi-feature data from water, air, and soil sources in the coastal environment. This multi-feature collaboration enhanced the prediction accuracy compared with that of single-feature models and further amplified the early warning capabilities enabled by multi-model fusion.

Keywords

Cooperative Early Warning, Ensemble Learning, Multi-Feature Fusion, Multi-Model Fusion

1. Introduction

Since the economic reform, China's GDP has grown rapidly due to rapid industrialization and urbanization. The acceleration of coastal and heavy industries has fostered developments in the coastal areas of the country. However, steel, oil refining, and chemical industries mainly distributed in coastal areas discharge industrial pollutants and domestic wastewater. This has caused serious pollution and threatened sustainable urban development. Therefore, early warning systems based on deep learning must be urgently developed to ensure effective and accurate supervision and governance of marine environment. Existing monitoring systems for marine pollution provide data on pollutant discharge; however, these data are not representative of coastal environmental conditions. Excessive amounts of data can cause information overload, thereby preventing efficient data processing that can eventually impact disaster management. Moreover, the migration and exchange of pollutants in water, air, and soil

Manuscript received January 17, 2024; first revision July 16, 2024; accepted August 18, 2024.
***Corresponding Author:** Peng Du (dupeng@dlou.edu.cn)
School of Marine Law and Humanities, Dalian Ocean University, Liaoning Dalian, China (yuxiuli@dlou.edu.cn, dupeng@dlou.edu.cn)

have not been sufficiently studied. By analyzing the factors that influence these mechanisms and identifying the correlation between these three media, coastal pollution management and emission reduction measures can be devised. Additionally, a water–air–soil coordinated warning system must be developed for coastal environments such that coastal ecological issues can be detected and mitigated in advance using modern and intelligent reasonable countermeasures. This will help the government to undertake timely decisions during emergencies.

Several methods that utilize statistical data have been proposed for air pollution prediction, which offer insights into evolving environmental conditions and facilitate informed decision-making and proposing mitigation strategies. Wu and Lin [1] proposed a hybrid VMD-SE-LSTM prediction model to reflect the characteristics of the original air quality index (AQI) sequence. Wang et al. [2] designed an L1-norm for air pollution monitoring and analysis. Wang and Yang [3] developed a warning system that integrated time-varying filtering-based empirical mode decomposition for air pollution forecasting. Song et al. [4] quantitatively and comparatively analyzed the effectiveness of environmental management in the coastal countries of Northeast Asia. They offered guidance methods and practical experience for developing a marine environmental management system in China. Urban air quality can also be predicted via numerical simulation and statistical methods [5]. Although numerical simulation is highly scientific, its application is limited by limited data; it is also difficult to use. Statistical methods can be categorized into simple empirical statistics and machine learning. Machine learning involves using models for air quality prediction based on pollution indicators such as meteorological indicators as well as remote sensing and terrain data. As mechanistic mathematical models are difficult to use, information fusion technology has emerged as a novel alternative for predicting the impacts of environmental pollution.

The optimization problem of a single model has relatively reached its performance limit in the process of continuous improvement of algorithm models used for pollution prediction. Consequently, ensemble learning is used for such predictions as it combines the predictions of multiple machine learning models or classifiers to offer highly accurate and stable results. For instance, Zhou [6] reported that low correlation among single models can enhance error correction in ensemble leaning, thereby yielding highly accurate and robust results. Ensemble learning is categorized into boosting, bagging, and stacking. The stacking algorithm effectively uses predictions generated by different classifiers as the input to the next layer of learning algorithms. It can integrate the learning mechanisms of different learning algorithms, establish prediction models based on the differences between them, and obtain the final prediction results by employing appropriate combination strategies. Although ensemble learning offers good results, it can be considerably improved [7] for wide applicability in remote sensing, facial recognition, disease detection, and other fields [8-11]. The fusion strategy that combines predictions from base-learners and meta-learners in stacking has been scarcely studied and must be further explored to enable multi-model fusion with ensemble learning.

The air quality in the coastal regions of China has considerably deteriorated due to high-pollutant emissions from various industries such as coal, metallurgy, and petrochemical. As these industries are crucial for China's economy, a comprehensive approach to managing pollutants must be devised. Air–water exchange is the most active pathway for the exchange of pollutants [12,13] such as petroleum hydrocarbons and microplastics. The migration and transformation of these pollutants have been analyzed in several studies, which revealed that they dissolve in water vapor or permeate and migrate to the soil under the action of gravity as a free phase [14-17]. Water and soil contamination exacerbate air pollution, particularly in coastal regions; therefore, these factors must be studied in depth to devise comprehensive and effective environmental strategies for holistically addressing coastal air pollution.

