Dempster-Shafer ensemble learning framework for air pollution nowcasting

Trung H. Le[1,*], *Huynh A.D.* Nguyen[1], *Quang P.* Ha[1,**], *Minh Q.* Tran[1], *Masrur* Ahmed[2], *Jing* Kong[2], *Xavier* Barthelemy[2], *Hiep* Duc[2], *Ningbo* Jiang[2], *Merched* Azzi[2], and *Matthew* Riley[2]

[1]Faculty of Engineering and IT, University of Technology Sydney, Australia
[2]Department of Climate Change, Energy, the Environment and Water, New South Wales, Australia

Abstract. Deep-learning has emerged as a powerful approach to significantly improve forecast accuracy for air quality estimation. Several models have been developed, demonstrating their own merits in some scenarios and for certain pollutants. In nowcasting, the prediction of air pollution over a small time period essentially demands accurate and reliable estimates, especially in the event cases. From these, selecting the most suitable model to achieve the required forecast performance remains challenging. This paper presents an ensemble framework based on the Dempster-Shafer theory for data fusion to identify the most accurate and reliable forecasts of air pollution obtained from multiple deep neural network models. Our framework is evaluated against three popular machine learning methods, namely, LightGBM, Random Forest, and XGBoost. Experiments are conducted on two horizons: 6-hour and 12-hour predictions using real-world air quality data collected from state-run monitoring stations and low-cost wireless sensor networks.

1 Introduction

Access to a reliable source of multi-hazard early warnings can save lives [1]. Nowcasting, which provides weather predictions up to a short period of time ahead, is one of the integral components in such systems, facilitating timely detection of natural hazards, as highlighted by the United Nations [2]. Not just supporting long-term sustainability in cities, nowcasting can also significantly contribute to microclimate management, providing local information for public dissemination and, in cases of extreme events, community-specific dispatch of responsive measures. Due to the stochastic and volatile nature of the atmospheric environment, achieving forecast accuracy in short time horizons is a challenging task, as a particular model for it is often suitable for different spatial and temporal scenarios [3]. Therefore, researchers have recently explored various deep learning techniques to effectively predict weather events [4], especially to blend those learning models for the best forecast performance.

For model ensembling, the Romanian National Meteorological Administration developed NowDeepN [5], where neural network-based deep learning (DL) algorithms are merged to increase the capability of accurately forecasting heavy precipitation and hail by leveraging radar data. To this end, data fusion from multiple remote sensing sources collected by the National Oceanic and Atmospheric Administration and the China Meteorological Administration has contributed to the development of NowcastNet [6], integrating the physics-informed layer with neural models. In Spain, researchers have implemented an ensemble DL framework to address the issue of fog-related accidents, demonstrating advantages over individual models in nowcasting fog events [7].

To improve the highly-demanded accuracy in real-time forecasting time-series, as of air quality, it requires a versatile and robust method that can all integrate advantages of each existing technique to efficiently handle uncertainty, random fluctuations and nonlinearity in various scenarios [8]. In this context, the Dempster-Shafer evidence theory (DSET) has proven to be an effective data fusion approach, where it can amalgamate predictions from an ensemble of learning-based models to enhance the forecast performance, e.g., precipitation classification and rainfall estimation in Algeria [9]. In another research, DSET was employed to blend Random Forest (RF) and Support Vector Machine models for flood susceptibility forecasts to improve the ultimate accuracy in the face of a multitude of conflicting flood conditioning factors [10]. The ground for adopting DSET in nowcasting could stem from its notable ability to assimilate environmental parameters and forecasts. Given the heterogeneity of urban air quality, where a certain model may respond differently to localized ambient conditions, especially within a small prediction horizon, there is a clear need for a reliable tool for merging the forecasts at a decision level, such as those Dempster-Shafer-based (DS) models.

In this paper, we propose a dynamic ensemble framework for model selection based on DS data fusion of various models and estimation sources for air pollution nowcasting in urban areas. As a result, the most accurate nowcast among member models is selected based on certain performance metrics. Through implementation on real-world air quality datasets and extensive statistical analy-

*e-mail: Hoang.T.Le@student.uts.edu.au
**e-mail: Quang.Ha@uts.edu.au

sis, the proposed framework demonstrates its feasibility and effectiveness in estimating air quality in some Sydney suburbs. The contributions of this work include (i) an architecture of the ensemble framework for air quality forecasting equipped with a DS mechanism to yield the most accurate forecasts from models and monitoring network observations, (ii) the DS-based algorithm developed based on performance metrics for dynamic selection of the best learning-based models, and (iii) nowcasting results for air pollution in some Sydney suburbs benchmarked with state-run monitoring stations.

