

Article

Assessing the Potential of Artificial Intelligence (Artificial Neural Networks) in Predicting the Spatiotemporal Pattern of Wildfire-Generated PM_{2.5} Concentration

Mojgan Mirzaei ^{1,*}, Stefania Bertazzon ¹, Isabelle Couloigner ^{1,2}  and Babak Farjad ^{3,4}

¹ Department of Geography, University of Calgary, Calgary, AB T2N 1N4, Canada; bertazzs@ucalgary.ca (S.B.); icouloig@ucalgary.ca (I.C.)

² Department of Ecosystem and Public Health, University of Calgary, Calgary, AB T2N 1N4, Canada

³ Centre of Excellence (COE), Environmental Monitoring and Science Division, Government of Alberta, Calgary, AB T2N 1N4, Canada; bfarjad@ucalgary.ca

⁴ Department of Geomatics Engineering, University of Calgary, Calgary, AB T2N 1N4, Canada

* Correspondence: mojgan.mirzaei@ucalgary.ca

Abstract: To evaluate the health effects of wildfire smoke, it is crucial to identify reliable models, at fine spatiotemporal resolution, of exposure to wildfire-generated PM_{2.5}. To this end, satellite-driven aerosol optical depth (AOD) measurements are widely used in exposure models, providing long and short-term PM_{2.5} predictions. Multiple regression models, specifically land use regression (LUR), incorporating AOD images have shown good potential for estimating long-term PM_{2.5} exposure, but less so for short-term predictions. In this study, we developed artificial neural networks (ANNs) and, in particular, multilayer perceptron (MLP) by integrating ground-based PM_{2.5} measurements with AOD images and meteorological and spatial variables. Moreover, we used spatial- and temporal-ANNs to investigate and compare the ANNs' ability to predict different PM_{2.5} concentration levels caused by abrupt spatial and temporal changes in fire smoke. The study herein analyzes and compares the viability of previously established neural network approaches in predicting short-term PM_{2.5} exposure during the 2014–2017 wildfire seasons in the province of Alberta, Canada. The performance of ANNs is also compared to classical models, including simple correlation (PM_{2.5} vs. AOD) and multiple linear regression (MLR) including meteorological and land-use predictors (MET_AOD_LUR). Our study shows that ANN achieved a 15% to 113% R² increase compared to competing models.

Keywords: wildfire; AOD; PM_{2.5}; ANN; MLR; spatiotemporal modelling; smoke exposure; public health



Citation: Mirzaei, M.; Bertazzon, S.; Couloigner, I.; Farjad, B. Assessing the Potential of Artificial Intelligence (Artificial Neural Networks) in Predicting the Spatiotemporal Pattern of Wildfire-Generated PM_{2.5} Concentration. *Geomatics* **2021**, *1*, 18–33. <https://doi.org/10.3390/geomatics1010003>

Received: 6 November 2020

Accepted: 27 December 2020

Published: 11 January 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 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

Severe air pollution due to wildfire smoke poses the risk of detrimental impacts on human health and the environment. Since 2010, possibly in connection with climate change, wildfires and biomass burning has increased, becoming more problematic for population health and the environment [1]. Toxic smoke from wildfires threaten not only the population living within the fire region but also those living in distant areas, because smoke can travel thousands of kilometres. Vulnerable populations, including individuals with respiratory illnesses, cardiovascular disease, and asthma, as well as elderly people, pregnant women, and children, may experience more severe short-term and long-term impacts [2].

Wildfire smoke composition, which depends on multiple factors such as the type of burning material, moisture content, fire temperature, and wind conditions, consists of a large amount of fine particulate matter and harmful gases such as nitrogen oxides and sulphur oxides [1]. Wildfire smoke is a significant contributor to particulate matter [3], and, among wildfire emissions, fine particles are the most harmful and the most concern-

ing components for human health [2]. Exposure to smoke-generated PM_{2.5} is strongly associated with mortality and respiratory morbidity [1,4,5].

The province of Alberta, Canada, with a population of about 4.4 million, has experienced one or more wildfire smoke episodes almost every summer during the fire season. Smoke drifts over Alberta from seasonal wildfires in southern Alberta, the province of British Columbia (BC), as well as from the United States, leading to poor air quality across the province (Alberta Health Services (<https://www.albertahealthservices.ca/>)). Wildfire events are likely to continue to occur in upcoming fire seasons, therefore it is important to quantify PM_{2.5} levels at a fine spatial and temporal resolution to: (1) evaluate the risk posed to the Alberta population by smoke-generated PM_{2.5} exposure; (2) provide the population, and particularly the vulnerable groups, with timely warnings; and (3) assess the association between the level of PM_{2.5} concentration and the health problems reported during each smoke episode.