In this study, a multi-model fusion technology that integrates detailed information on coastal

environmental pollution is proposed for predicting the AQI in coastal areas. This technology uses water, air, and soil pollutant emission data from different coastal enterprises to analyze their impact on environmental pollution. By leveraging the unique learning principles and observation nuances of various machine learning algorithms, a collaborative monitoring and alerting framework for coastal environments was developed that focused on water, air, and soil properties. This framework employs the multi-model fusion technology for providing ecological security early warning in coastal environments.

2. Research Method

Soil contaminants in coastal environments undergo complex transformations and migrations influenced by wind and water dynamics and migrate into the atmosphere and aquatic systems, causing air, surface water, and groundwater pollution. Similarly, waterborne pollutants transform and disperse into the atmosphere and soil, causing further contamination. As these pollution sources are interdependent, poor coastal air quality is not caused not only by atmospheric emissions alone but also by water and soil pollutants. Herein, air, water, and soil pollution data were integrated and a collaborative early warning model was developed based on multi-feature information fusion. Real-time monitoring data were leveraged to develop an optimal pollutant diffusion model for accurately predict the impact of pollutants on the coastal air environment, thereby fostering a more comprehensive and proactive approach to environmental management.

2.1 Proposed Methods

Taking model information as the core, a framework for collaborative monitoring and early warning of water–air–soil multi-feature is established, focusing on solving prominent coastal environmental problems. By combining information from different modalities, multi-model fusion enabled achieving good prediction accuracy. By obtaining early warnings of the coastal ecological environment, the government can proactively respond to pollution in such areas. A prediction method based on supervised learning and multi-model fusion was proposed. A selection strategy for ensemble learning multi-model was proposed to explore the collaborative warning of water–air–soil multi-feature based on multisource information fusion technology for coastal enterprises' pollution source emissions on surrounding air environmental pollution. The prediction accuracy was considerably improved by employing multi-model selection strategies and multi-feature fusion methods. The proposed framework can flexibly and effectively coordinate the early warnings of atmospheric heavy pollution events in coastal areas, thereby guiding enterprise pollution source emissions.

2.1.1 Evaluation criteria and influencing factors

To prevent pollution and safeguard the quality of coastal environments, the potential consequences of environmental pollutants emitted from construction projects on coastal regions must be analyzed thoroughly. This study focused on the Liaodong Peninsula, a region characterized by a continental monsoon climate within the northern temperate zone. This region experiences continental and marine climates due to its unique geographical position, i.e., surrounded by sea on three sides, rendering it susceptible to both marine and monsoon influences. Wastewater and waste residue also considerably impact the coastal air quality besides atmospheric pollutants. Herein, the primary factors that contribute to coastal environmental pollution have been studied based on the characteristics of water, air, and soil

emission sources and meteorological factors. The emission sources mainly come from pollution sources in production and daily life. The diffusion of pollutants is influenced by the meteorological environment in the region such as atmospheric turbulence. Atmospheric dilution and diffusion vary with meteorological conditions, resulting in different influence ranges and intensities of pollutants. The main meteorological factors that affect atmospheric conditions include temperature, pressure, humidity, wind direction, and wind speed.

2.1.2 Coastal air collaborative early warning model based on multi-feature and multi-model fusion

For predicting coastal ecological environment pollution, the air quality dynamics are intricately intertwined with both pollutant emission sources and meteorological variables. For predicting coastal air pollution, the analysis scope must be broadened beyond solely examining the emission characteristics of atmospheric pollutants. To this end, the emission profiles of water and soil pollutants and meteorological factors such as temperature, humidity, and wind direction were comprehensively evaluated herein. This multi-feature approach enabled providing highly accurate and timely air pollution level warnings, thereby enhancing the understanding and preparedness during disasters in coastal regions. A collaborative coastal air quality early warning system based on multi-feature and multi-model fusion was developed, as shown in Fig. 1.