The remainder of the paper is organized as follows. Section 2 introduces the DL models integrated into the proposed framework and the development of the ensemble model architecture as well as the DS-based algorithm for dynamic selection of best-performing predictions based on multi-metrics criteria. Following that, we present the nowcasting results of the proposed framework on the urban air quality monitoring network in Section 3 and discuss its overall performance. A conclusion is drawn in Section 4.

2 Dempster-Shafer-based ensemble nowcasting framework

2.1 Deep learning models

An ensemble model combines multiple machine learning (ML) or DL models to improve predictive performance. Each model, or "learner," captures distinct patterns in the data, making the ensemble more capable of handling different scenarios [11]. This approach ensures robustness and reliability by leveraging the strengths of individual models to adapt to varying temporal dynamics. This adaptability is essential for accurate air quality nowcasting in urbanized areas, where many localized environmental factors, such as traffic flows, household emissions from energy usage and interactions between changing weather conditions, govern the shift in pollutant levels.

In the proposed framework, each DL model chosen as a member learner exhibits its own strengths in predicting certain pollutant patterns from the input data. The following learners are incorporated into the ensemble architecture and their configurations are tabulated in Table 1:

- **1D-CNN (Convolutional Neural Network)**: A robust convolutional neural network for capturing short-term temporal patterns.
- **Artificial Neural Network (ANN)**: A general network can model all types of data including time series.
- **Long Short-Term Memory (LSTM)**: A DL recurrent neural network that captures long-term patterns of time series.
- **Gated Recurrent Unit (GRU)**: A type of lightweight LSTM with less number of internal gates.
- **Bidirectional LSTM (BiLSTM)**: A type of LSTM network with learning capacity in two directions.
- **Convolutional LSTM (CNN-LSTM)**: A hybrid DL model with robustness of spatial-temporal learning.

Table 1: Member learners of ensemble learning framework

Member learners	Configuration
1D-CNN	Conv1D(64) - MaxPooling - Dense(64)
ANN	Dense(100, 2 layers)
LSTM	LSTM(128) - Dense(64)
GRU	GRU(128) - Dense(64)
BiLSTM	BiLSTM(128) - Dense(64)
CNN-LSTM	Conv1D(64) - LSTM(128) - Dense(64)

The selection, after tuning, of hyperparameters for each DL model in the ensemble learning framework, is presented in Table 2. These hyperparameters are shared across all member learners, ensuring consistency in the training and prediction processes while preserving each model's particular suitability in handling temporal distributions.

Table 2: Hyperparameters for ensemble learning framework

Hyperparameters	Values and types
Input layer (historical data)	12
Output layer (prediction horizon)	6-12
Epoch	50
Batch size	512
Learning rate	0.001
Patience (Early stopping)	5
Loss function	Mean squared error
Optimizer	Adam

2.2 DS framework for ensemble learning

Multiple strategies for constructing decision levels in ensemble learning have recently been established [12]. Given the advantages and disadvantages of each nowcasting model, and in the face of air pollution volatility in urban conditions, for instance, particulate matter (PM) originating from anthropogenic sources, leveraging the forecast from all models may still encounter inaccuracy in the final estimation. As such, an ensemble model that employs a dynamic model selection mechanism to utilize the capabilities of each learner offers the adaptability in real-time for various environmental conditions, and thus, can improve forecast accuracy [13]. For this, the Dempster-Shafer-based ensemble learning framework (DSEL) is designed to enhance overall predictive performance by dynamically selecting the most reliable forecast from the diversity of DL learners to meet the nowcasting requirements.

The architecture of DSEL, depicted in Fig. 1, demonstrates the flow of information from individual nowcasts produced by member learners to the DS selection process for determining the best-performing model. Going in-depth, key components of this architecture include the calculation of performance indices between each learner's predictions and real observations from monitoring instruments. The DS algorithm, which is defined in 2.3, receives those metrics and then computes the similarity level for each model. Finally, the predictions that align most closely with real-world measurements are made the concluding nowcasting results of DSEL.

