sustainability

MDPI

Article

A Novel Hybrid Approach: Integrating Bayesian SPDE and Deep Learning for Enhanced Spatiotemporal Modeling of $PM_{2.5}$ Concentrations in Urban Airsheds for Sustainable Climate Action and Public Health

Daniel Patrick Johnson [1,*], Niranjan Ravi [2], Gabriel Filippelli [3] and Asrah Heintzelman [4]

[1] Department of Geography, Indiana University, Indianapolis, IN 46202, USA
[2] Department of Electrical and Computer Engineering, Indiana University, Indianapolis, IN 46202, USA; ravin@iu.edu
[3] Department of Earth and Environmental Science, Indiana University, Indianapolis, IN 46202, USA; gfilippe@iupui.edu
[4] Department of Earth Sciences, Indiana University, Indianapolis, IN 46202, USA; asrah@milestoneventuresinc.com
* Correspondence: dpjohnso@iu.edu

Abstract: This study introduces a novel hybrid model combining Bayesian Stochastic Partial Differential Equations (SPDE) with deep learning, specifically Convolutional Neural Networks (CNN) and Deep Feedforward Neural Networks (DFFNN), to predict $PM_{2.5}$ concentrations. Traditional models often fail to account for non-linear relationships and complex spatial dependencies, critical in urban settings. By integrating SPDE's spatial-temporal structure with neural networks' capacity for non-linearity, our model significantly outperforms standalone methods. Accurately predicting air pollution supports sustainable public health strategies and targeted interventions, which are critical for mitigating the adverse health effects of $PM_{2.5}$, particularly in urban areas heavily impacted by climate change. The hybrid model was applied to the Pleasant Run Airshed in Indianapolis, Indiana, utilizing a comprehensive dataset that included $PM_{2.5}$ sensor data, meteorological variables, and land-use information. By combining SPDE's ability to model spatial-temporal structures with the adaptive power of neural networks, the model achieved a high level of predictive accuracy, significantly outperforming standalone methods. Additionally, the model's interpretability was enhanced through the use of SHAP (Shapley Additive Explanations) values, which provided insights into the contribution of each variable to the model's predictions. This framework holds the potential for improving air quality monitoring and supports more targeted public health interventions and policy-making efforts.

Keywords: PM2.5 concentrations; air quality modeling; Bayesian SPDE; deep learning; Convolutional Neural Networks (CNN); Deep Feedforward Neural Networks (DFFNN); hybrid models

check for updates

Citation: Johnson, D.P.; Ravi, N.; Filippelli, G.; Heintzelman, A. A Novel Hybrid Approach: Integrating Bayesian SPDE and Deep Learning for Enhanced Spatiotemporal Modeling of $PM_{2.5}$ Concentrations in Urban Airsheds for Sustainable Climate Action and Public Health. *Sustainability* **2024**, *16*, 10206. https://doi.org/10.3390/su162310206

Academic Editors: O. V. Giannico, Roberta Zupo and Maria Lisa Clodoveo

Received: 23 October 2024
Revised: 8 November 2024
Accepted: 11 November 2024
Published: 22 November 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The critical issue of air quality and its direct impact on public health has helped drive significant interest in advanced methods for environmental monitoring. Particulate matter smaller than 2.5 μm ($PM_{2.5}$), known for its ability to penetrate deep into the lungs and enter the bloodstream, is particularly concerning due to its links to a wide array of health problems, from respiratory and cardiovascular diseases to neurological disorders and adverse pregnancy outcomes. The complex health impacts of $PM_{2.5}$, which vary across spatial and temporal dimensions, necessitate advanced analytical approaches to accurately model its distribution and contributing factors. Our research responds to this need by providing a comprehensive modeling framework that enhances understanding of $PM_{2.5}$ exposure and supports more effective public health management. Moreover, accurate

predictions of air pollution concentrations are critical for advancing global sustainability, aligning this work with Sustainable Development Goals (SDG) 3 (health and well-being) and SDG 13 (climate action).

Despite progress in air quality modeling, existing studies often fail to address the non-linear interactions and spatial dependencies inherent in $PM_{2.5}$ data, particularly in urban environments. Traditional geostatistical models can handle spatial dependencies but struggle with non-linear interactions, while machine learning models capture non-linearities but often neglect spatial correlations. To bridge this gap, we propose a hybrid approach that combines Bayesian Stochastic Partial Differential Equations (SPDE) with Convolutional Neural Networks (CNNs). This innovative method integrates spatial-temporal dependencies and non-linear relationships in $PM_{2.5}$ data, offering greater accuracy and interpretability in highly variable urban airsheds. Our model leverages the strengths of both frameworks, outperforming traditional methods such as kriging and standalone neural networks.

Advancements in geostatistical techniques, particularly the use of Integrated Nested Laplace Approximations (INLA) in conjunction with SPDE, have enhanced our ability to analyze environmental data with high precision. This framework efficiently handles spatial-temporal variability and enables detailed exploration of the complex interplay between environmental pollutants and health impacts. By leveraging SPDE's spatial modeling capabilities, we can generate insights essential for effective environmental health surveillance and policy interventions.

The choice of Bayesian Stochastic Partial Differential Equations (SPDE) and deep learning (CNN and DFFNN) models leverages each method's unique strengths to address the complex spatial and non-linear dynamics of $PM_{2.5}$ in urban environments. Bayesian SPDE offers a rigorous spatial framework, capturing spatial dependencies through the Matérn field and enabling detailed exploration of spatial-temporal variability in $PM_{2.5}$ concentrations [1,2]. By integrating deep learning, specifically CNN and DFFNN, our model can address non-linear interactions that traditional geostatistical methods may overlook [3,4]. This combination thus fills a critical gap in the literature, where models often focus on either spatial dependency or non-linearities, but not both.

By fusing Bayesian SPDE and CNNs, we offer a novel methodology that surpasses the limitations of traditional spatial trend modeling and deep learning approaches. Our model effectively integrates spatial structure and non-linear dynamics, enhancing prediction accuracy and offering a more granular approach to environmental data analysis. The insights generated by this framework hold significant implications for air quality management, providing refined tools for monitoring, prediction, and policymaking. This paper details the methodologies, implementation, and broader implications of this integrated approach, establishing a foundation for more advanced environmental modeling.

2. Background

2.1. $PM_{2.5}$ and Its Health Effects

Particulate matter smaller than 2.5 micrometers in diameter ($PM_{2.5}$) is a significant public health hazard due to its ability to penetrate deep into the respiratory system and enter the bloodstream [5]. Research has established strong associations between $PM_{2.5}$ exposure and a wide array of adverse health outcomes, including respiratory diseases, cardiovascular conditions, and neurological impairments [5–7]. Chronic exposure to $PM_{2.5}$ has been linked to increased risks of conditions such as asthma, chronic bronchitis, and heart disease [6,7]. Emerging evidence also suggests a strong association between $PM_{2.5}$ and adverse pregnancy outcomes, such as low birth weight and preterm births [8,9]. Additionally, neurodegenerative diseases and metabolic disorders are becoming focal points in research exploring the long-term effects of $PM_{2.5}$ exposure [10,11].

Despite the extensive knowledge of $PM_{2.5}$'s health impacts, accurately modeling its distribution remains challenging due to the complex interactions between environmental factors, including meteorological conditions, land use, and anthropogenic activities. These complexities emphasize the need for advanced models to capture both the spatial and temporal variability of $PM_{2.5}$ concentrations and the non-linear interactions among these covariates. Addressing $PM_{2.5}$ pollution is key to achieving global health equity and combating the adverse impacts of climate change. Public health efforts must align with climate action strategies to reduce air pollution, particularly in urban areas, in order to promote healthier, more sustainable living environments.

2.2. Advances in $PM_{2.5}$ Modeling

Recent advancements in air quality modeling have greatly enhanced our ability to predict $PM_{2.5}$ concentrations. These improvements are mainly due to the integration of diverse data sources such as satellite imagery, land use data, and meteorological variables, combined with sophisticated statistical and machine learning techniques [12,13]. For example, mobile monitoring combined with neural networks has been shown to provide fine-grained predictions of air pollution [3,14].

While traditional geostatistical methods like kriging have been widely used, they often struggle with the non-linearities and spatial heterogeneities inherent in $PM_{2.5}$ dynamics [15]. To address this, hybrid models that combine geostatistical approaches with machine learning methods have demonstrated substantial improvements in accuracy [12,16]. These hybrid models are particularly effective at capturing the interactions between environmental factors and the spatial distribution of $PM_{2.5}$ [17].

2.3. Bayesian SPDE in Air Quality Modeling

To overcome the limitations of traditional spatial models, the Stochastic Partial Differential Equation (SPDE) approach, applied within a Bayesian framework, has gained traction as a robust tool for spatial-temporal modeling. This method is particularly valuable in environmental and epidemiological studies, where spatial dependencies and uncertainties must be accounted for [1,18]. The SPDE framework enables the construction of spatially continuous models from discrete data points, offering a more precise alternative to classical geostatistical techniques [19].

The flexibility of the SPDE approach is evident in its ability to model spatial fields with complex spatial-temporal structures. For example, SPDE has been successfully applied to air quality modeling by capturing fine-scale spatial variations in $PM_{2.5}$ concentrations [2]. Moreover, incorporating geographic and meteorological covariates has been shown to enhance the predictive accuracy of SPDE models [18]. However, while SPDE excels in modeling spatial structures, it faces challenges in handling highly non-linear relationships between covariates, particularly in complex urban environments.

2.3.1. Mathematical Framework of SPDE

The SPDE approach represents a spatial field using a Gaussian Random Field (GRF), solving a specific SPDE. A common model within this framework is the Matérn field, widely used for its flexibility and physical interpretability [20,21]. The Matérn covariance function is often employed because it effectively represents spatial dependencies [19]

To solve the SPDE numerically, the finite element method is typically applied, discretizing the spatial domain into a mesh. This computational grid allows for the precise modeling of spatial-temporal correlations in $PM_{2.5}$ data, offering improvements over traditional methods like kriging [1].

2.4. Advances in Neural Network Applications for Air Quality Modeling

Artificial Neural Networks (ANNs) and Convolutional Neural Networks (CNNs) have emerged as powerful tools for addressing the limitations of traditional spatial models, particularly in capturing non-linear relationships in large datasets. ANNs are well-suited for air quality modeling due to their ability to handle high-dimensional inputs and capture complex patterns in meteorological and pollution data [6,14]. CNNs, initially developed for image recognition, have demonstrated significant potential in geospatial data analysis, particularly in urban areas where land-use patterns contribute to air pollution variability [4,22].

While ANNs and CNNs excel at modeling non-linear interactions and identifying localized pollution hotspots, they do not naturally capture broader spatial dependencies, a critical component in air quality modeling. Thus, hybrid models that leverage both the spatial rigor of SPDE and the non-linear modeling capabilities of neural networks are needed.

2.5. Transition to Hybrid Approach

By integrating the strengths of Bayesian SPDE and neural networks, such hybrid models provide a comprehensive solution for air quality modeling. SPDE offers robust spatial-temporal predictions by accounting for spatial autocorrelation and uncertainties, while neural networks excel at capturing the non-linear interactions between environmental and meteorological factors [17]. The synergy between these methods allows for both local feature extraction and global spatial trend modeling, improving predictive accuracy across multiple spatial scales.

This hybrid approach improves model performance and enhances the interpretability of the results. Such models play a crucial role in mitigating air pollution, a significant factor in climate change, and ensuring public health resilience in the face of environmental stressors. By incorporating SPDE outputs into a neural network framework, the model offers a more transparent understanding of the spatial and environmental factors driving $PM_{2.5}$ concentrations. Additionally, the use of SHAP analysis further enhances the transparency of neural networks, allowing for actionable insights in public health and urban planning.

Dimensionality Benefits of the Hybrid Approach

One key benefit of the hybrid approach is its ability to navigate the high-dimensional nature of air quality data, which typically integrates satellite imagery, land use data, and sensor networks. The combined use of SPDE and neural networks allows for a comprehensive analysis of localized and broader-scale factors influencing air quality [2]. This model is particularly well-suited for capturing urban pollution hotspots while accounting for larger-scale spatial trends, such as regional wind patterns and topography.

3. Methodology

3.1. Spatial Domain

The Pleasant Run Airshed (PRAS), located in Indianapolis, IN (Marion County), serves as the spatial domain for this study. With a diverse mix of land covers, PRAS predominantly consists of developed land, where medium-intensity development accounts for approximately 32% of the area. Overall, 94.79% of the airshed is characterized by various levels of development, with smaller portions allocated to cultivated crops (1.64%) and deciduous forests (1.49%). Despite this classification, urban tree canopies are underreported due to spatial resolution limitations in the Landsat imagery (30 m resolution), which complicates the accurate representation of mixed land use areas.

Figures 1 and 2 present PRAS's geographic location and land cover distribution, respectively, with $PM_{2.5}$ sensor locations highlighted to depict the monitoring network across the airshed.