Evaluating the health impact of PM_{2.5} emissions from wildfire is dependent on accurately quantifying smoke-generated PM_{2.5} concentration; however, ground-based PM_{2.5} measurements are spatially sparse. Satellite measurements of aerosol optical depth (AOD) are known to be appropriate complements to air quality (AQ) stations' measurements to improve the spatial coverage and resolution of PM_{2.5} concentration. They have been widely used in studies utilizing different methodologies [6–15].

A comprehensive review of various techniques used to predict PM_{2.5} levels using AOD has been summarized in a systematic study by Youssouf et al. (2014) [1]. This study included a variety of models ranging from simple linear regression to more complex approaches, such as chemical air quality models and machine learning approaches. Generally, such studies provide information on the relationship between ambient PM_{2.5} and AOD, but not during fire smoke periods. Thus far, only a few studies have been published aimed at predicting wildfire-generated PM_{2.5} [16–19].

Identifying the most appropriate model for a particular problem is not straightforward and depends on numerous factors. Two of the most substantial factors are the complexity of the problem and understanding the problem at hand [20,21]. There is a negative relationship between the complexity of a problem and its theoretical understanding. In general, provided that a complete perception of the problem and adequate data are available, the most suitable solution for addressing a problem is a full physical model [20,21]. In complex problems, and when large datasets, numerous variables, and nonlinearity are involved, artificial networks, particularly multilayer perceptron (MLP), are more appropriate [20]. Some prior assumptions concerning the data distribution, linearity, and stationarity may need to be considered by traditional regression models such as multiple linear regression. Many forms of nonlinear regression and model forms deal with the violation of each assumption (e.g., Spatial AutoRegression (SAR), Geographically Weighted Regression (GWR), etc.). However, when the numbers of observations and variables are very large, the problem may be the simultaneous presence of different nonlinear distributions, or simply the difficulty of correctly identifying the distribution of each variable. In such situations, non-parametric machine learning approaches may be more suitable [22]. One of the main advantages of an MLP approach is its ability to deal with nonlinear systems [23]. The effective ability of artificial neural networks (ANNs) as an alternative approach to more traditional statistical techniques has been previously proven in various studies [23,24].

An ANN is constructed based on a biologically inspired computational approach, with a set of processing neurons that are connected with coefficients called weights [25,26]. ANN has been applied to studies in the field of air quality to forecast and predict air pollution and also to detect traffic-based air pollution [20,27–32].

In our previous studies, we approached the relationship between smoke-related PM_{2.5} and AOD with land use regression [18] and linear mixed effect (LME) models [16]. Although each approach has merit, a significant problem was the presence of gaps in the data and the need to integrate them with data from different sources. There may be no alternatives to filling gaps in observations with model estimates and gaps in satellite

coverage with satellite data at the different spatial or temporal resolution, however these strategies yield inconsistent variables that hardly conform to known statistical distributions.

For these reasons, we are proposing to assess a machine learning approach in comparison with those models. The main objective of the present study was to determine the prediction ability of ANNs, and more specifically MLP, to quantify spatial and temporal wildfire-generated PM_{2.5} using remote sensing data and relevant ancillary variables. To the best of our knowledge, machine learning methods have not been used for PM_{2.5} prediction during wildfire episodes.

We examined the performance of ANNs for different levels of PM_{2.5} concentration over four consecutive fire seasons throughout Alberta. The ANNs were developed temporally (for each year separately) and spatially (for different regions). We also compared the performance of the ANN with that of a mixed (MLE) model [16] by training the 2017 ANN on the data for the entire 4-year period and using a very similar set of input parameters. Multiple linear regression (MLR), in particular meteorological–AOD–land use regression (MET_AOD_LUR) models were also developed as benchmarks to compare and investigate the short-term prediction performance of neural networks for wildfire-specific air pollution.

2. Materials and Methods

2.1. Study Domain and Period

The spatial domain of our study included the province of Alberta, Canada (Figure 1, left), with an area of 661,848 km² and a population of about 4.4 million (<https://www.alberta.ca/municipal-census.aspx>). The province of Alberta includes ten airshed zones. More than 70 air monitoring stations located across the province provide data contributing to the Air Quality Health Index (AQHI) [33] (AAC 2006). The temporal domain included August and September from 2014 to 2017, which included severe smoky periods due to the occurrence of seasonal wildfires.