2.3 Dynamic model selection mechanism

In this section, we elaborate on the model selection approach based on DSET at the decision-making unit of the ensemble learning framework. The frame of discernment (FoD) in DSET lays the foundation for the ensemble model selection mechanism. FoD declares a finite set of all possible hypotheses of the concerned problem, which are the individual learners here. We denote the FoD as the set

$$H = \{p_j, j = 1, \ldots, N_e\}, \tag{1}$$

where N_e is the total number of nowcasting estimators, and p_j denotes the probability of model j achieving predictions as similar as real observations.

For evaluation, we consider the performance-indicative metrics, where each provides different insights into the relation of prediction results with real observations. They are listed in the following,

- The Root Mean Squared Error (RMSE):

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}, \tag{2}$$

- The Mean Absolute Error (MAE):

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|, \tag{3}$$

- Pearson correlation coefficient (r):

$$r = \frac{\sum_{i=1}^{n}(y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}\sqrt{\sum_{i=1}^{n}(\hat{y}_i - \bar{\hat{y}})^2}}, \tag{4}$$

where n is the prediction horizon, y_i, $\hat{y}_i$, $\bar{y}$, and $\bar{\hat{y}}$ are, respectively, the real observations from monitoring instruments, the ensemble nowcasting results, the mean of measurements and the mean of predicted values.

The metrics RMSE, MAE, and correlation coefficient r are popular in DL time-series forecasting. However, the selection of a best-performing learner based on a single metric may overlook other indications, leading to some bias. To identify the best-fit prediction from an individual learner, DSET is tasked with the fusion of multiple evaluation metrics to reach a balance by using a multi-criteria selection process. Here, the metrics are considered as sources of evidence for the development of the basic probability assignment (BPA) in DSET. Quantifying the degree of support for the similarity between forecast values and observations, a BPA for each member learner is assigned based on the evidence provided by the model's performance indices. It is essential to note that the mass, or probability, associated with a certain hypothesis in the defined FoD must be in the range between 0 and 1. Therefore, mathematically, the mass function, denoted as m, is portrayed as

$$m : 2^H \to [0, 1]\,,$$

where it must satisfy

$$\begin{cases} m(\emptyset) = 0, \\ \sum_{p\subseteq H} m(p) = 1. \end{cases} \tag{5}$$

For computation, we first formulate a reference matrix that represents the FoD and BPAs. The reference matrix

$$R = [r_{jk}],\ k = 1, 2, ...K, \tag{6}$$

where K is the number of evaluation metrics used in the multi-criteria analysis. Since we are concerned with the resemblance of forecasts to observational values while considering multiple metrics, the sampling vector $S = [s_k]$ is derived from the assumption of exact alignment of forecasts to observations where the evaluation statistics are ideal. In other words, the sampling vector containing the optimal values of the metrics where they indicate the best performance.

In compliance with the conditions (5), we took the first step in deriving BPAs by the quantification of the similarity between nowcasting values and the ideal scenario of exact fit with measured values from monitoring devices, which is expressed as

$$d_{jk} = |s_{jk} - r_{jk}|. \tag{7}$$

The multi-metrics-fused similarity probability associated with an individual learner j is calculated as

$$q_{jk} = \frac{d_{jk}^{-1}}{\sum_{k=1}^{K} d_{jk}^{-1}}, \tag{8}$$

which establish the probability matrix $Q = [q_{jk}]$.

The statistical values of the evaluation metrics are transformed into the degree of forecasts aligning with actual observations in the matrix Q. However, the entries in each column only represent the supporting degree from one metric. Here, the Dempster's rule of combination in DSET is imperative in combining evidence from multiple metrics to derive a balanced and inclusive confidence level of a learner closely matching observations. If evidence from different metrics is represented by a mass function m, then given any two mass functions m_1 and m_2, the Dempster's rule aggregates them into a joint belief m_{12} by using the orthogonal sum ($\oplus$). Thus, we have

$$m_{12}(B \cap C) = m_1(B) \oplus m_2(C) = \frac{\sum m_1(B)\, m_2(C)}{1 - K_c}, \tag{9}$$

if $B \cap C \neq \emptyset$, and $m_{12}(\emptyset) = 0$, in which

$$K_c = \sum_{B\cap C=\emptyset} m_1(B)\, m_2(C)\,. \tag{10}$$

where K_c represents the level of conflict between pair-wise BPAs.