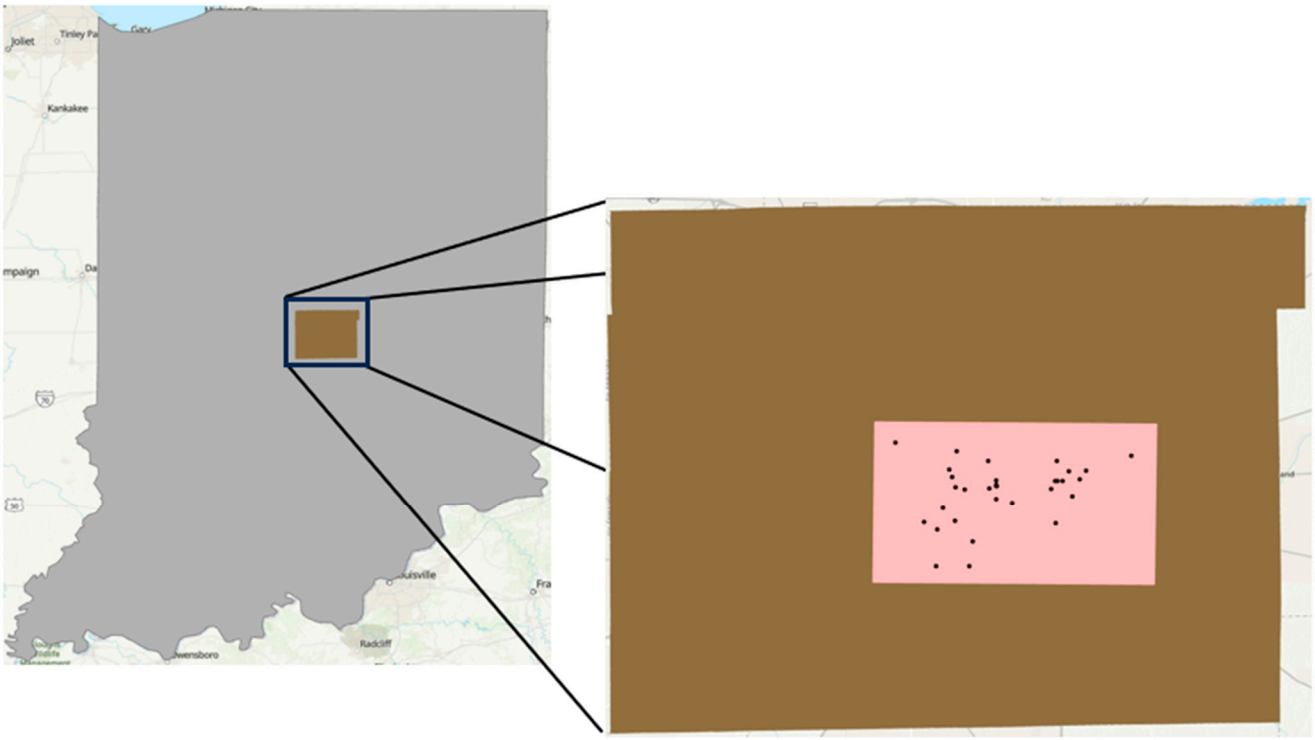


Figure 1. Pleasant Run Airshed (PRAS) is located in Indianapolis, Indiana (Marion County, IN, USA), Black points represent PM sensor locations.

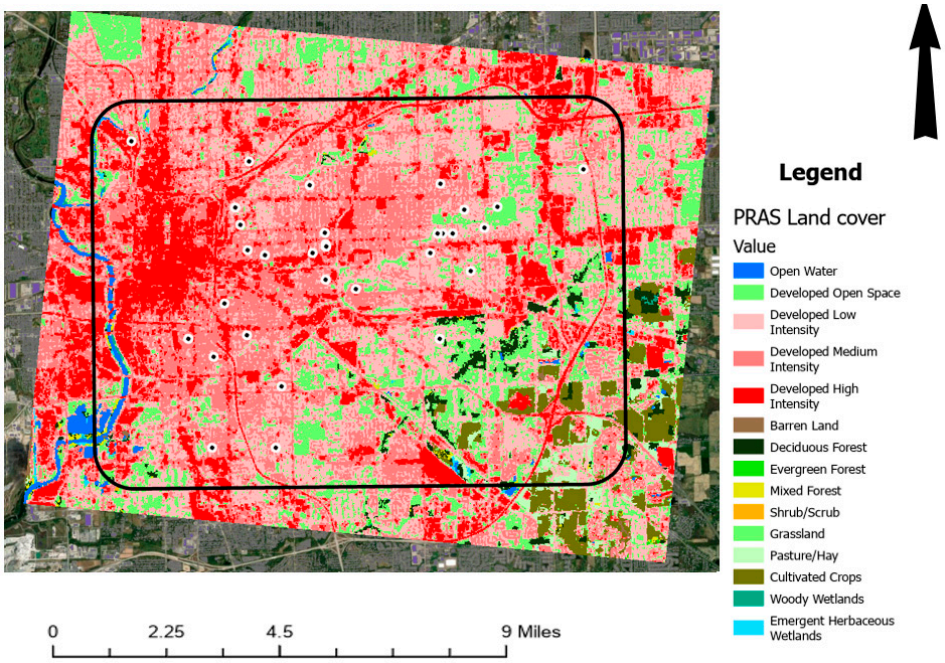


Figure 2. Land cover of the PRAS with $PM_{2.5}$ sensor locations highlighted [23].

3.2. Wind Patterns

Wind direction and speed are critical factors in pollutant dispersion. A wind rose diagram (Figure 3) illustrates the prevailing wind conditions in PRAS during the study period, with winds predominantly coming from the southwest at an average speed of 1.75 meters per second. The inclusion of such data is essential for modeling $PM_{2.5}$ dispersion, as higher wind speeds (2 to 4 meters per second) often facilitate the movement of pollutants across the urban landscape. See Table 1.

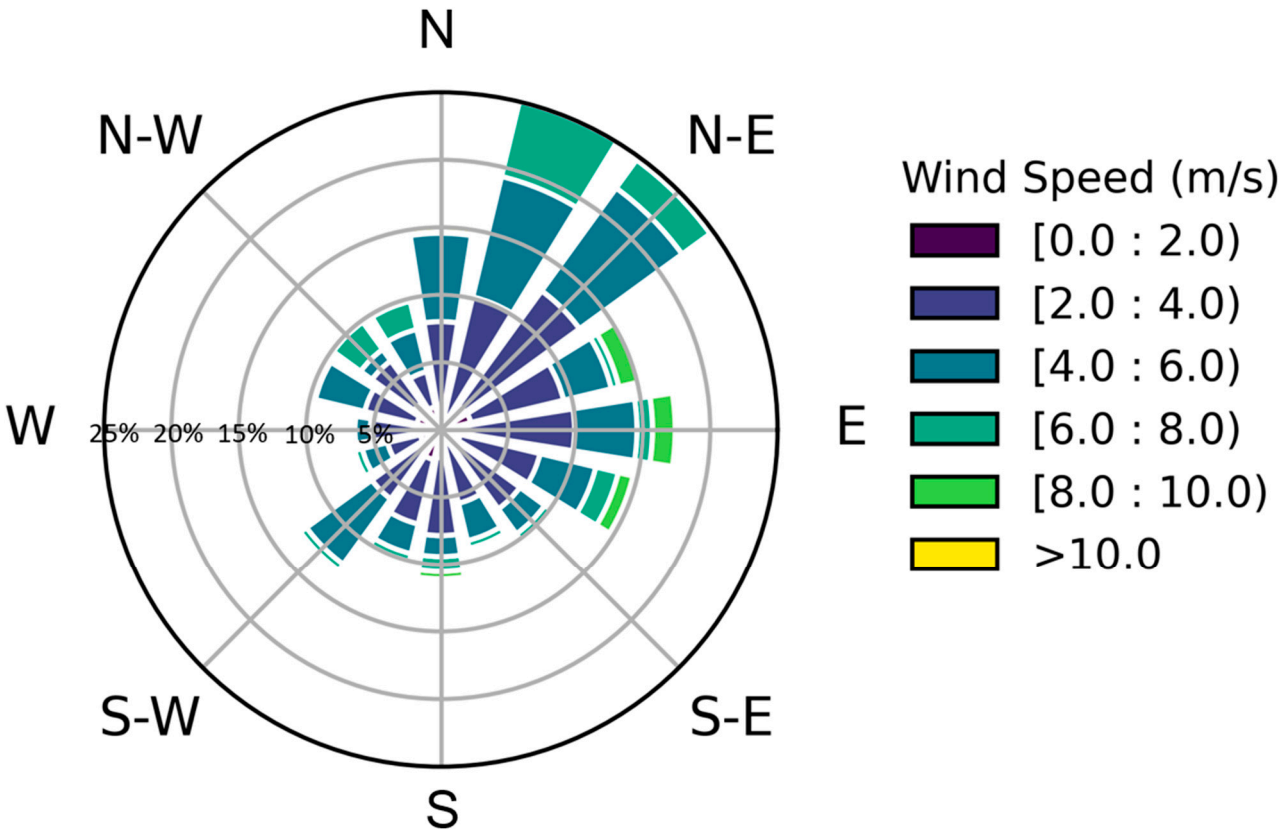


Figure 3. Wind rose of the PRAS wind data during the study period. Data sourced from the North American Land Data Assimilation System (NLDAS) and the MODIS satellite system [24,25].

Table 1. Area and Percent Area of cover types in the Pleasant Run Airshed. Source: National Land Cover Database (NLCD) [23,24].

Land Cover Type	Area in km^2	Percent Area
Open Water	1.05	0.78%
Developed Open Space	16.07	11.90%
Developed Low Intesity	39.96	29.58%
Developed Medium Intensity	46.65	34.54%
Developed High Intensity	25.35	18.77%
Barren Land	0.11	0.08%
Deciduous Forest	2.01	1.49%
Evergreen Forest	0.02	0.01%
Mixed Forest	0.24	0.18%
Shrub/Scrub	0.01	0.01%
Grassland	0.08	0.06%
Pasture/Hay	1.07	0.79%
Cultivated Crops	2.21	1.64%
Woody Wetlands	0.05	0.04%
Emergent Herbaceous Wetlands	0.19	0.14%

3.3. Data Collection

3.3.1. $PM_{2.5}$ Data Collection

A network of 32 citizen scientists was initially established to host continuous PA-II-SD sensors (<1-year-old) with Wi-Fi capabilities, facilitating real-time data collection across the Pleasant Run Airshed (PRAS) [26]. Data collection was centered around a defined area within PRAS, bounded by a minimum bounding rectangle around existing $PM_{2.5}$ sensors, and extended by a 1 km buffer to address potential edge effects. This area was segmented into a 250 m-resolution grid, creating 2160 distinct grid cells to support localized environmental assessments. Sensor readings were aggregated to weekly values and assigned to intersecting grid cells, with meteorological and land cover data processed in parallel to yield grid-based averages and sums (cf: Table 2).

Table 2. Descriptive statistics of variables used for training (grid cells in which a PM sensor was located).

Variable	Mean	Standard Deviation	Minimum	Maximum
$PM_{2.5}$ in $\mu g/m^3$	16.08	5.89	3.69	37.77
Convective Fraction	0.09	0.11	0.00	0.47
Potential Energy (in J/kg)	349.15	492.95	0.00	2048.36
Evaporation (in kg/m^2)	0.18	0.10	0.01	0.33
Longwave Radiation (in W/m^2)	322.00	51.21	233.83	414.57
Precipitation (in kg/m^2)	0.14	0.14	0.00	0.70
Pressure (in Pa)	98,827.24	383.86	97,977.33	99,866.88
Shortwave Radiation (in W/m^2)	178.82	77.80	59.50	316.06
Specific Humidity (Mass fraction)	0.01	0.00	0.00	0.02
Temperature (in C)	12.91	9.98	−6.27	27.34
windu	0.86	1.08	−1.84	2.91
windv	0.67	1.23	−2.29	3.46
AOD 47µm	174.21	101.42	22.67	530.00
AOD 55 µm	121.07	71.95	15.33	373.00
Cropland	0.00	0.00	0.00	0.00
Developed High Intensity (%)	3.60	6.17	0.00	27.45
Developed Medium Intensity (%)	35.71	13.10	10.80	59.91
Developed Low Intensity (%)	35.71	13.10	10.80	59.91
Developed Open Space (%)	6.55	9.69	0.00	28.71
Forest: Deciduous (%)	0.00	0.00	0.00	0.00
Forest: Evergreen (%)	0.00	0.00	0.00	0.00
Forest: Mixed (%)	0.00	0.00	0.00	0.00
Grassland (%)	0.00	0.00	0.00	0.00
Open Water (%)	0.00	0.00	0.00	0.00
Pasture (%)	0.04	0.20	0.00	1.00
Distance to Nearest Road (in m)	17.27	6.55	10.30	34.51

From 1 November 2018 to 31 October 2019, a total of 30 sensors was actively monitoring $PM_{2.5}$ levels. However, due to connectivity issues with some sensors and a manufacturer-issued power cord recall in February 2019, only 23 sensors consistently produced usable data, resulting in a temporary 3-week data gap. Consequently, the finalized dataset included 50 complete weeks of data (see Table 3).

Table 3. Descriptive statistics of covariates after zonal calculations were performed for all grid cells in the study area.