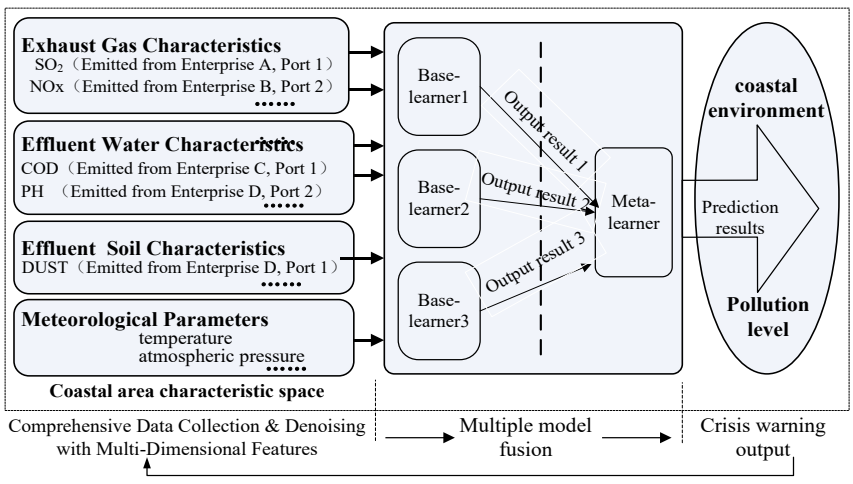

Fig. 1. A multi-feature and multi-model fusion approach for establishing a collaborative coastal air quality early warning system.

The technical approach for developing this system can be summarized as follows. First, the data on diverse water–air–soil and meteorological factors emanating from coastal industrial pollution sources undergo comprehensive acquisition and denoising procedures. Then, a collaborative warning system was developed by leveraging multi-model fusion that was tailored to the water–air–soil characteristics of coastal ecosystems. Herein, an array of characteristic parameters such as air pollutants (SO_2, NOx, CO, and smoke); waste water contaminants (sulfides and petroleum); solid pollutants (particulate matter); and meteorological factors (wind speed, wind direction, atmospheric pressure, and temperature) were considered as a parameter set. These were used for assessing the level of atmospheric pollution within coastal monitoring areas in terms of AQI, which is a quantifiable metric for air quality assessment.

2.2 Stacking Ensemble Learning Fusion Strategy

2.2.1 Selection of multi-model fusion

First, a single model was applied to forecast the coastal air pollution using the mean absolute error (MAE) as a metric to evaluate the prediction performance of the model on the test set.

$$\mathrm{MAE} = \frac{1}{m}\sum_{i=1}^{m} |(y_i - \hat{y}_i)| \tag{1}$$

where y_i and $\hat{y}_i$ denote the real and predicted values of AQI.

Aiming at the differences between different learning algorithms in training and reasoning, a prediction method based on a multi-model fusion ensemble learning mode was proposed to achieve better learning performance than a single algorithm. It fully leveraged the advantages of each model. To achieve optimal ensemble learning performance, the learning capabilities of multi-model algorithms were analyzed. As show in Fig. 2, XGBoost, gradient boosting machine (GBM), and kernel ridge regression (KRR) exhibit lower MAE values. This indicated their enhanced performance, and they were classified as "strong learners." In contrast, decision trees (DT), linear regression with Lasso regularization, and support vector regression (SVR) models exhibited weaker performance and were classified as "weak learners."

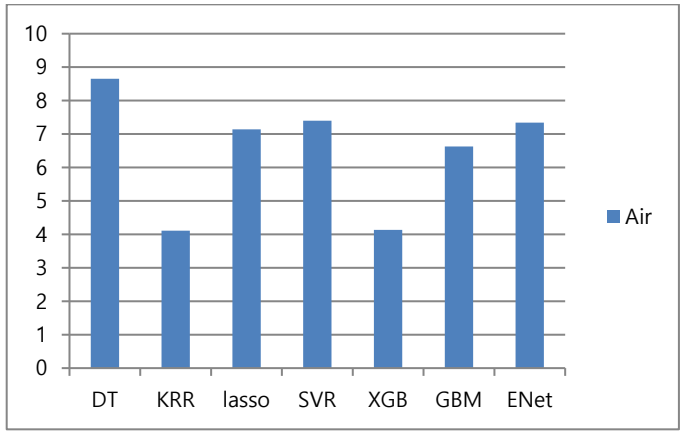

Fig. 2. Efficacy multiple models in predicting air quality by MAE.

2.2.2 Multi-model fusion strategy

When training the model on excessive data, a more powerful combination strategy known as the "learning method" that combines basic and meta-learners can be used. If the basic learner has strong ability, its predicted results are not considerably different than the real result; this indicates their linear relationship. To prevent model over fitting and reduce the impact of classification errors of these learners, the meta-learner uses a relatively simple linear regression model to yield accurate prediction results [18]. However, if the base learner has weak learning ability, the predicted results may differ considerably from the real results and may not be linearly related. In such cases, the meta-learner chooses a nonlinear regression model. When the base-learner has strong learning ability and the meta-learner is nonlinear, it is equivalent to repeating the learning twice. Such cases were therefore not considered in the fusion strategy design. The impact of stacking ensemble learning with only one layer of structural learners

(Stack_1) on the prediction performance was also analyzed, i.e., the case wherein the base learner is a strong learner and there is no meta-learner. The strength of the base-learner and the linearity of the meta-learner, as well as how different combinations affect the final prediction performance of stacking algorithm, were determined to classify the base-learner and meta-learner used in multi-model fusion strategy. Table 1 shows the multi-model fusion strategy of the stacking algorithm.