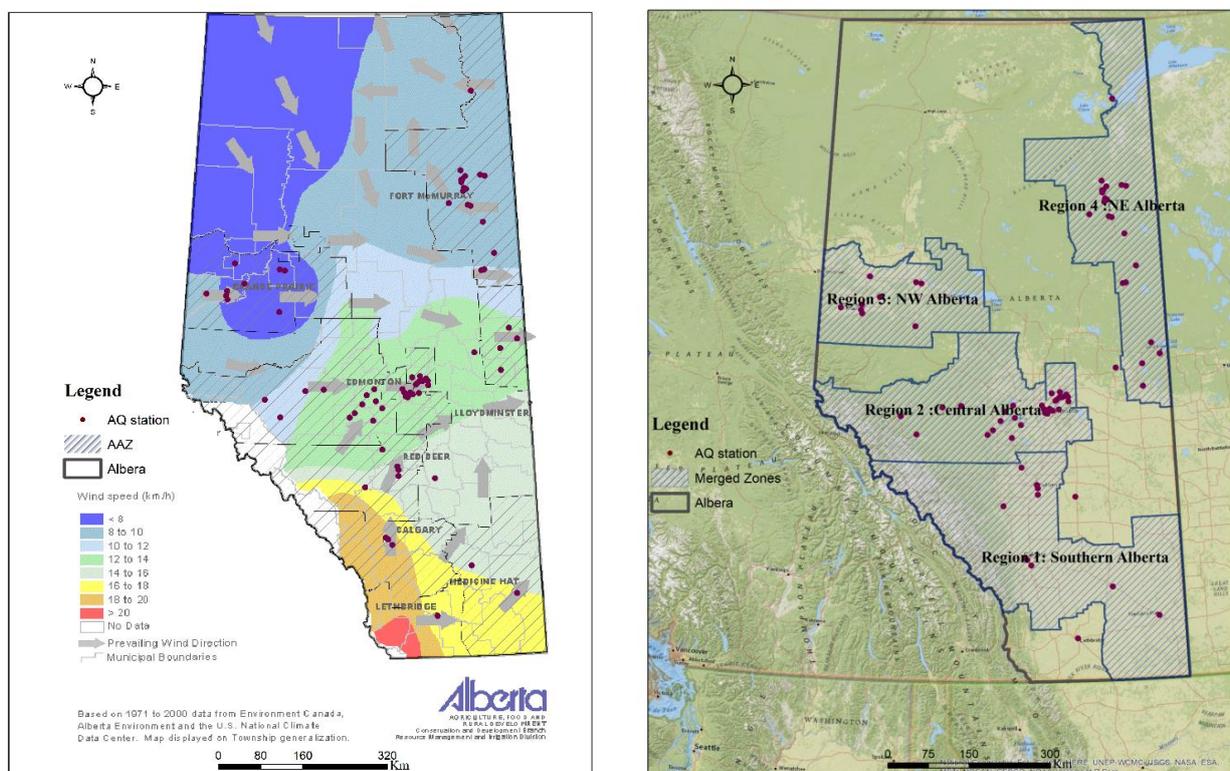


Figure 1. Left: Alberta airshed zones (AAZs), location of AQ stations, and 30-year mean wind speed and prevailing direction for summer in Alberta. Right: four regions including southern, central, NW and NE, made by merging AAZs to make spatial artificial neural networks (ANNs).

Two approaches were used to allocate the data to the ANNs. The first was temporal: the dataset was divided into years, and for each year, an ANN was developed and trained separately for the province, resulting in four separate temporal ANNs.

The second was spatial: Alberta is relatively large and diverse, therefore smoke levels are likely to vary across the province based on the proximity to the fire source and the ground topology, as well as wind speeds and direction. The study region was divided into four regions—southern, central, northwest (NW), and northeast (NE), based on the borders of airshed zones with sufficient data (Figure 1, right)—and four separate ANNs were developed and trained for each region.

2.2. Data

PM_{2.5}, temperature, wind: Ground-based PM_{2.5} measurements (the dependent variable), and meteorological predictors, i.e., temperature (TEMP), wind speed (WSP_{E-W} , WSP_{N-S}), and wind directions were collected at the location of continuous AQ stations for each fire season. The number of stations with available PM_{2.5} measurement ability varied for each year, and ranged from 24 stations in 2014 to 46 stations in 2017.

Hourly PM_{2.5} concentrations and temperature values were averaged into daily mean PM_{2.5} concentration and temperature.