Variable	Mean	Standard Deviation	Minimum	Maximum
Convective Fraction	0.09	0.11	0.00	0.47
Potential Energy (in J/kg)	351.02	493.97	0.00	2048.36
Evaporation (in kg/m^2)	0.18	0.10	0.01	0.33
Longwave Radiation (in W/m^2)	322.18	51.18	233.83	414.57
Precipitation (in kg/m^2)	0.14	0.14	0.00	0.70
Pressure (in Pa)	98,828.70	391.81	97,977.33	99,866.88
Shortwave Radiation (in W/m^2)	179.18	77.75	59.50	316.06
Specific Humidity (Mass fraction)	0.01	0.00	0.00	0.02
Temperature (in C)	12.95	9.96	−6.27	27.34
windu	0.86	1.09	−1.84	2.91
windv	0.67	1.24	−2.29	3.46
AOD 47µm	172.37	99.59	19.00	786.17
AOD 55µm	119.76	70.61	13.00	569.00
Cropland	1.44	6.90	0.00	69.16
Developed High Intensity (%)	9.96	14.64	0.00	69.40
Developed Medium Intensity (%)	24.24	15.26	0.00	67.08
Developed Low Intensity (%)	24.24	15.26	0.00	67.08
Developed Open Space (%)	12.23	14.32	0.00	69.42
Forest: Deciduous (%)	1.04	4.65	0.00	49.95
Forest: Evergreen (%)	0.02	0.27	0.00	5.24
Forest: Mixed (%)	0.15	1.19	0.00	21.92
Grassland (%)	0.03	0.32	0.00	9.65
Open Water (%)	0.61	3.78	0.00	50.65
Pasture (%)	0.64	4.13	0.00	67.12
Distance to Nearest Road (in m)	50.03	57.16	7.85	530.83

While a sample size of 23 sensors per week may appear limited, it reflects the practical constraints inherent in many urban monitoring networks. Coverage limitations, often due to budget or logistical constraints, can lead to sparse sensor networks in real-world urban settings. The design of this study's sensor network mirrors these challenges, making it representative of typical urban air quality monitoring conditions (Requia et al., 2019) [16]. Despite the limited number of sensors, spatial models like Bayesian SPDE are well-suited to such contexts, as they leverage spatial correlations to fill data gaps and improve prediction reliability. Similar studies (e.g., Adams & Kanaroglou, 2016) [12] demonstrate that such models effectively capture spatial patterns even when observational data is sparse.

Furthermore, the hybrid model employed here is designed to address the challenges of limited sensor density by capturing complex relationships between environmental variables. The Bayesian SPDE approach, combined with the model's neural network components, enhances prediction accuracy by integrating spatial and non-linear relationships within the dataset. While future research could benefit from expanded sensor networks or data imputation techniques to improve sensor coverage in under-monitored regions, the current framework provides a robust, real-world representation of urban air quality monitoring, reflecting the feasibility of applying this hybrid approach in diverse urban contexts.

3.3.2. Covariate Data Collection

Covariates such as meteorological data (e.g., wind speed, temperature) and Aerosol Optical Depth (AOD) measurements were obtained from the North American Land Data Assimilation System (NLDAS) and the MODIS satellite system [24,25]. These covariates were aggregated, and zonal statistics were applied to match the grid cells used in the $PM_{2.5}$ data collection, ensuring consistency in the spatial framework.

Tables 2 and 3 provide detailed descriptive statistics for the variables used in the analysis, showcasing the range of environmental and land use factors included.

3.3.3. Training, Validation, and Interpolation Datasets

The first prepared dataset, detailed in Table 2, was divided for model training (80%) and validation (20%). Due to the absence of Forest, Cropland, and Open Water land cover types in sensor-located grid cells, these categories were excluded from the analysis. The dataset comprised 1150 records spanning 21 variables pre-split and post-validation; the interpolation dataset was structured into a long-form format with 20 variables across 108,000 records (50 weeks × 2160 grid cells).

3.4. Model Development

The analysis incorporated several types of spatial data, all processed and managed using R or Python and their relevant library packages (INLA, sf, raster, rgee, keras, tensorflow, shap) [21,27–32]. Typical geospatial processing was performed to load the datasets into R and Python platforms. A spatial join operation was used to merge the sensor point locations with their respective cells of the 250 m grid.

3.4.1. Bayesian SPDE Model Development

There are several steps necessary to facilitate the SPDE modeling. Initially an irregular finite mesh is created over the study area as the foundation for the Matérn field used in the spatial model. The SPDE was specified using the Matérn covariance function, which as shown in Section 2.3.1, is suitable for capturing continuous spatial variation. Parameters such as range κ\kappa$\kappa$, variance τ\tau$\tau$, and smoothness α\alpha$\alpha$ were configured to reflect the physical and statistical properties of the spatial process. To account for temporal dependencies in $PM_{2.5}$ concentrations, a second-order random walk (RW2) was applied, capturing the autocorrelation between consecutive weeks. This ensures that the model doesn't treat each week independently but acknowledges the persistence of air quality conditions over time. The model estimation involved setting up a stacked model configuration for both the estimation and prediction phases. This configuration allowed for handling large datasets efficiently, incorporating spatial dependencies, and executing complex Bayesian hierarchical models within a reasonable computational timeframe. The SPDE models were executed on Indiana University's High-Performance Computing System's BigRed200. The environmental and anthropogenic covariates were included as fixed effects to adjust for potential confounders in the $PM_{2.5}$ concentration data. These covariates were standardized across the study period to ensure their comparability and influence on the model output were appropriately scaled. Given the computational intensity of the INLA-SPDE approach, particularly with a large dataset and complex model structure, optimizations such as the use of PARDISO™ solver version 7.2 were implemented. This solver significantly enhances the computational efficiency of matrix operations inherent in the INLA approach.

3.4.2. DFFNN Model Development

In this study, we implemented a Deep Feedforward Neural Network (DFFNN) due to its ability to handle large, multi-dimensional datasets and capture complex, non-linear relationships between environmental variables and $PM_{2.5}$ concentrations. DFFNNs are particularly well-suited for tasks involving regression on continuous data like air quality, where traditional statistical models may fall short. Their versatility allows them to map intricate interactions between covariates such as temperature, wind speed, and land use, improving predictive accuracy beyond what simpler models like linear regression can achieve. By utilizing multiple hidden layers and non-linear activation functions, DFFNNs enable the model to better capture the subtle and complex influences that various covariates exert on $PM_{2.5}$ levels. This deep learning approach complements our spatial modeling techniques by providing a robust mechanism to learn from diverse environmental inputs, which are often non-linear and highly interdependent. In this study, we implemented a DFFNN using the Keras and TensorFlow libraries in R to predict particulate matter ($PM_{2.5}$) concentrations (our target) from the set of predictors compiled into the training dataset [27,28]. The net-

work architecture was designed to mitigate overfitting using L2 regularization, a common technique that penalizes large weights in the model's parameters [33].

Model Configuration

The DFFNN was initialized as a sequential model, which consists of a linear stack of layers ideal for feedforward architectures. The network's input layer was designed to accept inputs corresponding to the number of predictors. Following the input layer, the first hidden layer contained 64 densely connected neurons using the ReLU (Rectified Linear Unit) activation function. ReLU was selected for its ability to introduce non-linearity without significantly affecting the gradients, promoting a smooth gradient flow during training. This layer also included the aforementioned L2 regularization ($\lambda = 0.01$) to help prevent overfitting by penalizing the magnitude of the coefficients.

The second hidden layer mirrored the first, with 64 neurons, ReLU activation, and L2 regularization, maintaining consistency in the transformation space. The third hidden layer, however, reduced the number of neurons to 32 while keeping the same ReLU activation and regularization strategy. This reduction helped abstract the features into a lower-dimensional representation before the final output. The output layer was densely connected with a single neuron, using a linear activation function, which is typical for regression tasks, to predict the continuous $PM_{2.5}$ concentrations.

Model Compilation

The model was compiled with the mean squared error loss function, which is standard for regression problems as it minimizes the squared differences between the predicted and actual values, thus enhancing the model's accuracy in estimating continuous variables. The Adam optimizer was chosen for its adaptive learning rate capabilities, which make it effective in handling sparse gradients and different parameter scales. The model metrics included mean absolute error, providing an additional performance measure during training and validation.

Hyperparameter Tuning

A systematic hyperparameter tuning process was conducted to optimize the model's performance. This included fine-tuning parameters such as the learning rate, batch size, number of epochs, and regularization strengths (L2 regularization lambda values). Grid search and random search methods were employed to explore a range of values for each hyperparameter. The optimal set of hyperparameters was determined based on cross-validation performance on the training data, ensuring the model's generalization capability.

Training

Training was conducted over 400 epochs with a batch size of 256, incorporating early stopping to monitor validation loss and to further prevent overfitting. We employed a validation split of 20%, which allowed for the continuous evaluation of model performance on a subset of the training data that was not directly used for training, thus providing insights into the model's generalization capabilities. The configuration and training of the deep neural network were carefully designed to balance complexity and performance, aiming to achieve high accuracy in $PM_{2.5}$ prediction while avoiding overfitting through regularization and early stopping strategies.

3.4.3. CNN Model Development

Convolutional Neural Networks (CNNs) were incorporated into our model to enhance their ability to capture spatial hierarchies and localized patterns in $PM_{2.5}$ concentrations across the urban landscape. CNNs are highly effective at processing geospatial data with inherent spatial structures, such as air pollution data influenced by land cover, meteorological conditions, and urban development. Unlike traditional methods, which may treat spatial

points as independent, CNNs leverage local spatial dependencies, making them ideal for identifying pollution hotspots and other localized phenomena in air quality datasets.

The strength of CNNs lies in their ability to recognize spatial features, such as urban-rural differences and the spatial distribution of pollution sources. This improves the model's predictive accuracy for $PM_{2.5}$ concentrations by efficiently capturing both fine-scale spatial variability and broader trends. Furthermore, CNNs excel at handling high-dimensional spatial data, a common characteristic of air quality studies with broad spatial and temporal dependencies.

Originally developed for image processing, CNNs have increasingly been applied to geospatial datasets, given their ability to extract spatial features through convolutional layers. In this study, CNNs were used to capture both localized and abstract patterns by processing input data hierarchically through multiple layers of convolutions and pooling. This approach allows the model to detect intricate patterns and relationships within environmental variables that Bayesian SPDE and DFFNN may overlook. By further leveraging CNNs, our model aligns with recent advancements in machine learning applications in atmospheric and environmental modeling sciences, reflecting a growing trend toward more sophisticated analytical frameworks for interpreting complex datasets.

To utilize our data in the CNN, the dataset was reshaped into a three-dimensional array to meet the input requirements. As with the DFFNN, the dataset was split into training and testing sets, with 80% of the data used for training and the remaining 20% for validation.

Model Configuration

The Convolutional Neural Network (CNN) architecture (Figure 4) was designed with multiple layers to capture local and abstract features from the input data. The initial convolutional layer used 512 filters with a kernel size of 3 and applied the ReLU activation function, serving as the primary feature extractor from the raw data. This was followed by a max pooling layer, which reduced dimensionality while extracting dominant features and decreasing the computational load for subsequent layers.

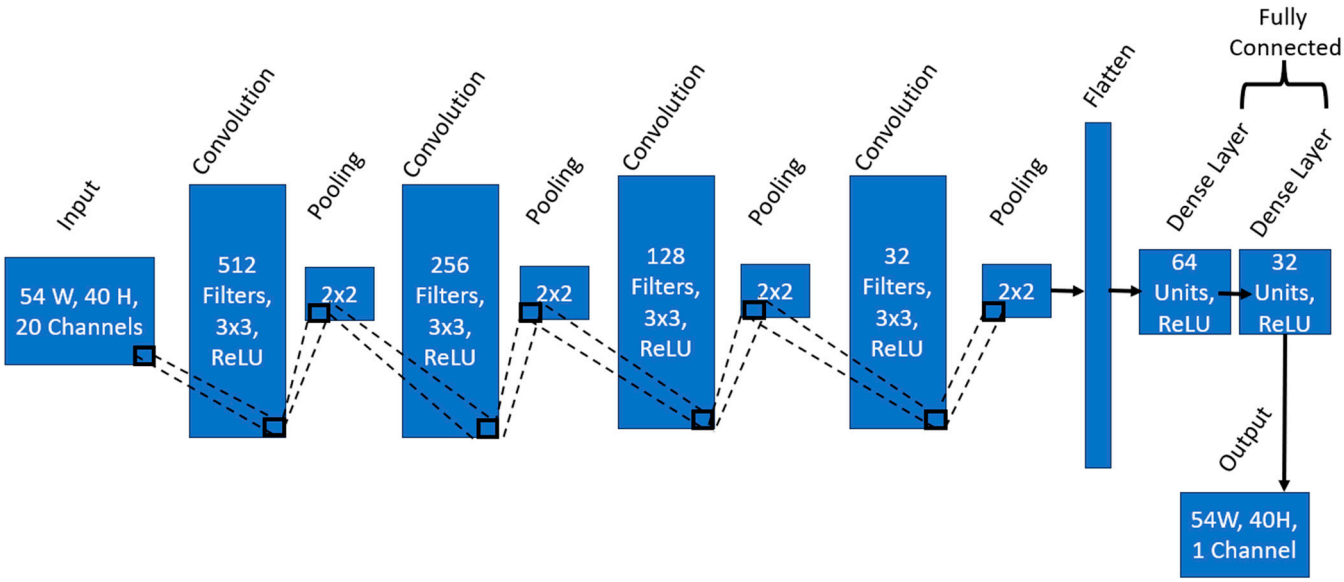


Figure 4. CNN model diagram showing all layers.

Several additional convolutional layers were added, each with diminishing filter sizes (256, 128, and 32), all utilizing the ReLU activation function. These layers progressively added depth to the model, enabling it to capture more complex and abstract patterns within the data. The output from the final pooling layer was then flattened into a one-dimensional array, making it suitable for feeding into the dense layers. The model concluded with a

single dense layer containing one unit, which produced the predicted $PM_{2.5}$ concentration, aligning with the regression nature of the task.