This multi-model fusion strategy improves the accuracy and stability of prediction. By combining various base learners, ensemble learning performance of the stacking algorithm was improved [19]. By selecting a base-learner with strong learning ability in the first layer, the predictive ability of the stacking model improved. Herein, the XGB algorithm was also used (Fig. 2). As the base-learner, the GBM adopted the boosting algorithm's ensemble learning method, with strong learning generalization ability. It completed causal mathematical interpretation and was used as the base learner herein. DT used the ensemble learning method of the bagging algorithm. SVR has unique advantages in solving small-sample, nonlinear, and high-dimensional pattern recognition problems, making it an effective regression tool. The second layer used this linear regression model to correct the inductive bias of several learning algorithms and reduce the risk of model overfitting.

Table 1. Multi-model fusion strategy of stacking algorithm

Integrated model	Integration strategy	Algorithms included
Stack_1	Base-learner Strong regression model	XGB; GBM; Light GBM
	Meta-learner	-
Stack_2	Base-learner Strong regression model	XGB; GBM; Light GBM
	Meta-learner linear regression model	Lasso regression
Stack_3	Base-learner Weak regression model	DT; Linear regression; SVR
	Meta-learner linear regression model	Lasso regression
Stack_4	Base-learner Weak regression model	DT; Linear regression; SVR
	Meta-learner Strong regression model	XGB

3. Empirical Research

3.1 Experimental Data

Herein, AQI data spanning from August 10, 2020 to June 30, 2022, in D City, were acquired. After pruning invalid records, a dataset of 643 observations was established to evaluate the efficacy of the model. Six key air pollutants were assessed, namely O_3, PM_{10}, $PM_{2.5}$, CO, SO_2, and NO_2, along with the ambient AQI sourced from the data center of the local environmental monitoring station. Emission characteristic data of industrial pollutants (air, water, and solid pollutants) were procured from the self-monitoring information disclosure platform of major pollutant-discharging units in Liaoning Province. Moreover, 33 enterprises located within the city's effective coastal region were focused on, and 204 emission data features were extracted from this platform; these comprised 128 gas emission features, 66 water emission features, and 10 soil emission features. By leveraging information fusion techniques, the study aimed to analyze and provide dynamic alerts on the coastal environmental pollution status within specified areas. The extensive dataset of pollutant emissions from coastal enterprises was used for this purpose. Table 2 shows the sample monitoring data.

Table 2. Sample data emission features and meteorological features in the study area

Monitor time	Air (A)-Water (W)-Soil (S) Emission characteristic parameters						Meteorological characteristic parameters		
	SO_2 -A1	NOx -A2	COD -W1	pH -W2	DUST -S1	…	Atmospheric pressure	Temperature	…
2020/8/10 0:00	11.49	67.58	19.67	7.1	2.60695	…	99.5	23.2	…
2020/8/10 2:00	2.3	73.1	19.21	7.13	2.53756	…	99.5	22.9	…
2020/8/10 4:00	7.79	74.82	19.49	7.28	2.64507	…	99.5	22.4	…
2020/8/10 6:00	11.63	69.91	18.46	7.35	2.54572	…	99.6	23.5	…
2020/8/10 8:00	22.09	68.96	18.07	7.37	2.34194	…	99.7	26.2	…
2020/8/10 10:00	3.37	62.27	29.36	7.28	2.25389	…	99.6	26.2	…
2020/8/10 12:00	3.26	60.32	23.33	7.29	2.2633	…	99.7	25.6	…
2020/8/10 14:00	3.1	66.48	21.41	7.26	2.27371	…	99.6	25.8	…
2020/8/10 16:00	2.63	67.59	19.34	7.17	2.15278	…	99.6	25.5	…
2020/8/10 18:00	18.31	70.73	37.04	7.21	2.33118	…	99.6	24.4	…
2020/8/10 20:00	9.95	70.66	40.55	7.23	2.42917	…	99.7	23	…
2020/8/10 22:00	15.11	58.52	26.11	7.22	2.60341	…	99.7	22.1	…
2020/8/11 0:00	23.08	60.74	24.11	7.21	2.5758	…	99.8	22.6	…
2020/8/11 2:00	5.21	57.75	24.94	7.21	2.49957	…	99.8	23.1	…
2020/8/11 4:00	4.1	57.83	24.32	7.22	2.47865	…	99.9	22.6	…
2020/8/11 6:00	20.07	68.38	22.33	7.17	2.41544	…	100	23.1	…
2020/8/11 8:00	16.35	66.07	23.08	7.14	2.30051	…	100.1	25.9	…
2020/8/11 10:00	9.18	60.89	22.94	7.13	2.02514	…	100.2	26.4	…
2020/8/11 12:00	9.67	60.57	22.49	7.13	2.10234	…	100.2	27.9	…
2020/8/11 14:00	5.93	55.34	22.5	7.12	2.24038	…	100.2	26.5	…