The wind is represented by two vectors in the east-west (E–W) and north-south (N–S) direction as follow:

$$WSP_{E-W} = WSP * \cos(\theta)$$

$$WSP_{N-S} = WSP * \sin(\theta)$$

where WSP denotes the hourly wind speed for each station, and θ is the wind direction angle. The daily average WSP_{E-W} and WSP_{N-S} were calculated and used as two input variables to develop the ANN. The background of Figure 1 (left) shows the 30-year mean speed and prevailing direction for Summer in Alberta [34].

AOD images: Daily 0.1 degrees (~10 km × 10 km) spatial resolution AOD images were collected from MODIS (MODerate resolution Imaging Spectrodiometer Instruments) Terra collection data [35]. More details about AOD image collecting and preprocessing can be found in our previous paper [18].

Elevation, latitude and longitude: A digital elevation map was created from the DMTI Spatial contour lines data of Alberta [36] using ArcGIS, and the elevation of each AQ station was extracted. Latitude and longitude were also used as input variables.

Normalized difference vegetation index (NDVI): NDVI is a greenness index used to measure the vegetation and biomass production of the area. More details can be found in [18].

2.3. Methods

Machine learning approaches such as neural networks, support vector machine, random forest, etc., have the ability to predict variables of interest when there is no complete theoretical description about the process, but a useful set of observations is available [37]. As noted, ANNs do not require prior assumptions on data distribution and can handle large amounts of data. Therefore, the ANN approach presents the potential of yielding reliable predictions despite inconsistent datasets for the analysis of smoke-related PM_{2.5}. ANN is also applicable for addressing complex problems, such as the spatiotemporal prediction of wildfire-generated PM_{2.5}, where numerous factors affect the PM_{2.5} concentration and the AOD–PM_{2.5} relationship. In addition to large gaps in the AOD images, temporal and spatial mismatching between the AOD measurements and the point-based PM_{2.5} data, and uncertainties in the data accuracy, there are concerns with the dependency of the PM_{2.5} concentration on unknown factors, which cannot be explicitly included in the model, however are likely to interact with the model variables and errors. Additionally, ANNs have a great capacity for solving nonlinear problems [26].

The main features of ANNs are their topology or architecture, activation function, and training algorithms. Network architecture, including the number of layers and nodes

in the hidden layers, is one of the initial steps in developing an ANN. The complexity of the model, the probable noise in the data, and the size of the training dataset are known as the most important factors affecting the optimum number of nodes in hidden layers [20].

The multilayer perceptron (MLP) network is a commonly used type of feed-forward ANN topology (where the data moves in only one direction, forward, from the input nodes, through the hidden nodes to the output nodes) that consists of at least three layers including a layer of input neurons (nodes), one or more hidden layers, and an output layer.

The selection of a sufficiently informative and significant set of input predictors to employ in the model is a key consideration for any model building procedure. The performance of the ANN is affected by uninformative input predictors, which yield weakly informative or redundant information. Therefore, it is necessary to have prior knowledge of the impact of different input data on the behaviour of output and network performance. Relevance is one of the key considerations for selecting an optimal set of inputs obtained through expert knowledge of data and model and having a reasonable input data set [38]. The predictors for each year have been finalized separately by forward stepwise linear regression model.

To evaluate and compare the performance of the developed models, two performance indices were used: the coefficient of determination (R^2) and the root mean square error (RMSE). The coefficient of determination shows the correlation between the observed values and the predicted values, with a value of one indicating a complete correlation between the observed and the calculated values. RMSE is the standard deviation of the residuals.

3. Results

3.1. Descriptive Statistics

Descriptive statistics of PM_{2.5} concentration and temporal predictors measured at the AQ stations for each year are summarized in Table 1 [18]. The table presents the average of each variable over all available stations (the number of stations varied over the years) and 61 days (August and September). The number, N, for each year represents the number of available stations for each year multiplied by the number of days with available data. The table shows that the highest daily concentration of PM_{2.5} (173 µg/m³) in 2017 when Alberta was dealing with the smoke-related to three large wildfires, and the second highest was measured in 2015 (111 µg/m³), when Alberta was under the smoke of the wildfires raging in the state of Washington (U.S.A.) (The event is known as the 2015 Pacific Northwest (PNW) wildfires, one of the most severe fire seasons in that region). The highest average PM_{2.5} level was 10 µg/m³, corresponding to the 2017 dataset. It can be seen that the lowest average and maximum daily PM_{2.5} concentrations was measured in 2016 (which was a very rainy season, explaining the low PM_{2.5} levels).