Model Compilation

The model was compiled with the Adam optimizer, mean squared error loss function, and standard regression platform metrics. Training was conducted over 400 epochs with a batch size of 256, incorporating early stopping to monitor validation loss and prevent overfitting.

Hyperparameter Tuning

As with the DFFNN, a comprehensive hyperparameter tuning process was also applied to the CNN model to ensure optimal performance. Key hyperparameters, such as the number of convolutional layers, number of filters, filter sizes, learning rates, and dropout rates, were systematically varied. Techniques, such as grid search and random search, were utilized to identify the most effective hyperparameter settings. The final hyperparameters were selected based on the best cross-validation performance metrics, ensuring the model's robustness and accuracy.

Training

Training was conducted over 400 epochs with a batch size of 256, incorporating early stopping to monitor validation loss and prevent overfitting. We employed a validation split of 20%, which allowed for the continuous evaluation of model performance on a subset of the training data that was not directly used for training, thus providing insights into the model's generalization capabilities. The configuration and training of the convolutional neural network were carefully designed to balance complexity and performance, aiming to achieve high accuracy in $PM_{2.5}$ prediction while avoiding overfitting through regularization and early stopping strategies.

3.5. Model Analysis

Each modeling method, including Bayesian SPDE, DFNN, CNN, and their hybrid configurations, was evaluated using 999 permutations to assess performance in predicting $PM_{2.5}$ concentrations. This extensive simulation was conducted on the high-performance computing (HPC) systems at Indiana University (BigRed200™) to manage the computational demands of these models. The permutation process involved randomizing data inputs and partitioning configurations to mitigate any bias from data order or composition.

Performance metrics—Root Mean Squared Error (RMSE), Relative Root Mean Squared Error (rRMSE), and Mean Absolute Error (MAE)—were collected and aggregated to provide a comprehensive view of each model's accuracy, stability, and reliability across various scenarios. This robust methodology tests the variability in model performance due to stochastic elements and data-handling strategies.

3.5.1. Model Interpretability Using SHAP Values

SHAP (Shapley Additive Explanations) values were used to interpret and explain the predictive model's outputs. SHAP values decompose a model's prediction into contributions from each feature, allowing us to understand how individual factors influence the model's output. Each feature's SHAP value represents its impact on changing the output from the base (average) model prediction to the actual model prediction for a given instance.

SHAP values provide both local and global interpretations. Locally, a SHAP value explains a feature's contribution to a specific prediction, making it a powerful tool for detailed case-by-case analysis. Globally, aggregated SHAP values across all data points offer insights into the overall importance of each feature within the model. SHAP summary plots and mean SHAP value plots help elucidate which features are most influential and how their varying levels affect $PM_{2.5}$ predictions, providing a clear, quantifiable measure

of each feature's impact. By employing SHAP values, we ensure a robust framework for interpretability, helping foster further technical enhancements and policy decisions for effective air quality management to promote environmental sustainability.

3.5.2. Beyond Feature Engineering

This study transcends traditional feature engineering, which typically relies on manually crafted features that can limit a model's ability to capture intricate relationships between variables. Instead, we integrated advanced Bayesian (INLA-SPDE) methods with state-of-the-art Deep Feedforward Neural Networks (DFFNNs) and Convolutional Neural Networks (CNNs) in a hybrid approach. While feature engineering primarily enhances the inputs to machine learning models, our framework creates a dynamic process where the INLA-SPDE captures spatial-temporal structures and feeds these structured outputs into neural networks. This synergy enables the model to detect complex, non-linear interactions, and spatial dependencies that engineered features alone would fail to uncover. Accurate air quality predictions can guide targeted interventions in urban areas, supporting both climate mitigation efforts and public health strategies. In particular, our hybrid model offers insights that could inform policies aimed at reducing pollution exposure, particularly in vulnerable communities facing the dual threats of poor air quality and climate change.

The integration of INLA-SPDE with DFFNNs and CNNs is not merely an exercise in feature creation but a comprehensive framework that enhances the predictive capability and interpretability of environmental models. The INLA-SPDE framework generates structured outputs incorporating spatial and temporal covariates, which serve as dynamic inputs for our DFFNN and CNN. This process goes beyond adding engineered features by embedding rich spatial statistical information into the learning process of neural networks, enabling the detection of intricate patterns that simpler models would likely overlook.

4. Results

4.1. Descriptive Statistical Analysis of $PM_{2.5}$ and Covariates

Throughout our study period, from 1 November 2018 to 31 October 2019, Table 2 details the descriptive statistics for the variables used in model training derived from the grid cells where $PM_{2.5}$ sensors were located. The table summarizes the mean, standard deviation, minimum, and maximum values for $PM_{2.5}$ concentrations alongside various meteorological and land-use covariates. $PM_{2.5}$ levels had an average concentration of 16.08 $\mu g/m^3$, with notable variability, as indicated by a 5.89 $\mu g/m^3$ standard deviation.

The dataset reflects a range of values, with the highest recorded $PM_{2.5}$ level at 37.77 $\mu g/m^3$, capturing typical urban conditions and episodic pollution events. To put these values into context, the World Health Organization's recommended guideline for daily average $PM_{2.5}$ levels is 15 $\mu g/m^3$, while the U.S. Environmental Protection Agency (EPA) sets the National Ambient Air Quality Standard (NAAQS) at 35 $\mu g/m^3$ for 24 h exposure. The elevated levels observed in some weeks of the study exceed these thresholds, highlighting the potential for adverse health effects, particularly for vulnerable populations such as individuals with pre-existing respiratory or cardiovascular conditions.

Meteorological factors, such as potential energy, evaporation, and longwave radiation, showed significant variation, indicative of their dynamic nature and potential influence on $PM_{2.5}$ dispersion. AOD measurements at both 47 μm and 55 μm wavelengths were also included, providing insight into particulate matter's distribution in the atmosphere.

Table 3 extends the analysis with descriptive statistics of covariates' post-zonal calculations for all grid cells within the study area. This comprehensive approach captures the broader environmental context, with variables such as wind speed, temperature, and various land cover types revealing the complex interactions affecting $PM_{2.5}$ levels over the entire airshed.

The average distance to the nearest road was 17.27 m, underlining urban proximity as a potential factor in $PM_{2.5}$ distribution. However, the data also suggest degrees of hetero-

geneity within the urban landscape, as demonstrated by the range of values across different land cover types, from developed high-intensity areas to open water and grasslands.

These descriptive statistics provide a foundation for understanding the multivariate conditions under which the $PM_{2.5}$ sensors operated and highlight the intricate relationships that our hybridized modeling approach aims to simulate.

Our study's weekly averaged $PM_{2.5}$ measurements are depicted in Figure 5, highlighting the variation in $PM_{2.5}$ levels across all sensor locations within the urban airshed. $PM_{2.5}$ concentrations exhibited a discernible downward trend throughout the observation period, with episodic peaks suggesting a temporal variability influenced by environmental or anthropogenic factors. The peak concentration of 37.7 $\mu g/m^3$ was observed in Week 4, while the minimum value of 3.69 $\mu g/m^3$ occurred during Week 6. Notably, the highest $PM_{2.5}$ values were predominantly recorded during the third, fourth, and seventh weeks, aligning with specific environmental conditions or urban activities yet to be fully explored. The end of the study period marked a reduction in both the average $PM_{2.5}$ levels and their variability, indicating a possible seasonal influence on particulate concentrations. However, this aspect was not the focus of the current analysis.

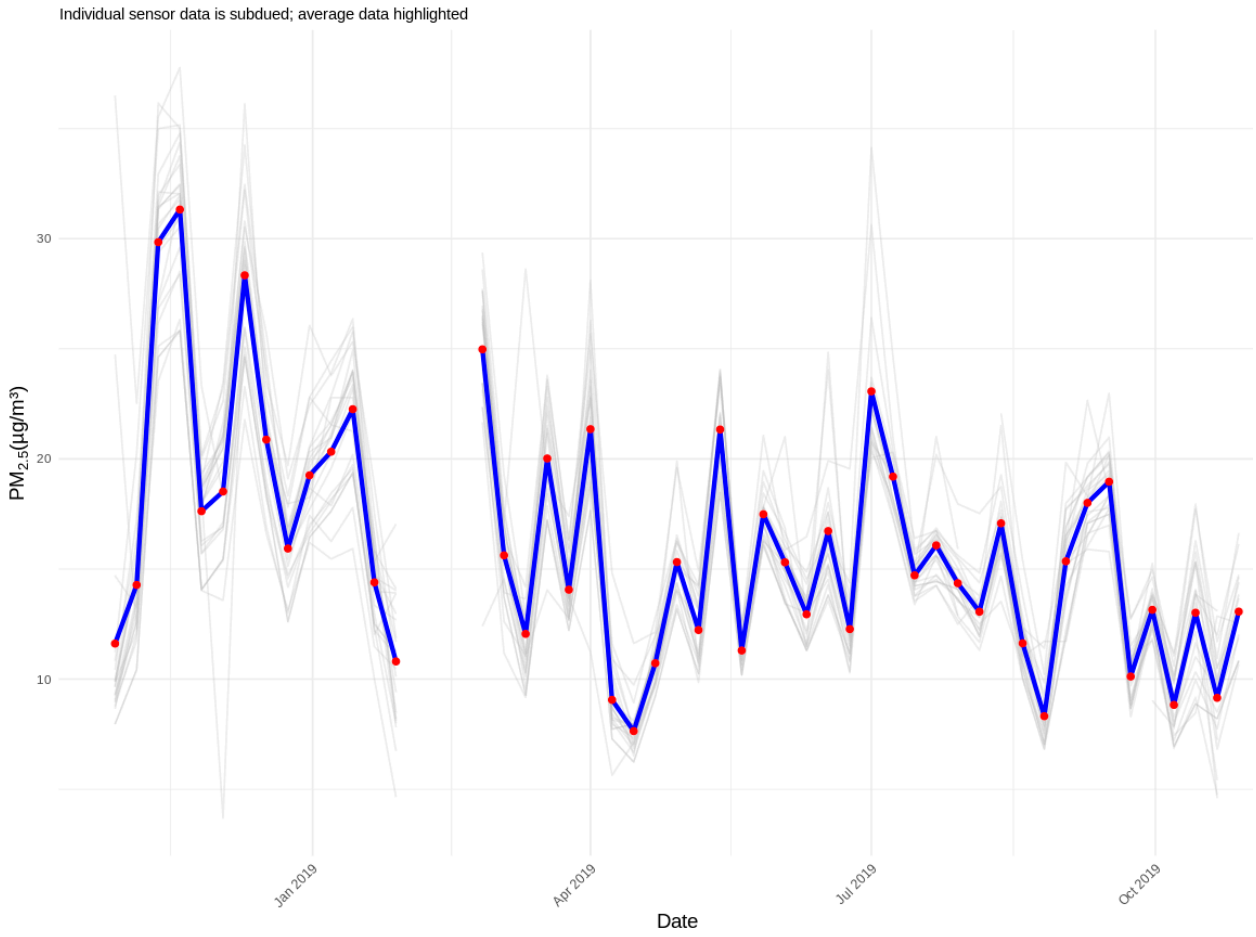


Figure 5. Highlighted average weekly $PM_{2.5}$ levels across all sensors ($\mu g/m^3$) with values at individual sensors subdued. Note: values were missing for most of February due to sensor malfunctions.

4.2. Bayesian SPDE Model and Spatial Structure

Table 4 provides a detailed overview of the mean posterior values for variables included as covariates in the Bayesian SPDE model. Variables statistically significant within the 95% Credibility Envelope are highlighted in bold, underscoring their substantial impact on the spatial distribution and variability of $PM_{2.5}$ concentrations. These results underscore

the intricate relationships between $PM_{2.5}$ levels and various environmental and anthropogenic factors, informing the refined spatial analysis provided by the SPDE framework.

Table 4. Mean posterior values for variables included as covariates in the Bayesian SPDE model. Bold variables are statistically significant within the 95% Credibility Envelope.

Variable	Mean	Standard Deviation	2.5% Cred. Int.	97.5% Cred. Int.
Intercept	−0.09	0.15	−0.36	0.18
Convective Fraction	**0.10**	**0.04**	**0.02**	**0.17**
Potential Energy	**0.53**	**0.08**	**0.38**	**0.69**
Evaporation	**−1.42**	**0.20**	**−1.82**	**−1.03**
Longwave Radiation	**0.56**	**0.22**	**0.12**	**1.00**
Precipitation	**−0.35**	**0.03**	**−0.41**	**−0.29**
Pressure	**0.21**	**0.03**	**0.15**	**0.27**
Shortwave Radiation	**0.49**	**0.14**	**0.21**	**0.77**
Specific Humidity	**−0.88**	**0.24**	**−1.36**	**−0.41**
Temperature	**0.46**	**0.20**	**0.07**	**0.85**
AOD 47 µm	−0.54	1.61	−3.70	2.62
AOD 55 µm	0.79	1.60	−2.36	3.93
Developed High Intensity (%)	−0.02	0.03	−0.09	0.05
Developed Medium Intensity (%)	0.02	22.36	−43.82	43.87
Developed Low Intensity (%)	0.02	22.36	−43.82	43.87
Developed Open Space (%)	−0.09	0.05	−0.18	0.00
Distance to Nearest Road (in m)	0.01	0.04	−0.07	0.10
Wind Speed	**−0.16**	**0.03**	**−0.22**	**−0.10**
Wind Direction	**−0.08**	**0.02**	**−0.13**	**−0.03**

Figure 6 offers a visual representation of the spatial effect as determined by the Bayesian SPDE model within the PRAS. The left side of the figure shows the posterior mean, pinpointing regions where the geographical location exerts a distinct influence on $PM_{2.5}$ concentration estimates. The west-central portion of the PRAS displays a pronounced spatial effect, suggesting the presence of localized factors uniquely affecting $PM_{2.5}$ levels in this area. This area coincides with the developed high-intensity land cover/land use and the Indianapolis central business district.