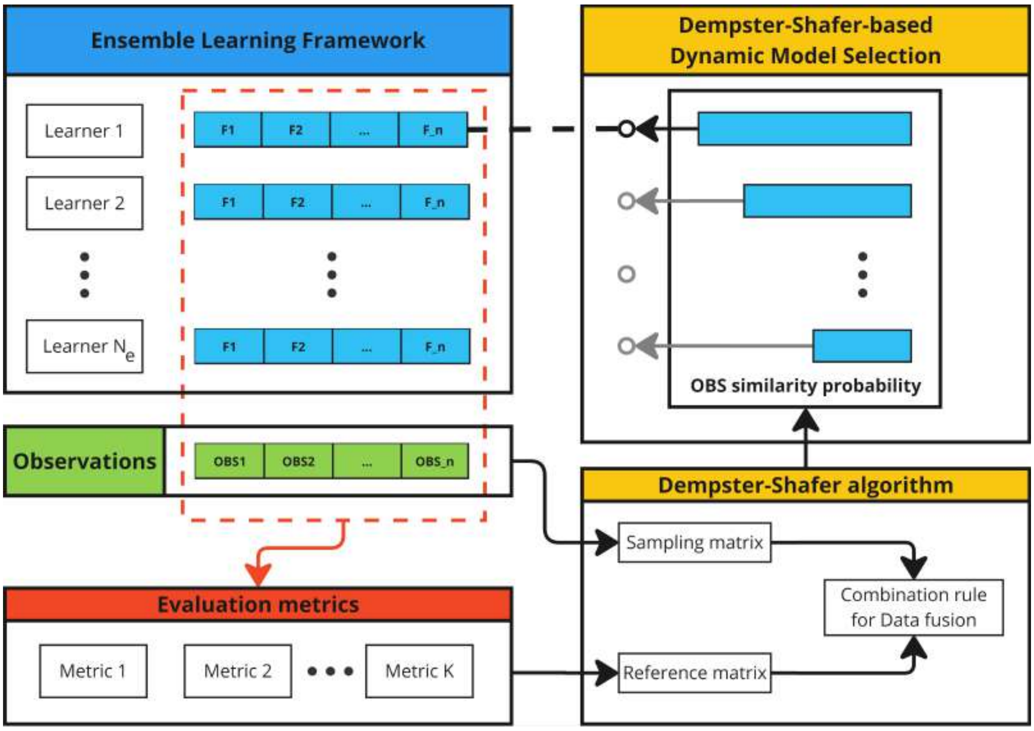


Figure 1: Dempster-Shafer ensemble learning framework for nowcasting: the architecture

The pseudocode presented in Algorithm 1 summarizes for the DS-based model selection mechanism and describes the step-by-step procedure for dynamically selecting the best model at each nowcasting interval. By incorporating DSET in the decision-making layer of DSEL, the proposed ensemble learning framework gains the necessary flexibility to continuously varying input data to improve the accuracy required for nowcasting.

3 DS ensemble learning implementation for urban air quality nowcasting

3.1 Multi-scale air quality monitoring network

Rapid urban growth in smart cities necessitates microclimate management to address challenges posed by climate change and safeguard public health. Multi-scale environmental monitoring combining reference-grade air quality monitoring stations (AQMS) and low-cost air quality sensors becomes more and more popular, e.g., in Europe [14] and Australia [15]. To this end, our proposed DSEL framework has been implemented to perform along such networks in urban settings within the city of Sydney, New South Wales, Australia. Our sites of interest, illustrated in 2, include the monitoring station located in Liverpool suburb at the coordinates -33.93132°S, 150.90727°E and PurpleAir sensors (PAS) located in Lidcombe suburb at coordinates -33.88143°S, 151.04676°E. These data sources are under the management of the local authority, enabling a quality-assured, continuous and real-time information stream.