3.2 Experiment and Analysis

The gas, water, soil emission features and meteorological indicators were consolidated into a comprehensive parameter set to assess the likelihood of atmospheric pollution at the monitoring stations located within core coastal residential zones. The ultimate output, denoted as parameter y, represents the AQI that provides a quantitative measure of environmental health. In addition to the representative machine learning algorithm provided for single model prediction and the multi-model fusion strategy (Table 1), the stacking algorithm was trained and tested via five-fold cross-validation. The performance of various integrated models was evaluated based on MAE, and the results are shown in Fig. 3.

The combination strategies of ensemble learning have different impacts on prediction performance. Among them, Stack_2, which uses a strong learner as the input to the base-learner and linear regression as the meta-learner, yields the best prediction results. Traditionally, ensemble learning fully leverages the advantages of multiple models and adopts various strategies to integrate different learning outcomes to improve prediction results. However, herein, the MAE did not improve considerably under Stack_4 and the prediction results were not better than that obtained using a single learner. These results indicate that the combination strategy of base-learners and meta-learners impacts the

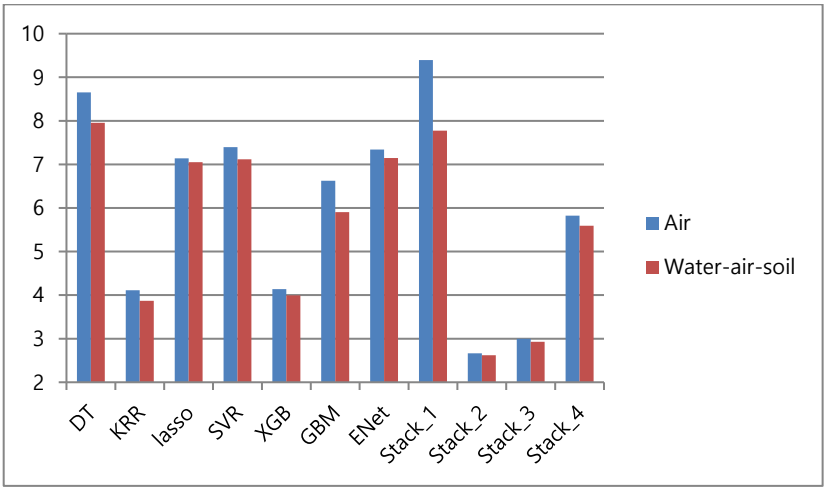

Fig. 3. Performance of machine learning algorithm in water–air–soil multi-feature cooperative early warning system (MAE).

results of ensemble learning. Stack_2 yields better prediction results for most scenarios. Stack_2 had higher MAE than high-performance single models. This indicated that not every type of stacking offers "quality assurance." Multi-model fusion may not necessarily yield good prediction results when applied to certain specific and specific scenarios.

The red bars in Fig. 3 show that with all other parameters constant, the characteristic data of air pollutant emissions has been broadened to encompass multi-feature emissions spanning water, air, and soil, facilitating the assessment of pollution potential in coastal core residential areas' current monitoring points. A comparative analysis revealed that the MAE for water–air–soil multi-feature collaborative warnings was generally lower than that of air pollutant emission warnings, indicating superior predictive performance. This underscores that atmospheric environmental forecasting is influenced by not only air pollutant emissions but also pollutant migration, transformation, as well as wastewater and waste residue discharges; these parameters collectively impact environmental quality and may induce secondary pollution. Consequently, the establishment of a water–air–soil multi-feature collaborative warning model enhances prediction outcomes.