The right side of the figure, illustrating the posterior standard deviation, provides a measure of the model's uncertainty in estimating these effects. The increased standard deviation at the airshed's periphery indicates less confidence in the model's predictions, suggesting an edge effect. Despite efforts to address this edge effect through a strategic mesh designation, which aims to ensure comprehensive spatial coverage and minimize boundary-related distortions in the model. Furthermore, there is less certainty in areas near the central business district, likely due to a planetary boundary interaction with the heightened surface friction due to the scale of the built environment [30].

Figure 6 offers a visual representation of the spatial effect as determined by the Bayesian SPDE model within the PRAS. The left side of the figure shows the posterior mean, pinpointing regions where geographical location exerts a distinct influence on $PM_{2.5}$ concentration estimates. The west-central portion of the PRAS displays a pronounced spatial effect, suggesting the presence of localized factors uniquely affecting $PM_{2.5}$ levels in this area. This area coincides with the developed high-intensity land cover/land use and the Indianapolis central business district.

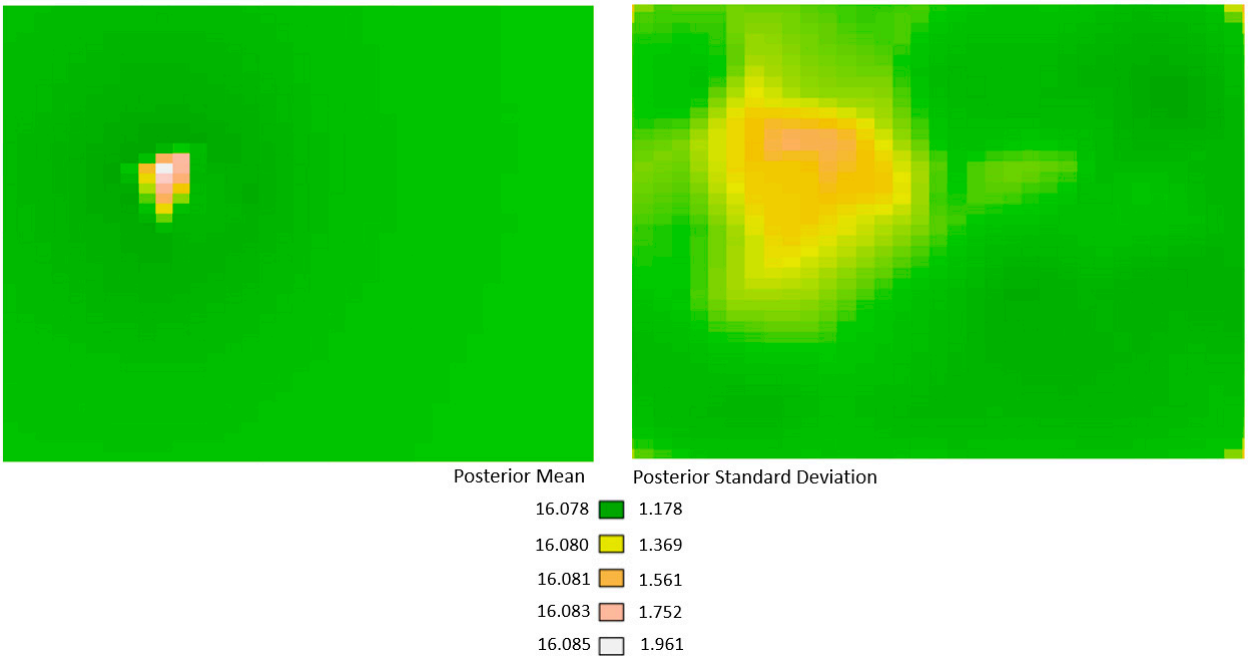


Figure 6. Posterior mean and posterior standard deviation of the spatial random field ($\mu g/m^3$).

The right side of the figure, illustrating the posterior standard deviation, provides a measure of the model's uncertainty in estimating these effects. The increased standard deviation at the airshed's periphery indicates less confidence in the model's predictions, suggesting an edge effect. Despite efforts to address this edge effect through strategic mesh designation, which aims to ensure comprehensive spatial coverage and minimize boundary-related distortions in the model. Furthermore, there is less certainty in areas near the central business district, likely due to planetary boundary interaction with heightened surface friction due to the scale of the built-environment [34].

These insights reveal the model's nuanced ability to discern the complex spatial structure of $PM_{2.5}$ concentrations, highlighting both the efficacy and the limitations of the current modeling approach in capturing environmental variability within urban airsheds.

4.3. Model Performances

Table 5 summarizes the performance metrics of the five modeling architectures used for interpolating $PM_{2.5}$ concentrations after 999 permutations of each model. The evaluation is based on the three different error metrics listed in Section 3.5: Root Mean Squared Error (RMSE), Relative Root Mean Squared Error (rRMSE), and Mean Absolute Error (MAE). Each metric provides a different perspective on model accuracy and precision:

1. Bayesian SPDE Model: This model exhibits moderate variability. Its mean RMSE is 4.68, with values ranging from 4.24 to 4.98. Its mean rRMSE is 0.14, with variations between 0.13 and 0.15. It records a mean MAE of 3.58, spanning from 3.12 to 3.94. This model is suitable for environments where interpretability is as crucial as accuracy.
2. DFFNN (Deep Feedforward Neural Network): Recording the highest errors among the models, the DFFNN shows a mean RMSE of 6.26, ranging from 5.93 to 7.24. The mean rRMSE stands at 0.19, varying from 0.18 to 0.22, and the mean MAE is 5.02, with values between 4.43 and 5.28. These figures suggest less suitability for precise $PM_{2.5}$ prediction tasks under the tested conditions.

3. CNN (Convolutional Neural Network): Demonstrating significantly better accuracy, the CNN achieves a mean RMSE of 3.87, ranging from 3.25 to 4.21. Its mean rRMSE is 0.12, with a span from 0.10 to 0.13, and the mean MAE is 2.55, with figures ranging from 1.98 to 2.76. This model is particularly adept at capturing spatial hierarchies and dependencies in $PM_{2.5}$ data.
4. Bayesian SPDE—DFFNN Hybrid: This hybrid model has a mean RMSE of 4.44, ranging from 4.14 to 5.12. The mean rRMSE is 0.13, varying between 0.12 and 0.13, and it has a mean MAE of 3.23, with figures from 2.90 to 3.45. It outperforms the standalone DFFNN and is somewhat effective, though it does not achieve the performance level of the CNN or its hybrid with Bayesian SPDE.
5. Bayesian SPDE—CNN Hybrid: Outperforming all other models, this hybrid achieves a mean RMSE of 2.92, ranging from 2.67 to 3.99. The mean rRMSE is 0.09, with a span from 0.08 to 0.10, and the mean MAE is 1.59, with values from 1.34 to 2.03. By integrating Bayesian SPDE with CNN, this model harnesses the strengths of stochastic modeling and deep learning to deliver superior prediction accuracy.

Table 5. Model performance metrics measured in modeled $PM_{2.5}$ deviation from $PM_{2.5}$ realization in $\mu g/m^3$. rRMSE is calculated by RMSE/range of $PM_{2.5}$ realization. 999 permutations of each model.

	RMSE			**rRMSE (Unitless)**			**MAE**		
Model	**Mean**	**Min**	**Max**	**Mean**	**Min**	**Max**	**Mean**	**Min**	**Max**
Bayesian SPDE	4.68	4.24	4.98	0.14	0.13	0.15	3.58	3.12	3.94
DFFNN	6.26	5.93	7.24	0.19	0.18	0.22	5.02	4.43	5.28
CNN	3.87	3.25	4.21	0.12	0.10	0.13	2.55	1.98	2.76
SPDE-DFNN Hybrid	4.44	4.14	5.12	0.13	0.12	0.15	3.23	2.90	3.45
SPDE-CNN Hybrid	2.92	2.67	3.99	0.09	0.08	0.12	1.59	1.34	2.03

4.4. Hybridized Bayesian SPDE—CNN Model Outputs

Figure 7 shows a SHAP (Shapley Additive Explanations) summary plot that quantifies each feature's contribution to the Model 5′s output averaged over 999 permutations. Meteorological factors like Shortwave Radiation, Temperature, and Precipitation, along with contextual variables such as Developed Land Intensity and Aerosol Optical Depth (AOD), are ranked by their impact on the model. The broader dispersion of SHAP values for variables like Bayesian Prediction and Shortwave Radiation highlights their significant and varied influence on $PM_{2.5}$ concentration predictions.

While Bayesian Prediction incorporates covariates that are also included in the SHAP analysis, the SHAP framework accounts for multicollinearity by ensuring that feature contributions are not overestimated. The Bayesian prediction variable serves as a composite measure, summarizing spatial-temporal patterns captured by the SPDE framework. SHAP allows us to disentangle the unique contributions of other covariates, such as meteorological conditions, without double-counting their effects.

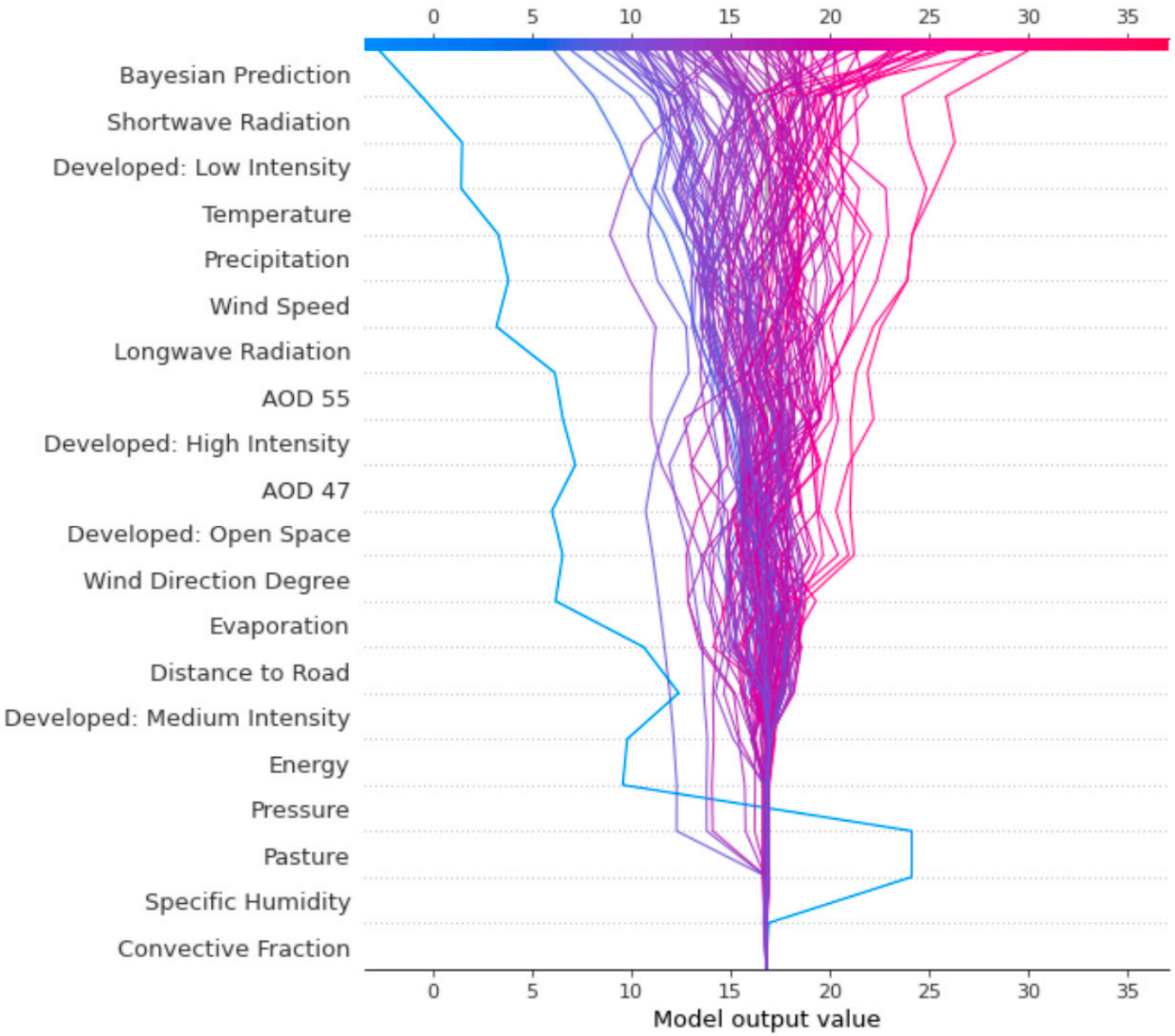


Figure 7. SHAP model output values for each instance (230 from the test set).

Figure 8 provides a violin plot that captures the spread and density of predictions across different variables to explore the distribution of Model 5's outputs. The tails represent output variability, while the thicker sections indicate where most predictions are concentrated. Bayesian Prediction consistently shows the highest median values, reflecting its dominant role in driving the model's results.