We retrieved the data of Liverpool AQMS through the publicly available online portal of the local environmental agency. This dataset spans the period from January 1, 2018 to September 30, 2023. As for the PAS in Lidcombe, we collected the data from March 1, 2021 to June 30, 2022. The focal point of both datasets is hourly $PM_{2.5}$ concentrations, the primary target in this study, serving for the training and evaluation of the proposed DSEL framework against other ensemble models for cross-comparison purposes. The training and nowcasting of $PM_{2.5}$ for all ensemble models were performed on an Interactive High Performance Computing (iHPC) server equipped with the NVIDIA A2 GPU to ensure uniformity in the learning process and support rigorous evaluation afterward.

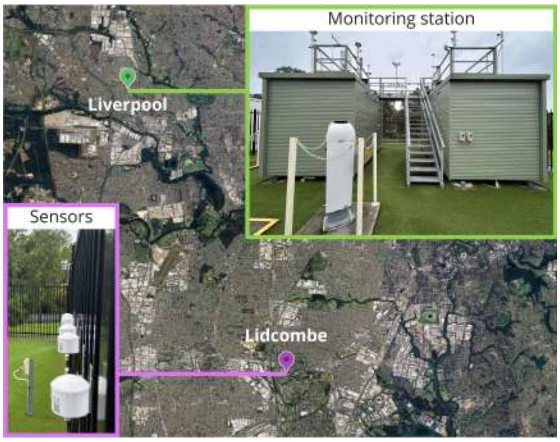


Figure 2: Surveyed urban air quality monitoring instruments in Sydney

Algorithm 1 Dynamic Dempster-Shafer-based model selection for ensemble learning

function STATISTICS($ensemble\_nowcast$, obs)
  for $k = 1 : K$ **do**
    Compute evaluation metrics $stats_k$ (Eq. 2 3, 4)
    Compute intermediate mass $inter\_mass[j]$ from pair-wise hypotheses (Eq. 9)
  end for
end function
function MATRIX_FORMATION($ensemble\_nowcast$, obs)
  for $j = 1 : N_e$ **do**
    $stats \leftarrow$ STATISTICS($ensemble\_nowcast_j$, obs)
    $ref\_mat[j] \leftarrow stats$
  end for
  $samp\_mat$ contains ideal statistics
end function
function COMBINED_MASS($discounted\_mass$)
  for $j = 1 : N_e$ **do**
    Compute conflict coefficient K_c (Eq. 10)
    Compute intermediate mass $inter\_mass[j]$ from pair-wise faults (Eq. 9)
  end for
  $combined\_mass \leftarrow inter\_mass[N_e]$
end function
function DSET_ALGORITHM(ref_mat, $samp\_mat$)
  Compute distances (Eq. 7)
  Compute probabilities (Eq. 8)
  $similarity\_prob \leftarrow$ COMBINED_MASS($ensemble\_nowcast_j$, obs)
end function
function MODEL_SELECTION($similarity\_prob$)
  if $max(similarity\_prob) == similarity\_prob_j$ **then**
    $best\_model \leftarrow j$
  end if
  return best-fit nowcast $ensemble\_nowcast_{best\_model}$
end function

3.2 Air pollution nowcasting with DS ensemble learning: Results and Evaluation

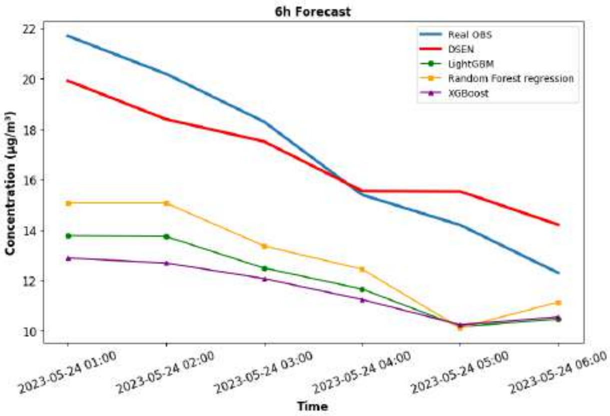
To demonstrate the application of our proposed framework on real-world datasets, we present the final results of DSEL in the form of 6-hour nowcasting and 12-hour very short-term forecasting for the reference-grade station in Liverpool and PAS in Lidcombe. Both horizons provide necessary information for quick responses and early warning in local communities [16]. The performance of our proposed framework is compared with different ensemble ML methods LightGBM, RF regression and XGBoost. They were trained on the exact datasets as DSEL. Figure 3 presents the time-series prediction from DSEL and other ensembles plotted alongside real observations from professional-grade and cost-effective instruments. Specifically, Fig. 3a and Fig. 3c show a general trend of transitioning from a *Fair* to *Good* level of $PM_{2.5}$ concentration as categorized by the local authority [17]. While all ensemble models considered here are able to recognize the constant inclination of observations, DSEL presents the ability to trace the magnitude of its target when the others show significant biases in magnitude. The decreasing tendency of $PM_{2.5}$ in the 12-hour forecast accompanies fluctuations in the form of sudden spikes and valleys. Despite the challenges posed, the 12-hour DSEL forecasts can adapt well to such abrupt changes to match observational values at Liverpool station, as depicted in Fig. 3c.