The impact of characteristic factors on collaborative warning was further evaluated by defining *Influe*, where *featuresMAE* represents the single-feature model performance (MAE) and *coordinatingMAE* represents the collaborative feature model performance (MAE).

$$\text{Influe} = \frac{featureMAE - coordinatingMAE}{coordinatingMAE}. \tag{2}$$

The performances of different models in predicting atmospheric pollutants, water pollutants, and solid pollutants are compared in terms of MAE in Table 3. Among them, influe_air (%), influe_water (%), and influe_soil (%) represent the degree to which the sample prediction accuracy of the water–air–soil collaborative model has been improved in training compared to the air, water, and soil feature models, respectively.

The improvements of the water, air, and soil feature models were averaged to determine the model performance. The degree of impact is shown in Fig. 4.

Table 3. Prediction and improvement performance of different pollutant features

	influe_air (%)	influe_water (%)	influe_soil (%)
DT	8.09	26.79	54.14
KRR	5.89	–10.25	13.90
lasso	1.24	0.63	3.45
SVR	3.81	0.75	4.17
XGB	3.44	–4.20	4.70
GBM	10.86	9.01	23.4
ENet	2.64	0.00	3.24
Stack_1	17.27	33.96	52.98
Stack_2	1.73	–0.78	6.19
Stack_3	2.35	0.03	4.12
Stack_4	3.99	0.78	6.09

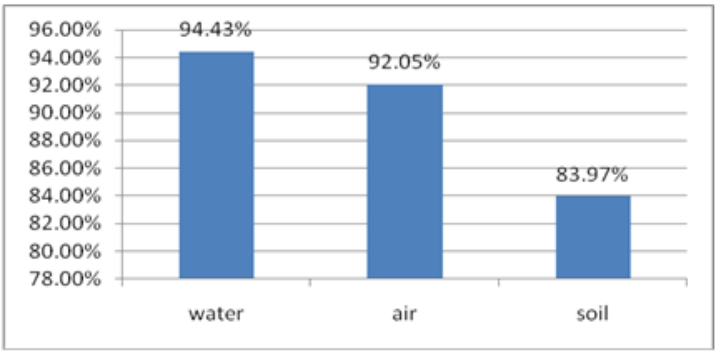

Fig. 4. Performance comparison of different feature factors in the model.

Water–air–soil multi-feature model yields better prediction results compared with single pollutant emission characteristics. The degree of influence of the analyzed features were in the following order from highest to lowest: air > water > soil.

The performance of the ensemble model is determined based on the goodness of fit (R^2). This metric reflects the degree of fitting between the regression line and the observed value. The closer the R^2 value is to 1, the better the fit.

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (\bar{y} - \hat{y}_i)^2}. \tag{3}$$

Fig. 5 shows the prediction performance of multiple features used in the proposed warning model and the single-feature (water, air, soil). Stack_2 performed well and Stack_4 performed poorly. The developed warning model with water–air–soil multi-features exhibited better prediction performance than the single pollutant emission characteristic model.

The impact of selection strategies for primary and secondary learners on different ensemble learning outcomes was discussed in the previous section. The proposed model performed effectively on the fixed dataset and on other data, indicating its robustness and enhanced generalization ability. Therefore, the dataset was further divided using different strategies, and the performance of Stack_2 was considered as the benchmark. In Strategy 1, the dataset was divided into two datasets with high and low temperatures. In Strategy 2, the dataset was divided into two datasets with high and low wind speeds.

Table 4. Improvement in MAE of Stack_2 for different temperatures and wind speeds

	High_temperature (%)	Low_temperature (%)	High_wind_speed (%)	Low_wind_speed (%)
DT	65.36	38.32	55.36	60.88
KRR	24.43	18.79	3.86	27.18
lasso	21.46	–3.52	–0.07	8.93
SVR	63.31	29.48	49.14	33.34
XGB	78.62	31.26	26.49	25.74
GBM	56.53	31.51	50.89	39.64
ENet	87.15	78.26	79.00	80.54

As shown in Table 4, the prediction bias (MAE) of Stack_2 improved, except for lasso. This indicates that the ensemble learning model in Stack_2 can exhibit better prediction performance in most scenarios than most base models.