Feature contributions, shown in the bar chart in Figure 9, are depicted with average SHAP values (after 999 permutations) for each variable in Model 5. Bayesian Prediction emerges as the most influential factor, with an average impact of +2.27, followed closely by other environmental and meteorological variables. This visualization offers a straightforward comparison of each feature's contribution to the model's accuracy.

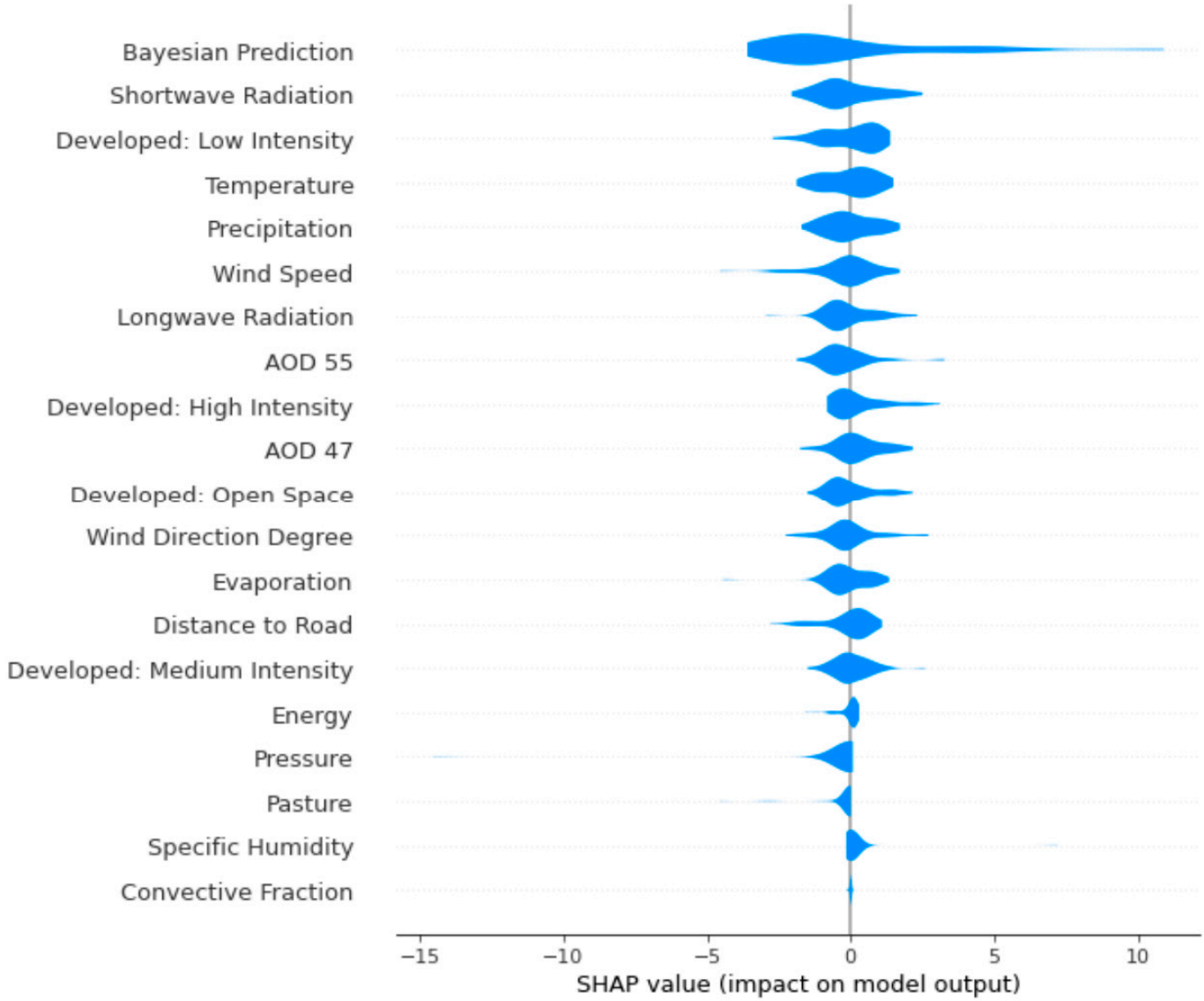


Figure 8. Summary plots of SHAP values for included variables after 999 permutations Base Value = 15.85.

Figure 10 displays maps that capture spatial-temporal fluctuations in predicted $PM_{2.5}$ concentrations (from Model 5) across the Pleasant Run Airshed for 4 randomly selected weeks within the study period. These maps, covering some key periods in the study, reveal distinct patterns in air quality variation. Elevated concentrations in November 2018 may be attributed to seasonal shifts or episodic pollution events, while lower levels in June 2019 suggest favorable meteorological conditions that likely facilitated pollutant dispersion such as higher wind speeds and precipitation than average.

Overall, the spatial variation in $PM_{2.5}$ concentrations highlights the model's ability to capture fine-scale differences across the airshed. The hybridized Bayesian SPDE—CNN approach used in Model 5 achieved a mean RMSE of 2.92 $\mu g/m^3$, rRMSE of 0.09, and MAE of 1.59, marking a high level of predictive precision and accuracy. This demonstrates the strength of integrating Bayesian spatial statistics with deep learning for robust environmental analysis and underscores the capabilities of each input approach.

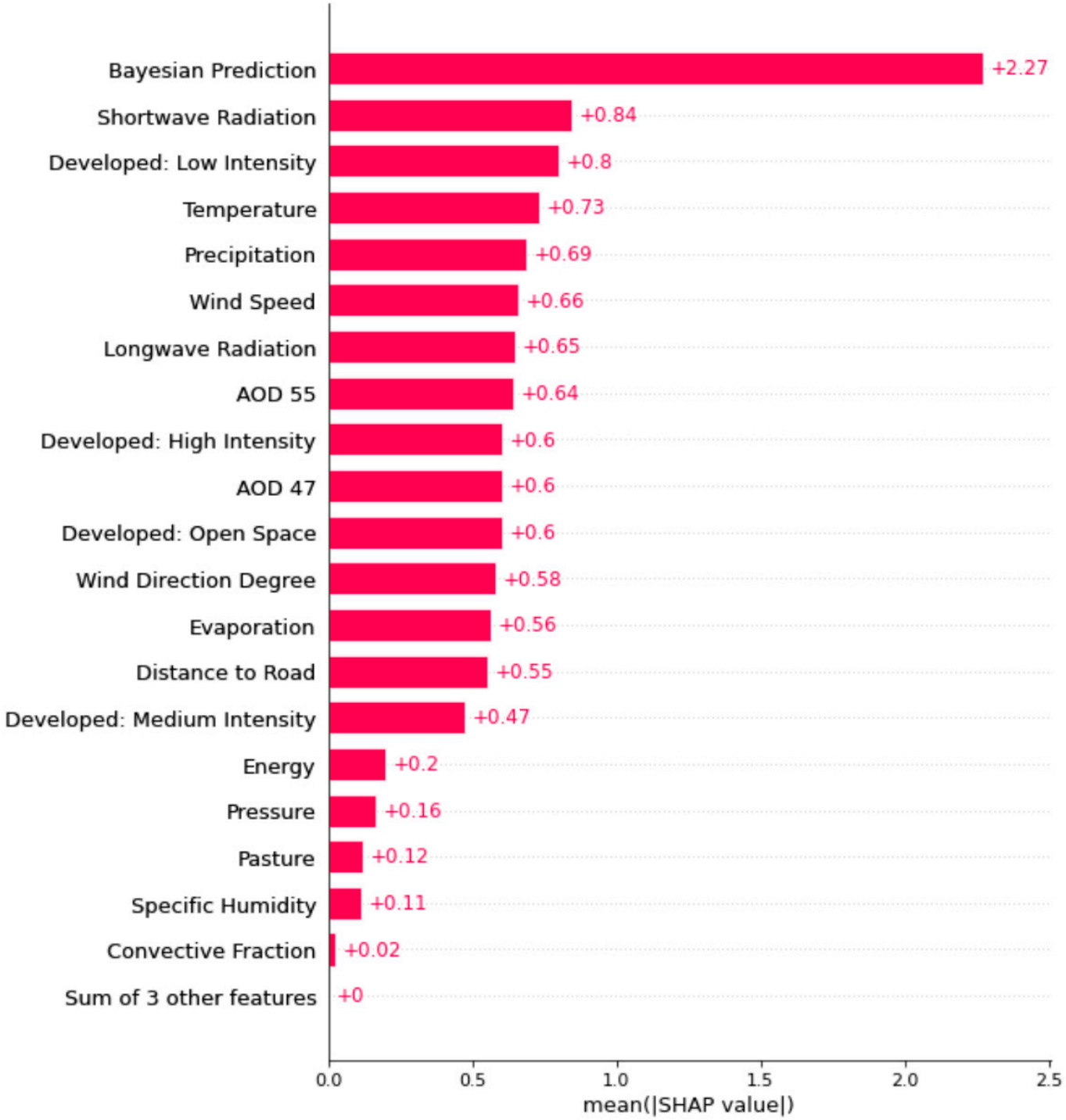


Figure 9. Variable importance after 999 Permutations (Base Value = 15.85).

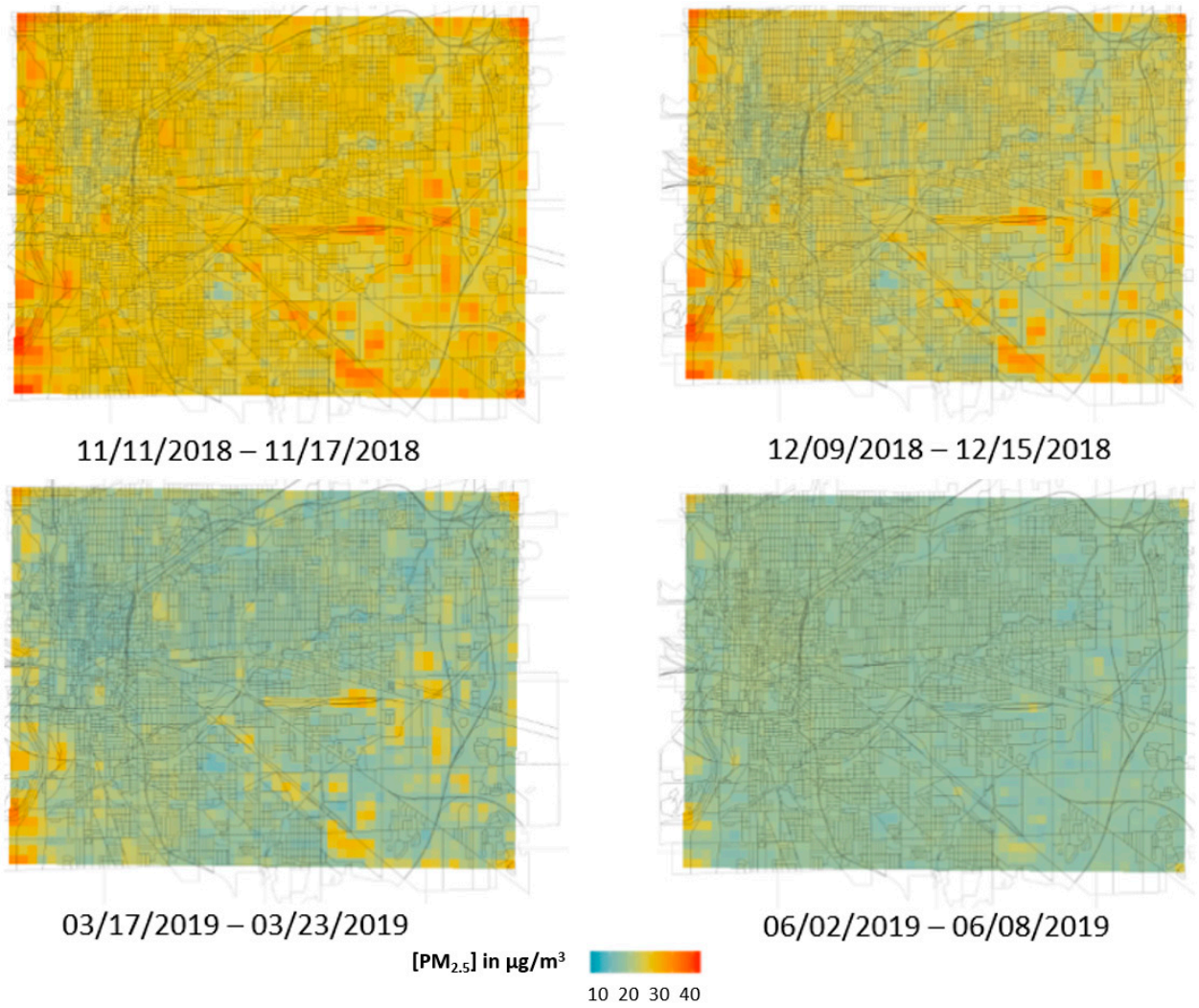


Figure 10. Estimated $PM_{2.5}$ concentration using the hybridized Bayesian SPDE—CNN model. Mean RMSE = 2.92 $\mu g/m^3$.

5. Discussion

5.1. Interpretation of Results

Our study presents a significant step forward in air quality modeling by integrating Stochastic Partial Differential Equations (SPDE) within a Bayesian framework, alongside advanced Artificial Neural Networks (ANNs) and deep learning techniques. This hybrid approach provides a robust methodological foundation for accurately predicting $PM_{2.5}$ concentrations and reflects recent advancements in environmental data modeling. By coupling the statistical precision of the SPDE framework with the adaptability and scalability of deep learning, we effectively address many complex spatial and temporal dynamics of air pollutants, capturing both fine-scale local variations and broader environmental trends.