The measurements from low-cost sensors are expected to have sharper edges than their counterparts collected by regional monitoring stations due to inherent differences in hardware and deployment sites. The PAS readings shown in Fig. 3b illustrate the rapid fluctuation of $PM_{2.5}$ values. It is observed that DSEL predictions can identify and follow the general trend of the PAS measurements in spite of its high instability, whereas predictions from other ensemble models struggle to capture the extrema and instead converge toward the overall mean value. Figure 3d portrays a different scenario with a steep and jagged decline of the $PM_{2.5}$ ground-truth. In this case, it is visible that DSEL forecasts can follow the general downward trend from PAS, especially in the middle of the prediction window. At the two ends, the forecast shows some biases compared to observations, likely due to extreme shifts in $PM_{2.5}$ levels. The underperformance of ML ensemble models is exhibited in cases of considerably high concentrations of PM throughout the test datasets. However, at relatively lower concentrations, the distinction is not as significant and in a comparable level with DSEL. To comprehensively evaluate the performance, we tabulated the statistics in Table 3 and 12-hour forecasting in Table 4.

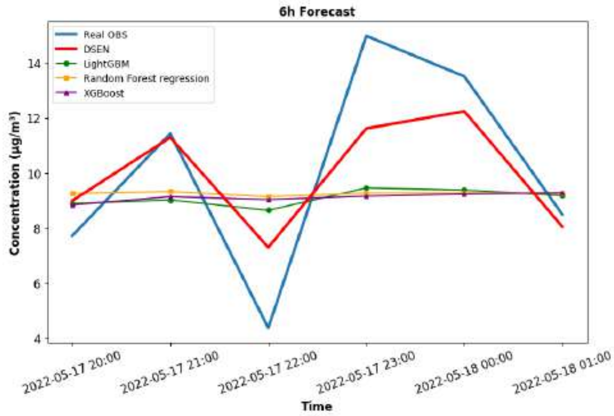
Through statistical analysis, the results for 6-hour horizon produced by DSEL have significantly lower RMSE and MAE compared to other ensemble nowcasting models, meaning the nowcast from DSEL closely matches the magnitude of real observations, with RMSE ranging from 1.44 to 1.97 $\mu g/m^3$ and MAE between 1.29 to 1.57 $\mu g/m^3$ for the AQMS and PAS respectively. The correlation coefficient also indicates the outperformance resulting from the proposed DSEL, which is tabulated as 0.99 for AQMS and 0.93 for PAS. Within the 12-hour prediction horizon, the statistics of AQMS forecasts from different ensemble models are quite comparable. Nevertheless, it is acknowledged that DSEL forecasts offer better accuracy. In contrast, there is a clear distinction between the performance of DSEL in the 12-hour forecast of PAS data compared to other ensemble models. The RMSE and MAE of the 12-hour forecast on PAS data are 1.94 and 1.63 $\mu g/m^3$, within close proximity to their counterparts in the nowcasting, while maintaining a high correlation of 0.92. These results emphasize DSEL's high predictive accuracy even in a longer prediction window, especially in dealing with erratic data patterns of low-cost ambient sensors.