Table 1. Descriptive statistics of PM_{2.5} concentration, aerosol optical depth (AOD), and temporal predictive variables for each year (Var: variable, SD: standard deviation, TEMP: temperature, NDVI: Normalized difference vegetation index, WSP_{E-W}: wind speed in the east-west direction, WSP_{N-S}: wind speed in the north-south direction)

	Var.	Min	Mean	Max	SD		Min	Mean	Max	SD
2014 (N = 630)	PM _{2.5}	0.15	9	44	7.4	2015 (N = 772)	0.1	7.0	111	8.7
	AOD	0.001	0.2	2.3	0.3		0.02	0.18	2.6	0.3
	TEMP	2.3	16	28	5		4	15.0	29	5
	NDVI	0.3	0.5	0.8	0.1		0.28	0.54	0.5	0.1
	WSP _{E-W}	-28	-0.9	14	4.7		-5	-0.02	6	1.3
	WSP _{N-S}	-31	-1.6	27	5.6		-6	-0.03	5	1.3
2016 (N = 1480)	PM _{2.5}	0.04	4.5	35	2.61	2017 (N = 2329)	0.04	10	173	11.7
	AOD	0.0007	0.08	0.5	0.07		0.003	0.3	3.4	0.4
	TEMP	3	14	25	4		3	15.0	27	4
	NDVI	0.3	0.6	0.8	0.12		0.28	0.7	0.9	0.1
	WSP _{E-W}	-25	-0.3	22	5.4		-22	-5.0	14	4.1
	WSP _{N-S}	25	-3	22	5.2		-26	-2.0	23	6.4

Figure 2 presents the temporal variation of daily $PM_{2.5}$ concentration as boxplots throughout Alberta during the yearly study periods. The severe fire episodes that occurred in 2015 and 2017 are evident in Figure 2b,d. Figure 2b displays the elevated recorded $PM_{2.5}$ due to the 2015 Pacific Northwest PNW wildfires, discussed in [16,18]. Figure 2a (2014) also depicts some small variations in the level of measured $PM_{2.5}$ during the period; however, the maximum values ($<40 \mu\text{g}/\text{m}^3$) are much lower than the level of $PM_{2.5}$ recorded during the fire periods of 2015 and 2017.