The successful implementation of this model could support broad implications for the field of environmental modeling, particularly in urban settings where pollution levels fluctuate significantly due to a combination of anthropogenic activities and meteorological conditions. Through rigorous validation and the use of SHAP (Shapley Additive Explanations) values for model interpretation, we provide a transparent, straightforward, interpretable framework that not only helps with predictive capacity but can enhance our understanding of the driving factors behind $PM_{2.5}$ variability. This hybridized approach offers a powerful tool for air quality management and should help support data-driven public health policies and interventions. As cities worldwide face increasing pollution levels due to climate change and increased urbanization, models like this can play a crucial

role in guiding interventions that address both environmental and public health concerns. Furthermore, by identifying hotspots and trends in pollution data, this approach supports equitable urban planning and public health policies, helping to mitigate health risks in a sustainable and globally relevant manner.

5.2. SPDE Framework's Contribution to Air Quality Modeling

Integrating the SPDE approach in this study is a critical enhancement in modeling the spatial-temporal distribution of $PM_{2.5}$ concentrations. Similar to the findings of Wright et al. (2021), Fioravanti et al. (2021) and Chen et al. (2023), our model effectively captured the dynamic interactions between pollutants and geographic factors over time [2,18,19]. By leveraging the R-INLA package, we accurately accounted for fixed and time-varying effects, where previous geostatistical methods have often fallen short. The SPDE approach is compelling in its ability to model continuous spatial fields with high precision, enabling the detailed exploration of fine-scale spatial dependencies while incorporating temporal variability. This capability extends to modeling complex spatial-temporal interactions between different periods and locations. For instance, the value of a neighboring cell can influence a target cell in the present and across previous time periods. The SPDE framework excels at capturing this type of spatial-temporal structure, making it especially effective in handling such dynamic relationships.

Furthermore, this capacity to seamlessly integrate geographic and meteorological covariates, well demonstrated in the work of Chen et al. (2023) [18], underscores the flexibility and power of the SPDE framework for fine-scale environmental modeling. The model's ability to handle the inherent uncertainty in spatial predictions while incorporating critical factors such as meteorological conditions and land use represents a marked improvement over many of the traditional geostatistical interpolation methods.

5.3. Enhancements Through Artificial Neural Networks

Incorporating Artificial Neural Networks (ANNs) into our hybrid framework, particularly Convolutional Neural Networks (CNNs), enables the model to capture complex non-linear relationships between environmental variables, a standard limitation in purely statistical models. Our results align with the findings of Kurt et al. (2008) and Mao et al. (2021), who demonstrated the power of ANNs in handling large datasets and providing high-accuracy predictions [35,36]. By integrating CNNs, our model detects intricate spatial patterns essential for accurate air quality forecasting, particularly in urban environments with typical pollution hotspots.

The hybridization of Bayesian SPDE with CNN also allows us to overcome the weaknesses of purely data-driven models by providing structured outputs that are probabilistically grounded. This structured input into the CNN enhances its predictive capabilities while also retaining the interpretability of the spatial relationships inherent in the SPDE framework. The adaptability and scalability of ANNs, as emphasized by Berrocal et al. (2020), have proven instrumental in urban settings, where geographic complexity and variability in environmental conditions require a flexible, robust modeling approach [15].

5.4. SHAP Analysis and Model Interpretation

The SHAP analysis performed on our model provides key insights into the reliability and importance of each feature, ensuring that the model's predictions are both interpretable and transparent. Features like Bayesian Prediction and Shortwave Radiation exhibited broader distributions of SHAP values, highlighting their significant roles in influencing $PM_{2.5}$ concentration predictions. In this model, the variability suggests that these factors are critical in determining air quality outcomes and should be decomposed more thoroughly to understand the influences.

In Figure 7, the base value of 15.82 $\mu g/m^3$ serves as a reference point for interpreting the influence of individual features, indicating how each feature shifts the model's prediction away from this baseline. Features like 'Shortwave Radiation' and 'AOD 55' show

substantial individual contributions and interact with other environmental and meteorological variables, reflecting the multi-dimensional complexity of air pollution dynamics.

The analysis in Figures 8 and 9 further reinforces these findings, demonstrating the substantial positive influence of the SPDE Bayesian Prediction on the model's outputs. By integrating probabilistic insights from the SPDE model, the neural network gains a more nuanced understanding of the underlying patterns driving $PM_{2.5}$ concentrations. The distinct impacts of features such as temperature and developed land intensity emphasize the interplay between human activities and natural environmental processes, which are crucial for accurate air quality forecasting.

5.5. Challenges and Interdependencies

Despite the success of the hybrid Bayesian SPDE and CNN framework, several potential challenges emerged during the modeling process and should be carefully considered in any replications of the methodology.

1. Feature Redundancy and Over-attribution:
   Integrating outputs from the Bayesian SPDE model into deep learning architectures poses the risk of feature redundancy. For example, in cases where meteorological conditions appear in both models, there is a potential for over-attribution, where the same feature's impact might be counted twice. We mitigated this risk using SHAP (Shapley Additive Explanations) values to ensure proper feature attribution, making the model's predictions interpretable and reliable.
2. Handling Sparse Data in Spatial Domains:
   Sensor coverage is often uneven in urban environments, leading to sparse data points in certain regions. The Bayesian SPDE model helped predict continuous spatial fields, but integrating these predictions into the neural network layer posed challenges. Specifically, areas with fewer sensors had higher uncertainty, complicating accurate predictions in these zones. Future iterations may benefit from optimizing sensor placements or employing methods to reduce uncertainty in sparsely monitored areas. However, the SPDE approach is notably superior when dealing with sparsity.
3. Computational Complexity:
   The combination of Bayesian models and deep learning requires substantial computational resources. Processing large spatial datasets across many iterations involves significant memory and processing power. The model was optimized for high-performance computing systems to manage this, but further work is needed to streamline processes for more efficient training without sacrificing accuracy. If modeling was performed on a daily, rather than weekly, basis this would substantially increase the computational demand.
4. Boundary Effects and Edge Handling:
   Spatial models often exhibit boundary effects, where predictions near the study area's edges tend to be less reliable due to a lack of neighboring data points. Although mesh design and the inclusion of spatial covariates in this study have helped mitigate these effects, boundary areas remain challenging. Future work could further refine the mesh design by implementing denser meshing near the edges or applying a buffer zone approach. A buffer zone would extend beyond the core study area, allowing boundary points to incorporate influence from adjacent regions and thereby smooth transitions along edges.

These challenges provide critical insight into the current model's limitations and present avenues for further refinement, particularly regarding sensor placement, potential effects of feature redundancy, and computational optimization.

5.6. Comparative Analysis and Future Directions

Our study's integration of SPDE within a Bayesian framework, augmented by a Convolutional Neural Network (CNN), offers a sophisticated and comprehensive approach that extends the innovations seen in prior air quality modeling research. Achieving an average RMSE of 2.92 μg/m³, our model's precision is comparable to the ensemble-based model described by Di et al. (2019), which recorded an RMSE of 2.79 μg/m³ across the contiguous United States [37]. This parallel emphasizes the efficacy of combining advanced statistical methods with deep learning techniques to capture the nuanced spatiotemporal dynamics of $PM_{2.5}$ dispersion.

Building on the data integration strategies employed by Han et al. (2022) and Tang et al. (2017), our model incorporates diverse data sources, including satellite imagery, ground sensor readings, land use/land cover, and meteorological/planetary boundary layer variables [13,38]. Our approach further enriches the representation of spatial distributions and temporal variations in $PM_{2.5}$ concentrations, adding to the advancements in these studies. Furthermore, our work supports the findings of Chae et al. (2021), showcasing the superior predictive performance of CNN architectures in $PM_{2.5}$ interpolation by accurately capturing spatial hierarchies and non-linear patterns.

Our results also reinforce the significant influence of environmental and anthropogenic factors on $PM_{2.5}$ levels, aligning with findings from Lee (2019) and Requia et al. (2019) [16,39]. Integrating satellite-derived Aerosol Optical Depth (AOD) data with detailed land use and urban morphology analyses enhances our understanding of spatial heterogeneities, contributing to the effectiveness of our model for detailed exposure assessments and public health advisories. This methodological synergy echoes the integrated approach pioneered by Kloog et al. (2011), where high-resolution satellite data were used to estimate daily $PM_{2.5}$ levels relevant for more precise epidemiological studies [40]. Our research highlights how advanced statistical methods, combined with machine learning techniques, can drive sustainable outcomes in climate action and public health.

Our findings align with prior studies demonstrating the effectiveness of hybrid models in improving air quality predictions. Di et al. (2019) [37], as previously mentioned, achieved high accuracy using an ensemble-based model, with an RMSE of 2.79 μg/m³ across the contiguous United States, closely matching our hybrid model's performance (RMSE = 2.92 μg/m³). Similarly, Kloog et al. (2011) and Han et al. (2022) highlighted the value of integrating diverse datasets (e.g., satellite AOD and ground sensors) for spatial interpolation of PM2.5, an approach that our model advances by combining SPDE's spatial precision with CNN's ability to capture non-linearities [38,40]. The results underscore the hybrid model's robustness in capturing $PM_{2.5}$ variability, which prior approaches often addressed using single-method geostatistical or machine learning models.

The comparative success of our hybrid model underscores its potential as a robust methodological foundation for future studies. By fusing the statistical strengths of SPDE with the dynamic modeling capabilities of CNNs, we present a novel framework adept at handling the recognized complex interplay of factors influencing urban air quality. Our approach not only aligns with but also advances the current research trajectory, further helping to set a new precedent for enhancing the precision and applicability of air quality models across diverse urban and rural environments. The general availability of relevant data (except for in situ $PM_{2.5}$ concentrations) employed in this study, within the U.S. and globally, makes the methodology outlined here widely replicable. Furthermore, these data and software applications are free of charge.

Our findings also suggest that integrating SPDE with CNN mitigates the limitations inherent in using each approach individually. The promising results highlight a pathway for future research to explore deeper integrations of spatial statistical methods, machine learning and deep learning with statistical frameworks, enabling further improvements in predictive capabilities and a more profound understanding of the environmental and anthropogenic factors contributing to $PM_{2.5}$ levels. Furthermore, these methods could be

used for other air quality measures (such as ground level Ozone or nitrous oxides) or other types of continuous environmental phenomena.

Future studies could build upon this framework by incorporating additional covariates, such as real-time traffic flow or industrial emissions data, which would enhance model performance in highly dynamic urban areas. Additionally, expanding the model's scope to include different geographic locations would provide valuable insights into the generalizability and robustness of the hybrid framework. This hybrid model can also inform interdisciplinary applications, ranging from public health interventions to urban planning and policy development, underscoring the broader societal relevance of this research.

Future studies may explore data imputation methods to address limitations from sparse sensor data, as noted in prior urban air quality research (Han et al., 2022). Techniques like kriging, SPDE, or model-based imputation could expand sensor data coverage, improving spatial resolution and predictive accuracy in under-monitored areas. Additionally, on the model development side, incorporating transfer learning architectures, such as ResNet, or Residual Network, a type of deep neural network architecture designed to facilitate the training of very deep networks by introducing residual connections, could enhance the model's performance by leveraging pre-trained spatial feature extraction capabilities [41]. This approach may increase efficiency and accuracy, especially in regions with limited training data.

6. Conclusions

This study offers valuable contributions to air quality modeling by introducing a hybridized approach that integrates Bayesian SPDE and CNN to predict $PM_{2.5}$ concentrations. The hybrid model demonstrates substantial improvements in predictive accuracy, with a mean RMSE of 2.92 $\mu g/m^3$, underscoring the efficacy of combining traditional spatial statistical methods with advanced deep learning techniques to address complexities within environmental data and the phenomena they measure.

Our findings align with previous research, confirming the critical role of meteorological factors, such as wind speed, precipitation, and satellite measured Aerosol Optical Depth (AOD), in shaping the distribution and transport of $PM_{2.5}$. For example, the predominant southwest wind patterns identified in the study highlight the importance of incorporating regional climatology into predictive models to enhance the understanding of pollutant dispersion. Such insights are essential for developing targeted air quality interventions, particularly in urban settings with high pollution levels.

Beyond the meteorological and boundary layer factors, our study reveals the intricate relationships between environmental covariates and $PM_{2.5}$ levels. Using high-resolution, spatially, and temporally rich datasets allowed for detailed exploration of these dynamics, adding value to environmental monitoring efforts. However, limitations, such as the edge effects in spatial analysis and the uneven distribution of sensors in homogenous urban land-cover zones, suggest areas for improvement. Addressing these limitations in future research, particularly by optimizing sensor placement across diverse landscapes and better handling boundary conditions, could further refine model accuracy and generalizability.

Our model's capacity to capture both local and global spatial patterns and its adaptability to high-dimensional environmental data positions it as a flexible, scalable tool for urban air quality monitoring. Future research can build on this framework by incorporating additional data sources, such as traffic flow or industrial emissions, to enhance predictive performance in complex urban environments. These efforts align with the global sustainability agenda, specifically SDG 3 (Good Health and Well-Being) and SDG 13 (Climate Action), highlighting the role of air quality management in fostering healthier, more resilient cities. Future work should focus on expanding this hybrid approach to new geographic areas, ensuring its continued relevance in addressing the dual challenges of environmental degradation and public health.