4. Conclusion and Enlightenment

In this study, various learning principles and observation discrepancies of machine learning algorithms were analyzed in depth. Then, information fusion technology was leveraged for regional air pollution forecasting. By leveraging the advantages of each machine learning model, a multi-model fusion strategy was developed. Then, the influence of pollution emissions from coastal enterprises in Liaoning Province on the AQI was determined. The selection and fusion of multiple models and integration strategy employed for developing the collaborative early warning framework for coastal air environments were discussed. The proposed framework incorporates dynamic, data-driven water–air–soil multi-feature collaboration and accurately forecasting the environmental impact trends of emissions. Compared with those of a single-feature prediction model, the multi-feature multi-model framework exhibited higher MAE and R^2 values, as well as enhanced stability and fusion reliability. This framework was developed to solve coastal ecological and environmental problems, as well as improve the effectiveness and accuracy of environmental monitoring and governance. It offers a well-grounded forecast of the influence trajectory of emissions on the ambient air quality, thereby helping government agencies with a data-driven interpretation of mitigation measures. Ultimately, it helps in alleviating coastal ecological and environmental pollution and devise strategies for ensuring environmental quality management. The proposed system addresses the following critical issues:

1) It furnishes a solid data-based rationale for the regulation of emissions from coastal enterprises, thereby enhancing the effectiveness of pollution source control measures. Unlike traditional methods employed for calculating the AQI, the correlation between multi-feature pollutant emissions from coastal enterprises and environmental quality in surrounding areas was explored herein by employing big data fusion analysis. Based on the varying degrees of impact of these emissions, a list of key regulatory enterprises and key pollutant discharge outlets can be proposed. Drawing upon the outcomes of coastal environmental quality, this study advocates for the targeted regulation of corporate pollution discharges. It also quantifies the respective contributions of these pollution sources to the environmental quality, leveraging the degree of impact exerted by water, air, and soil pollutant features. It thus guides the formulation of scientific and efficient preventive governance strategies aimed at overseeing the discharge outlets of various enterprises situated along coastal regions.

2) The collaborative early warning model captures coastal air pollution dynamics and exhibits enhanced prediction accuracy. By adopting this model, scientific and effective responses can be obtained

to address severe pollution events in coastal environments. In instances where moderate to severe pollution warnings are issued for coastal areas, it signifies that the concentration of pollutants emanating from nearby enterprises may be elevated. Under such conditions, coupled with meteorological factors such as high temperatures and low humidity, the model underscores the importance of strengthening safety production management within the region to prevent the escalation of chain reactions such as marine pollution accidents. Furthermore, the predictive models facilitate the identification of potential hazards, thereby offering novel perspectives for early warning and dynamic collaborative governance of coastal ecological environment pollution incidents.

Conflict of Interest

The authors declare that they have no competing interests.

Funding

This research was supported by the Basic Research Project of Liaoning Provincial Department of Education (No. LJ112410158015) and the Basic Research Funds Project of Dalian Ocean University in 2024 (No. 2024JBYBR003).

References

[1] Q. Wu and H. Lin, "Daily urban air quality index forecasting based on variational mode decomposition, sample entropy and LSTM neural network," *Sustainable Cities and Society*, vol. 50, article no. 101657, 2019. https://doi.org/10.1016/j.scs.2019.101657

[2] B. Wang, Q. Jiang, and P. Jiang, "A combined forecasting structure based on the L1 norm: application to the air quality," *Journal of Environmental Management*, vol. 246, pp. 299-313, 2019. https://doi.org/10.1016/j.jenvman.2019.05.124

[3] J. Z. Wang and W. D. Yang, "Air quality early warning system based on nonlinear correction strategy," *Systems Engineering-Theory & Practice*, vol. 39, no. 8, pp. 2138-2151, 2019. https://doi.org/10.12011/1000-6788-2018-2470-14

[4] N. Song, Q. Wang, J. Huang, and X. Huang, "Comparative study on marine environmental management in major coastal countries in Northeast Asia," *Environmental Management in China*, vol. 11, no. 6, pp. 16-22, 2019. https://doi.org/10.16868/j.cnki.1674-6252.2019.06.016

[5] X. Li, P. Sun, Z. Wang, and Y. Li, "Dynamic comprehensive evaluation of environmental quality in coastal cities," *Environmental Engineering*, vol. 33, no. 10, pp. 108-112, 2015. https://doi.org/10.13205/j.hjgc.201510024

[6] Z. H. Zhou, *Machine Learning*, Beijing, China: Tsinghua University Press, 2016. pp. 171-196.