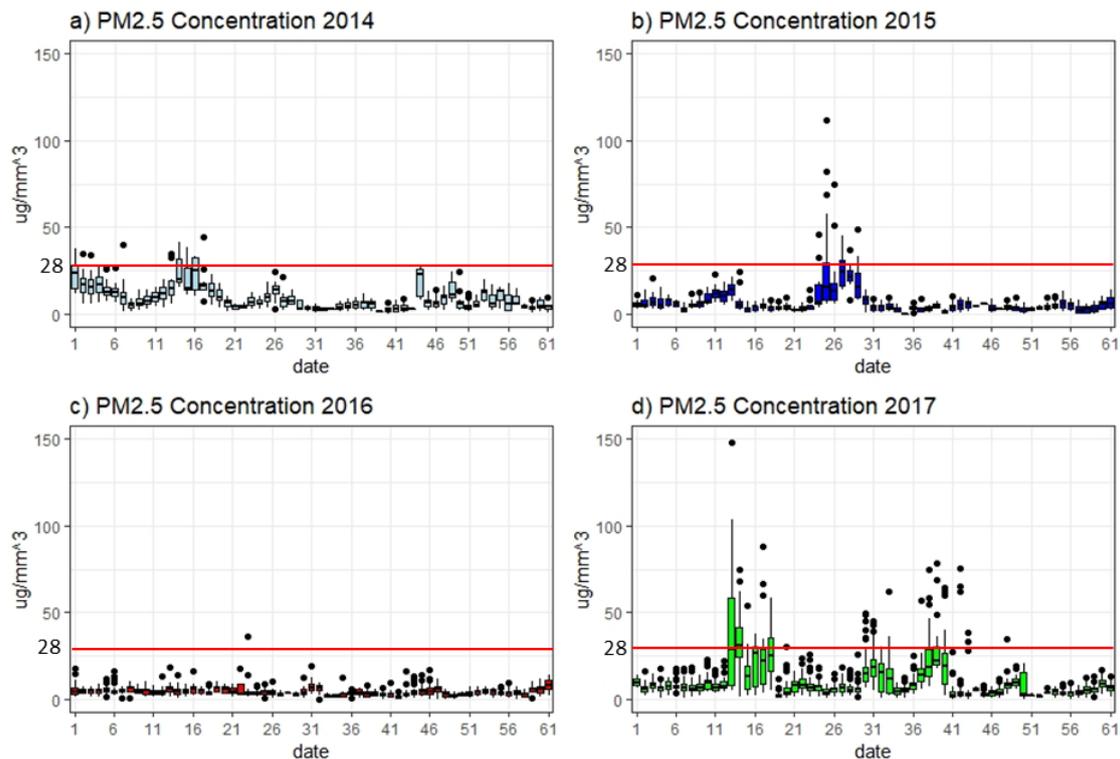


Figure 2. Temporal average daily $PM_{2.5}$ measurements over two months of the fire seasons in each year. The red line represents the three-year average of the annual 98th percentile of the daily 24-h average concentrations (<https://open.alberta.ca/dataset/684c673a-14a3-45b9-b5df-a158c8e83234/resource/0929eb23-9655-4046-b8df-0c0fc4e195fd/download/alberta-air-zones-report-2014-2016.pdf>).

The temporal variation of $PM_{2.5}$ concentration of 2016 (Figure 2c) shows that a rainy season can bring a very different variation of the $PM_{2.5}$ levels, and it constitutes an anomaly in our study: no significant elevated level of $PM_{2.5}$ was observed during this period, with all recorded values under $20 \mu\text{g}/\text{m}^3$. Consequently, the present study provided the opportunity of comparing the performance of an ANN model during fire seasons, including the high-level variations of $PM_{2.5}$, with a more normal (low level) variation of $PM_{2.5}$ concentration.

3.2. ANN Implementation Process

The results presented here have been obtained from the most suitable ANNs for each dataset amongst numerous ANN-built topologies that included different numbers of hidden layers, nodes in each layer, different training algorithms, test-train dataset, and variables (predictors). The ANN models were implemented, visualized, and tested using R [39].

Train and test dataset: To predict the spatial and temporal (regional and daily) $PM_{2.5}$ concentration, the ANNs were trained, temporally, for each of the four seasonal datasets, and, spatially, for each of the four regions, separately. This resulted in a total of eight models. Moreover, a single ANN was built to predict daily $PM_{2.5}$, which considered the pooled

dataset obtained by aggregating the four years of data for the whole study area. The topology of each network and the number of training samples varied for each dataset and was optimized using the trial-and-error approach. To validate the performance of the models, each dataset was randomly split into two datasets: 80% for training the model and 20% for testing it. As an additional validation approach, we also excluded data from one station and used that data to compare the predicted vs. recorded temporal PM_{2.5} values.

ANN architecture, training, and validation: Different ANN architectures were examined by changing the number of hidden layers, the quantity of the nodes on the hidden layers, the activation function, and the learning algorithm to identify the most suitable topology of ANN for each dataset separately. Figure 3 presents the final topology of the MLP used to predict daily PM_{2.5} in this study. To obtain the most optimal topology, up to two hidden layers and between 5 to 35 neurons in the hidden layer(s) were tested. The ideal number of neurons in the hidden layer(s) was decided by considering their RMSE and the point at which the network started to overfit. Based on the selected relevant variables (possible predictors), the input layer in all networks included eight neurons (nodes), as shown in Figure 3, with each node corresponding to one predictor. The training process of the final ANNs was repeated 100 times using randomly chosen training and test datasets to remove any effects of the random selection of the initial weights.