Additionally, while certain land cover types and anthropogenic factors showed less pronounced influence on $PM_{2.5}$ levels, this finding opens new avenues for investigating

the interactions between these factors and unmeasured variables, such as local emissions and population density. Expanding the scope of environmental and socio-economic data in future studies would likely lead to more nuanced models and advocate for more robust data-driven public health strategies.

Overall, this research highlights the potential of hybrid modeling approaches to improve environmental monitoring and decision-making. By integrating advanced spatial statistical models with deep learning architectures, we have contributed to developing a newer exemplar for air quality prediction, one that holds promise for more accurate, data-driven interventions. The ongoing refinement and application of these advanced hybridized methodologies will support public health outcomes and inform future environmental policies to mitigate urban air pollution and help develop a more sustainable and equitable environment.

Author Contributions: Conceptualization, D.P.J., G.F. and A.H.; Methodology, D.P.J. and N.R.; Software, D.P.J. and N.R.; Validation, D.P.J. and N.R.; Formal analysis, D.P.J. and N.R.; Investigation, D.P.J., G.F. and A.H.; Resources, D.P.J., N.R., G.F. and A.H.; Data curation, D.P.J. and A.H.; Writing—original draft, D.P.J., N.R., G.F. and A.H.; Writing—review & editing, D.P.J., G.F. and A.H.; Visualization, D.P.J.; Supervision, D.P.J.; Project administration, D.P.J.; Funding acquisition, D.P.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data contained in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Cameletti, M.; Gómez-Rubio, V.; Blangiardo, M. Bayesian Modelling for Spatially Misaligned Health and Air Pollution Data through the INLA-SPDE Approach. *Spat. Stat.* **2019**, *31*, 100353. [CrossRef]
2. Wright, N.; Newell, K.; Lam, K.B.H.; Kurmi, O.; Chen, Z.; Kartsonaki, C. Estimating Ambient Air Pollutant Levels in Suzhou through the SPDE Approach with R-INLA. *Int. J. Hyg. Environ. Health* **2021**, *235*, 113766. [CrossRef] [PubMed]
3. Alimissis, A.; Philippopoulos, K.; Tzanis, C.G.; Deligiorgi, D. Spatial Estimation of Urban Air Pollution with the Use of Artificial Neural Network Models. *Atmos. Environ.* **2018**, *191*, 205–213. [CrossRef]
4. Li, L. A Comparative Study of Three Deep Learning Models for PM2.5 Interpolation. In Proceedings of the 26th International Database Engineered Applications Symposium (IDEAS '22], Budapest, Hungary, 22–24 August 2022. [CrossRef]
5. Bell, M.L.; Dominici, F.; Ebisu, K.; Zeger, S.L.; Samet, J.M. Spatial and Temporal Variation in PM2.5 Chemical Composition in the United States for Health Effects Studies. *Environ. Health Perspect.* **2007**, *115*, 989–995. [CrossRef]
6. Bekesiene, S.; Meidute-Kavaliauskiene, I. Artificial Neural Networks for Modelling and Predicting Urban Air Pollutants: Case of Lithuania. *Sustainability* **2022**, *14*, 2470. [CrossRef]
7. Krittanawong, C.; Qadeer, Y.K.; Hayes, R.B.; Wang, Z.; Virani, S.; Thurston, G.D.; Lavie, C.J. PM2.5 and Cardiovascular Health Risks. *Curr. Probl. Cardiol.* **2023**, *48*, 101670. [CrossRef]
8. Finch, B.K.; Thomas, K.; Gibbons, J.R.; Beck, A.N. The Impact of Residing in a Gang Territory on Adverse Birth Outcomes: Evidence from Los Angeles. *J. Urban Health* **2021**, *98*, 233–247. [CrossRef]
9. Johnson, N.M.; Hoffmann, A.R.; Behlen, J.C.; Lau, C.; Pendleton, D.; Harvey, N.; Shore, R.; Li, Y.; Chen, J.; Tian, Y.; et al. Air Pollution and Children's Health—A Review of Adverse Effects Associated with Prenatal Exposure from Fine to Ultrafine Particulate Matter. *Environ. Health Prev. Med.* **2021**, *26*, 72. [CrossRef]
10. Ganguly, A.; Ghosh, S.; Shin, B.-C.; Touma, M.; Wadehra, M.; Devaskar, S.U. Gestational Exposure to Air Pollutants Perturbs Metabolic and Placenta-Fetal Phenotype. *Reprod. Toxicol.* **2024**, *128*, 108657. [CrossRef]
11. Wang, T.; Zhang, T.; Sun, L.; Li, W.; Zhang, C.; Yu, L.; Guan, Y. Gestational B-Vitamin Supplementation Alleviates PM2.5-Induced Autism-like Behavior and Hippocampal Neurodevelopmental Impairment in Mice Offspring. *Ecotoxicol. Environ. Saf.* **2019**, *185*, 109686. [CrossRef]
12. Adams, M.D.; Kanaroglou, P.S. Mapping Real-Time Air Pollution Health Risk for Environmental Management: Combining Mobile and Stationary Air Pollution Monitoring with Neural Network Models. *J. Environ. Manag.* **2016**, *168*, 133–141. [CrossRef] [PubMed]

13. Tang, M.; Wu, X.; Agrawal, P.; Pongpaichet, S.; Jain, R. Integration of Diverse Data Sources for Spatial PM2.5 Data Interpolation. *IEEE Trans. Multimed.* **2017**, *19*, 408–417. [CrossRef]
14. Cordova, C.H.; Portocarrero, M.N.L.; Salas, R.; Torres, R.; Rodrigues, P.C.; López-Gonzales, J.L. Air Quality Assessment and Pollution Forecasting Using Artificial Neural Networks in Metropolitan Lima-Peru. *Sci. Rep.* **2021**, *11*, 24232. [CrossRef] [PubMed]
15. Berrocal, V.J.; Guan, Y.; Muyskens, A.; Wang, H.; Reich, B.J.; Mulholland, J.A.; Chang, H.H. A Comparison of Statistical and Machine Learning Methods for Creating National Daily Maps of Ambient PM2.5 Concentration. *Atmos. Environ.* **2020**, *222*, 117130. [CrossRef] [PubMed]
16. Requia, W.J.; Coull, B.A.; Koutrakis, P. Evaluation of Predictive Capabilities of Ordinary Geostatistical Interpolation, Hybrid Interpolation, and Machine Learning Methods for Estimating PM2.5 Constituents over Space. *Environ. Res.* **2019**, *175*, 421–433. [CrossRef]
17. Cabaneros, S.M.; Calautit, J.K.; Hughes, B.R. A Review of Artificial Neural Network Models for Ambient Air Pollution Prediction. *Environ. Model. Softw.* **2019**, *119*, 285–304. [CrossRef]
18. Chen, J.; Miao, C.; Yang, D.; Liu, Y.; Zhang, H.; Dong, G. Estimation of Fine-Resolution PM2. 5 Concentrations Using the INLA-SPDE Method. *Atmos. Pollut. Res.* **2023**, *14*, 101781. [CrossRef]
19. Fioravanti, G.; Martino, S.; Cameletti, M.; Cattani, G. Spatio-Temporal Modelling of PM10 Daily Concentrations in Italy Using the SPDE Approach. *Atmos. Environ.* **2021**, *248*, 118192. [CrossRef]
20. Lindgren, F.; Havard, R.; Lindstrom, J. Explicit Link between Gaussian Fields and Gaussian Markov Random Fields: The Stochastic Partial Differential Equation Approach. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **2011**, *73*, 423–498. [CrossRef]
21. Rue, H.; Martino, S.; Chopin, N. Approximate Bayesian Inference for Latent Gaussian Models by Using Integrated Nested Laplace Approximations. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **2009**, *71*, 319–392. [CrossRef]
22. Li, W.; Kang, S.; Sun, Y.; Bai, W.; Wang, Y.; Song, H. A Machine Learning Approach for Air-Quality Forecast by Integrating GNSS Radio Occultation Observation and Weather Modeling. *Atmosphere* **2023**, *14*, 58. [CrossRef]
23. Dewitz, J.; U.S. Geological Survey. *National Land Cover Databae (NLCD) 2019 Products (Ver. 3.0)*; U.S. Geological Survey: Reston, VA, USA, 2021.
24. Wan, Z.; Hook, S.; Hulley, G. MOD11A1 MODIS/Terra Land Surface Temperature/Emissivity Daily L3 Global 1km SIN Grid V006. NASA EOSDIS Land Processes DAAC. 2015. Available online: https://lpdaac.usgs.gov/products/mod11a1v006/ (accessed on 3 March 2024).
25. Cosgrove, B.A.; Lohmann, D.; Mitchell, K.E.; Houser, P.R.; Wood, E.F.; Schaake, J.C. Real-time and retrospective forcing in the North American Land Data Assimilation System (NLDAS) project. *J. Geophys. Res. Atmos.* **2003**, *108*, 2002JD003118. [CrossRef]
26. Heintzelman, A.; Filippelli, G.M.; Moreno-Madriñan, M.J.; Wilson, J.S.; Wang, L.; Druschel, G.K. Efficacy of Low-Cost Sensor Networks at Detecting Fine-Scale Variations in Particulate Matter in Urban Environments. *Int. J. Environ. Res. Public Health* **2023**, *20*, 1934. [CrossRef]
27. Allaire, C. Keras: R Interface to 'Keras'. 2023. Available online: https://CRAN.R-project.org/package=keras (accessed on 24 April 2024).
28. Allaire, J.J.; Kalinowski, T.; Falbel, D.; Eddelbuettel, D.; Tang, Y.; Golding, N.; Google Inc.; Posit; PBC. Tensorflow: R Interface to "TensorFlow". 2024. Available online: https://cran.r-project.org/web/packages/tensorflow/index.html (accessed on 30 April 2024).
29. Aybar, C.; Qiusheng, W.; Bautista, L.; Yali, R.; Barja, A.; Ushey, K.; Ooms, J.; Appelhans, T.; Allaire, J.J.; Tang, Y.; et al. Rgee: R Bindings for Calling the "Earth Engine" API. 2021. Available online: https://CRAN.R-project.org/package=rgee (accessed on 1 December 2021).
30. Hijmans, R.J.; van Etten, J.; Sumner, M.; Cheng, J.; Baston, D.; Bevan, A.; Bivand, R.; Busetto, L.; Canty, M.; Fasoli, B.; et al. Raster: Geographic Data Analysis and Modeling. 2023. Available online: https://cran.r-project.org/web/packages/raster/index.html (accessed on 30 April 2024).
31. Pebesma, E. Simple Features for R: Standardized Support for Spatial Vector Data. *R J.* **2018**, *10*, 439–446. [CrossRef]
32. Pebesma, E.; Bivand, R. Spatial Data Science. Available online: https://r-spatial.org/book/ (accessed on 30 April 2024).
33. Krogh, A.; Hertz, J. A Simple Weight Decay Can Improve Generalization. In *Advances in Neural Information Processing Systems*; Morgan-Kaufmann: Burlington, MA, USA, 1991; Volume 4.
34. Donateo, A.; Contini, D. Correlation of Dry Deposition Velocity and Friction Velocity over Different Surfaces for PM2.5 and Particle Number Concentrations. *Adv. Meteorol.* **2014**, *2014*, e760393. [CrossRef]
35. Kurt, A.; Gulbagci, B.; Karaca, F.; Alagha, O. An Online Air Pollution Forecasting System Using Neural Networks. *Environ. Int.* **2008**, *34*, 592–598. [CrossRef]
36. Mao, W.; Wang, W.; Jiao, L.; Zhao, S.; Liu, A. Modeling Air Quality Prediction Using a Deep Learning Approach: Method Optimization and Evaluation. *Sustain. Cities Soc.* **2021**, *65*, 102567. [CrossRef]
37. Di, Q.; Amini, H.; Shi, L.; Kloog, I.; Silvern, R.; Kelly, J.; Sabath, M.B.; Choirat, C.; Koutrakis, P.; Lyapustin, A.; et al. An Ensemble-Based Model of PM2.5 Concentration across the Contiguous United States with High Spatiotemporal Resolution. *Environ. Int.* **2019**, *130*, 104909. [CrossRef]
38. Han, S.; Kundhikanjana, W.; Towashiraporn, P.; Stratoulias, D. Interpolation-Based Fusion of Sentinel-5P, SRTM, and Regulatory-Grade Ground Stations Data for Producing Spatially Continuous Maps of PM2.5 Concentrations Nationwide over Thailand. *Atmosphere* **2022**, *13*, 161. [CrossRef]

39. Lee, H.J. Benefits of High Resolution PM2.5 Prediction Using Satellite MAIAC AOD and Land Use Regression for Exposure Assessment: California Examples. *Environ. Sci. Technol.* **2019**, *53*, 12774–12783. [CrossRef]
40. Kloog, I.; Koutrakis, P.; Coull, B.A.; Lee, H.J.; Schwartz, J. Assessing Temporally and Spatially Resolved PM2.5 Exposures for Epidemiological Studies Using Satellite Aerosol Optical Depth Measurements. *Atmos. Environ.* **2011**, *45*, 6267–6275. [CrossRef]
41. He, K.; Zhang, X.; Ren, S.; Sun, J. Identity Mappings in Deep Residual Networks. In *Computer Vision—ECCV 2016*; Leibe, B., Matas, J., Sebe, N., Welling, M., Eds.; Springer International Publishing: Cham, Switzerland, 2016; pp. 630–645. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.