4 Conclusion

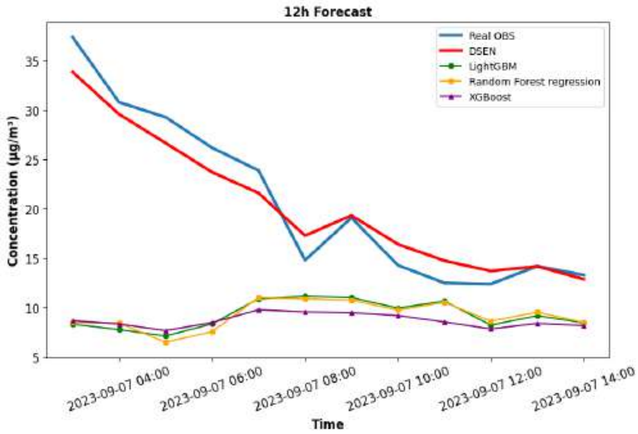
The ability to predict the status of particulate matter concentrations in urban areas for short time horizons is pivotal for supporting early warning systems in a metropolis. This paper presents a Dempster-Shafer-based ensemble learning (DSEL) framework for the multi-criteria model



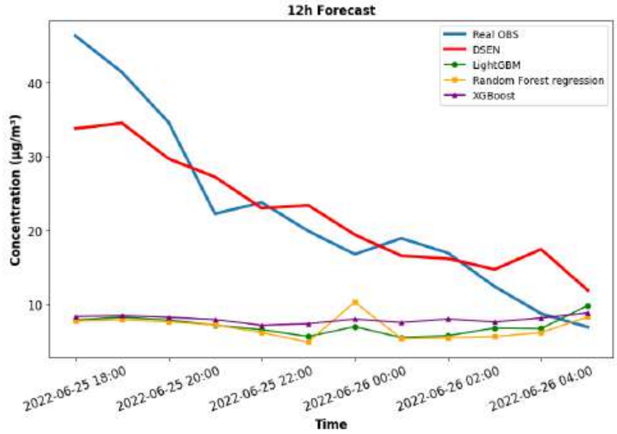
(a) Nowcasting results for AQMS



(b) Nowcasting results for PAS



(c) 12-hour forecasts for AQMS



(d) 12-hour forecasts for PAS

Figure 3: Comparison of real observations from monitoring instruments against the proposed DSEL, LightGBM, Random Forest regression and XGBoost.

Table 3: Performance comparison between different ensemble models on nowcasting of $PM_{2.5}$

Prediction horizon (hr)	Data source	Model	RMSE ($\mu g/m^3$)	MAE ($\mu g/m^3$)	Pearson's r
6	AQMS	DSEL	**1.442**	**1.294**	**0.990**
		LightGBM	4.084	3.662	0.817
		RF regression	4.207	3.605	0.799
		XGBoost	4.312	3.872	0.769
	PAS	DSEL	**1.972**	**1.570**	**0.935**
		LightGBM	3.503	3.037	0.916
		RF regression	3.671	3.193	0.706
		XGBoost	3.677	3.162	0.486

selection mechanism to achieve the best forecast performance. The predictions obtained from real-world datasets suggested that the proposed DSEL is capable of tracking the overall propensity observed by monitoring instruments and outperforms other ensemble ML models, particularly in the presence of high concentrations of the pollutant. The adaptability and accuracy of the framework consolidate its suitability in volatile environments. For future work, we plan to extend it to accommodate multivariate nowcasting and broaden the outputs to other key pollutants that may cause critical impacts on the environment and human health. The ultimate aim is to improve the computational efficiency and scalability of DSEL across the multi-scale monitoring network to provide fine-grained air quality nowcasting to the community level and comprehensively facilitate the informative decision-making process and timely responses from the stakeholders.

Table 4: Performance comparison between different ensemble models on very short-term forecasts of $PM_{2.5}$

Prediction horizon (hr)	**Data source**	**Model**	**RMSE ($\mu g/m^3$)**	**MAE ($\mu g/m^3$)**	**Pearson's r**
12	AQMS	DSEL	**3.778**	**3.030**	**0.960**
		LightGBM	3.908	3.362	0.921
		RF regression	4.702	4.064	0.958
		XGBoost	4.055	3.525	0.934
	PAS	DSEL	**1.943**	**1.637**	**0.924**
		LightGBM	2.974	2.337	0.840
		RF regression	2.845	2.202	0.839
		XGBoost	3.185	2.630	0.917

References

[1] C. Demir, S. Kim, Early Warnings Save Lives and Support Long-Term Sustainability, accessed: Sep. 19, 2024, https://www.undp.org/eurasia/blog/early-warnings-save-lives

[2] W.M.O. (WMO), EARLY WARNINGS FOR ALL, accessed: Sep. 19, 2024, https://library.wmo.int/idurl/4/58209