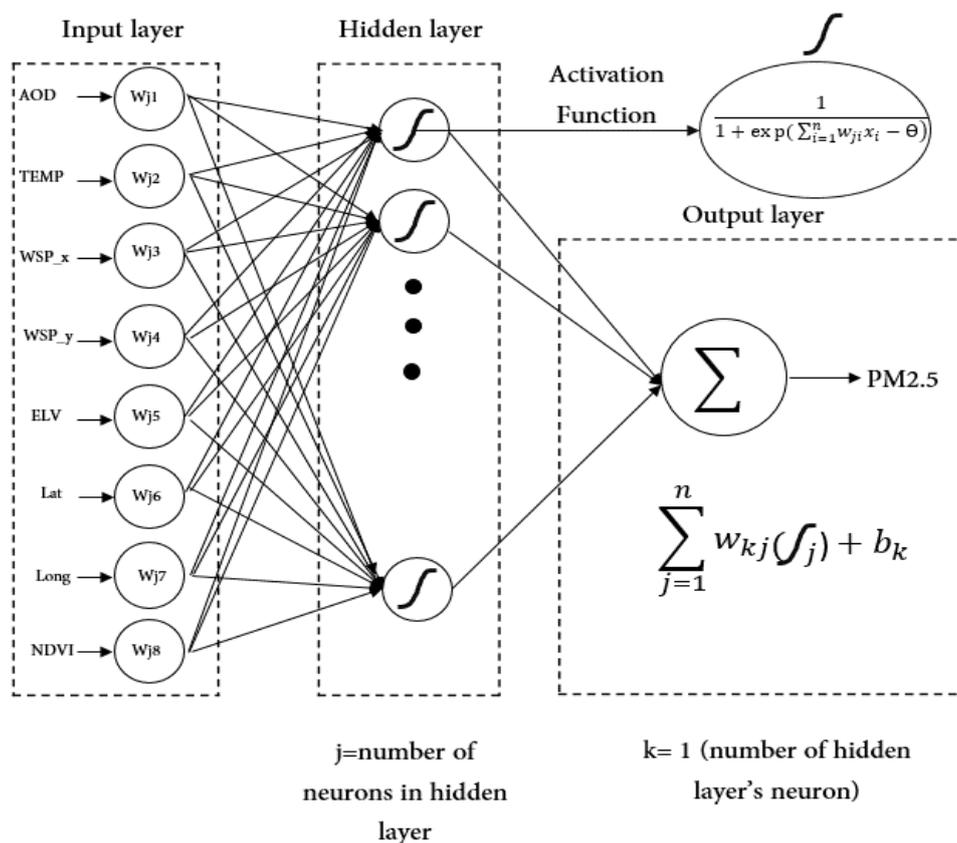


Figure 3. The structure of the ANN model in which the number of nodes on the hidden layer for each dataset was determined by the training process.

A logistic sigmoid activation function (Figure 3) was chosen for the hidden layer transformation function. We also scaled the values of each variable to a scalar between 0 and 1 to provide a similar impact to all input predictors.

In the case of the learning algorithm for MLP, different algorithms were examined, and the most suitable result was obtained by the resilient propagation (RPROP) learning function for all networks [40].

3.3. MLR Models and Temporal ANN Results

The predicted variables, coefficients, and the accuracy of the MLR models are presented in Table 2, and the performance of the developed temporal ANNs evaluated by R² and RMSE is provided in Table 3. The best MLR model was obtained for the 2014 dataset, and the weakest performance corresponds to the 2016 dataset. The best ANN performance based on RMSE and the R² test was achieved for the 2014 dataset, whereas the 2016 ANN showed the worst results.

Table 2. Results of the multiple linear regression (MLR) (MET_AOD_LUR) models, the benchmarks, for the yearly and pooled (four-year aggregated) data sets. The input data were normalized to 0–1 using min-max values, therefore the root mean square error (RMSE) is unitless.

	2014	R ² = 0.46 RMSE = 0.047	2015	R ² = 0.36 RMSE = 0.062
	Coef.	p-Value	Coef.	p-Value
Intercept	−0.10	9.52 × 10 ^{−7}	0.03	1.34 × 10 ^{−3}
AOD	0.40	<2 × 10 ^{−16}	0.35	<2 × 10 ^{−16}
TEMP	0.07	8.26 × 10 ^{−13}	0.10	1.33 × 10 ^{−14}
NDVI	−0.03	2.52 × 10 ^{−5}	−0.05	2.59 × 10 ^{−6}
WSP_X	-	-	-	-
WSP_Y	0.041	4.17 × 10 ^{−6}	-	-
lat	0.08	1.17 × 10 ^{−6}	-	-
long	0.06	1.07 × 10 ^{−7}	−0.03	1.93 × 10 ^{−2}
ELV	0.06	8.65 × 10 ^{−4}	-	-
	2016	R ² = 0.15 RMSE = 0.067	2017	R ² = 0.33 RMSE = 0.055
	Coef.	p-Value	Coef.	p-Value
Intercept	0.05	5.06 × 10 ^{−7}	0.02	7.09 × 10 ^{−3}
AOD	0.05	2.85 × 10 ^{−5}	0.30	<2 × 10 ^{−16}
TEMP	0.10	1.12 × 10 ^{−12}	0.06	1.06 × 10 ^{−15}
NDVI	−0.052	1.12 × 10 ^{−12}	−0.01	1.89 × 10 ^{−2}
WSP_X	-	-	−0.03	1.52 × 10 ^{−2}
WSP_Y	0.09	2.65 × 10 ^{−9}	−0.04	6.88 × 10 ^{−5}
lat	-	-	-	-
long	-	-	-	-
ELV	−0.02	2.89 × 10 ^{−2}	0.04	5.60 × 10 ^{−11}
	4-year dataset	R ² = 0.35 RMSE = 0.043		
	Coef.	p-Value		
Intercept	0.02	9.48 × 10 ^{−4}		
AOD	0.28	<2 × 10 ^{−16}		
TEMP	0.06	<2 × 10 ^{−16}		
NDVI	−0.02	5.70 × 10 ^{−14}		
WSP_X	−0.03	8.97 × 10 ^{−7}		
WSP_Y	−0.01	5.78 × 10 ^{−3}		
lat	0.01	6.15 × 10 ^{−2}		
long	-	-		
ELV	0.02	7.34 × 10 ^{−7}		