[7] C. Sun, M. Ma, Z. Zhao, and X. Chen, "Sparse deep stacking network for fault diagnosis of motor," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 7, pp. 3261-3270, 2018. https://doi.org/10.1109/TII.2018.2819674

[8] R. F. Alvear-Sandoval and A. R. Figueiras-Vidal, "On building ensembles of stacked denoising auto-encoding classifiers and their further improvement," *Information Fusion*, vol. 39, pp. 41-52, 2018. https://doi.org/10.1016/j.inffus.2017.03.008

[9] N. Demir and G. Dalkiliç, "Modified stacking ensemble approach to detect network intrusion," *Turkish Journal of Electrical Engineering and Computer Sciences*, vol. 26, no. 1, pp. 418-433, 2018. https://doi.org/10.3906/elk-1702-279

[10] Y. Guo, S. Liu, Z. Li, and X. Shang, "BCDForest: a boosting cascade deep forest model towards the classification of cancer subtypes based on gene expression data," *BMC Bioinformatics*, vol. 19, article no. 118, 2018. https://doi.org/10.1186/s12859-018-2095-4

[11] A. Khamparia, D. Gupta, N. G. Nguyen, A. Khanna, B. Pandey, and P. Tiwari, "Sound classification using convolutional neural network and tensor deep stacking network," *IEEE Access*, vol. 7, pp. 7717-7727, 2019. https://doi.org/10.1109/ACCESS.2018.2888882

[12] D. T. Ho, V. C. Engel, S. Ferron, B. Hickman, J. Choi, and J. W. Harvey, "On factors influencing air-water gas exchange in emergent wetlands," *Journal of Geophysical Research: Biogeosciences*, vol. 123, no. 1, pp. 178-192, 2018. https://doi.org/10.1002/2017JG004299

[13] X. Yu and W. Guan, "Research on the characteristics and influencing factors of air pollution in coastal industrial environment," in *2024 the 8th International Conference on Energy and Environmental Science (ICEES 2024)*. Cham, Switzerland: Springer, 2024, pp. 15-29. https://doi.org/10.1007/978-3-031-63901-2_2

[14] F. Tucca, T. Luarte, J. Nimptsch, S. Woelfl, K. Pozo, G. Casas, et al., "Sources and diffusive air–water exchange of polycyclic aromatic hydrocarbons in an oligotrophic North–Patagonian lake," *Science of the Total Environment*, vol. 738, article no. 139838, 2020. https://doi.org/10.1016/j.scitotenv.2020.139838

[15] L. J. Saunders, J. L. Rodriguez-Gil, S. S. Stoyanovich, L. E. Kimpe, M. L. Hanson, B. Hollebone, D. M. Orihel, and J. M. Blais, "Effect of spilled diluted bitumen on chemical air-water exchange in boreal lake limnocorrals," *Chemosphere*, vol. 291(Part 1), article no. 132708, 2022. https://doi.org/10.1016/j.chemosphere.2021.132708

[16] Y. A. Mahayskiy, A. A. Pavlov, and O. Miseckaite, "Crops water consumption and vertical soil moisture exchange," *AGROFOR International Journal*, vol. 6, no. 1, pp. 57-64, 2021. https://doi.org/10.7251/AGRENG2101057M

[17] Y. Luo, Q. Li, K. Yang, W. Xie, X. Zhou, C. Shang, Y. Xu, Y. Zhang, and C. Zhang, "Thermodynamic analysis of air-ground and water-ground energy exchange process in urban space at micro scale," *Science of the Total Environment*, vol. 694, article no. 133612, 2019. https://doi.org/10.1016/j.scitotenv.2019.133612

[18] K. M. Ting and I. H. Witten, "Issues in stacked generalization," *Journal of Artificial Intelligence Research*, vol. 10, pp. 271-289, 1999. https://doi.org/10.1613/jair.594

[19] X. Wang, X. Yu, and T. Wang, "Air pollution impact prediction of chemical industry park based on ensemble learning strategy," *Operations Research and Management Science*, vol. 30, no. 11, pp. 127, 2021. http://www.jorms.net/EN/10.12005/orms.2021.0360

Xiuli Yu **https://orcid.org/0000-0002-0128-1890**

She received B.S. degree in computer science from Dalian Ocean University in 2003, M.S. degree in management science from Dalian Maritime University in 2006, and Ph.D. degree in management science from Dalian University of Technology, in 2020. She is currently an associate professor in the Department of Human Resource Management, Dalian Ocean University, Liaoning, China. Her research interests include data mining and emergency management.

Peng Du **https://orcid.org/0009-0005-0645-6919**

She received B.S. degree in international economics and trade from Harbin University of Science and Technology in 2004, M.S. degree in technical economy and management from Harbin University of Science in 2007, and Ph.D. degree in management science from Harbin University of Science and Technology, in 2010. She is currently an associate professor in the Department of Human Resource Management, Dalian Ocean University, Liaoning, China. Her research interests include data mining and marine development and management.