[3] H.A.D. Nguyen, T.H. Le, M. Azzi, Q.P. Ha, Monitoring and estimation of urban emissions with low-cost sensor networks and deep learning, Ecological Informatics **82**, 102750 (2024). 10.1016/j.ecoinf.2024.102750

[4] Q. Liao, M. Zhu, L. Wu, X. Pan, X. Tang, Z. Wang, Deep Learning for Air Quality Forecasts: A Review, Current Pollution Reports **6**, 399 (2020). 10.1007/s40726-020-00159-z

[5] G. Czibula, A. Mihai, E. Mihuleţ, NowDeepN: An Ensemble of Deep Learning Models for Weather Nowcasting Based on Radar Products' Values Prediction, Applied Sciences **11**, 125 (2021). 10.3390/app11010125

[6] Y. Zhang, M. Long, K. Chen, L. Xing, R. Jin, M.I. Jordan, J. Wang, Skilful nowcasting of extreme precipitation with NowcastNet, Nature **619**, 526 (2023). 10.1038/s41586-023-06184-4

[7] C. Peláez-Rodríguez, J. Pérez-Aracil, A. de Lopez-Diz, C. Casanova-Mateo, D. Fister, S. Jiménez-Fernández, S. Salcedo-Sanz, Deep learning ensembles for accurate fog-related low-visibility events forecasting, Neurocomputing **549**, 126435 (2023). 10.1016/j.neucom.2023.126435

[8] B. Kang, P. Zhang, Z. Gao, G. Chhipi-Shrestha, K. Hewage, R. Sadiq, Environmental assessment under uncertainty using Dempster–Shafer theory and Z-numbers, Journal of Ambient Intelligence and Humanized Computing **11**, 2041 (2020). 10.1007/s12652-019-01228-y

[9] F. Belmahdi, M. Lazri, F. Ouallouche, K. Labadi, R. Absi, S. Ameur, Application of Dempster-Shafer theory for optimization of precipitation classification and estimation results from remote sensing data using machine learning, Remote Sensing Applications: Society and Environment **29**, 100906 (2023). 10.1016/j.rsase.2022.100906

[10] T. Gudiyangada Nachappa, S. Tavakkoli Piralilou, K. Gholamnia, O. Ghorbanzadeh, O. Rahmati, T. Blaschke, Flood susceptibility mapping with machine learning, multi-criteria decision analysis and ensemble using Dempster Shafer Theory, Journal of Hydrology **590**, 125275 (2020). 10.1016/j.jhydrol.2020.125275

[11] J. Wang, G. Song, A Deep Spatial-Temporal Ensemble Model for Air Quality Prediction, Neurocomputing **314**, 198 (2018). 10.1016/j.neucom.2018.06.049

[12] Y. Yang, H. Lv, N. Chen, A Survey on ensemble learning under the era of deep learning, Artificial Intelligence Review **56**, 5545 (2023). 10.1007/s10462-022-10283-5

[13] J.F. de Oliveira, E.G. Silva, P.S. de Mattos Neto, A Hybrid System Based on Dynamic Selection for Time Series Forecasting, IEEE Transactions on Neural Networks and Learning Systems **33**, 3251 (2022). 10.1109/TNNLS.2021.3051384

[14] M. Penza, D. Suriano, V. Pfister, M. Prato, G. Cassano, Urban Air Quality Monitoring with Networked Low-Cost Sensor-Systems, Proceedings **1**, 573 (2017). 10.3390/proceedings1040573

[15] T.H. Le, H.A.D. Nguyen, X. Barthelemy, T.T. Nguyen, Q.P. Ha, N. Jiang, H. Duc, M. Azzi, M. Riley, Visualization Platform for Multi-Scale Air Pollution Monitoring and Forecast, in *2024 IEEE/SICE International Symposium on System Integration (SII)* (2024), pp. 01–06

[16] Y. Wang, E. Coning, A. Harou, W. Jacobs, P. Joe, L. Nikitina, R. Roberts, J. Wang, J. Wilson, A. Atencia et al., Guidelines for Nowcasting Techniques (WMO, 2017), ISBN 978-92-63-11198-2

[17] New South Wales Government, About the air quality categories, accessed: Sep. 27, 2024, https://www.environment.nsw.gov.au/topics/air/understanding-air-quality-data/air-quality-categories