Table 3. Results including the R^2 for the training and test set obtained by the final ANN models and by the regression models for each year and pooled dataset (the best ANN model was recognized based on the lowest RMSE).

Temporal ANN	RMSE	Training R^2	Testing R^2	MLR Model	R^2	RMSE
2014	0.022	0.80	0.60		0.46	0.047
2015	0.039	0.63	0.45		0.36	0.062
2016	0.053	0.30	0.32		0.15	0.067
2017	0.036	0.61	0.60		0.33	0.055
All years	0.034	0.54	0.41		0.35	0.043

A comparison of R^2 and RMSE obtained by temporal ANNs and MLR models shown in Table 3 indicate that, for all datasets, the ANN model provided a higher prediction accuracy. The R^2 value of the four-year datasets increased by 30%, 25%, 113%, 81%, and 17%, respectively; the RMSE of all models also improved when the ANN was applied. Based on the multiple linear regression method, the 2014 model displayed the best performance. However, the accuracy of the model based on R^2 and RMSE was much lower than the neural network, which was also the best ANN for the remaining three years. The temporal ANNs results also showed a 10% to 92% increase in All_year_ANN R^2 , except for the 2016 ANN.

The scatterplots in Figure 4 display a visual representation of the measured versus the estimated $PM_{2.5}$ obtained by the training and testing dataset for each year. The regression and 1:1 reference lines are also presented in each plot. The regression equations are similar except for the 2016 data. The 1:1 line in the 2014 ANN plots suggests a low bias in the results. The poor performance of the 2016 ANN is also demonstrated by its scatterplot.

3.4. Spatial ANN Results

Similarly to the temporal model, the best spatial model was selected from the four regional ANNs using the lowest RMSE criterion. The results are presented in Figure 5. As shown in Figure 2, no significant wildfire episodes were recorded in 2016. Indeed, it was an unusually rainy summer in southern Alberta (<http://www.agriculture.alberta.ca/acis/climate-maps.jsp>), during which PM levels were unusually low (see Figure 2). As noted, the 2016 ANN yielded the worst temporal prediction. Hence, the 2016 data were not included when developing the spatial ANN for southern Alberta.

The spatial ANN models run on each region yielded significant results when 2016 was included (Figure 5): The northern regions, NE and NW, yielded the highest R^2 , (0.79 and 0.60 for testing, respectively, with training values approximately 0.70), and RMSE values approximately 0.020. Conversely, the central and southern regions yielded worse results, with testing R^2 at approximately 0.50 (training 0.53 and 0.62, respectively), and RMSE > 0.05.

Conversely, excluding the 2016 data, only the South region yielded significant results, with a better fit than excluding 2016, with testing $R^2 = 0.65$ (training $R^2 = 0.66$) and RMSE = 0.059.

Figure 5 displays the scatterplots of the predicted versus observed $PM_{2.5}$ values obtained by the training and testing datasets for each region, without 2016 for the southern region. As illustrated, in the northern areas where there were lower smoke-related $PM_{2.5}$ concentrations, the regression lines between the predicted and the measured values are closer to the 1:1 line.

3.5. Validation of Temporal ANN Predictions

To validate the temporal models and evaluate their prediction performance, the Lethbridge station was removed from the Southern region dataset (only for the 2015 and 2017 data), and the models were used to predict its recorded $PM_{2.5}$ concentrations for 2015 and 2017, as shown in Figure 6 Lethbridge AQ station, as the most southern AQ station in Alberta, was one of the smokiest stations during wildfire episodes in 2015 and 2017 due to its closer proximity to the source of fires originated in the U.S. Therefore, it was chosen

to validate the temporal models. It can be seen that the observed peaks of PM_{2.5} due to fire smoke were well predicted by the ANN models, even if the prediction was better for recorded levels below 40 µg/m³ and if the higher concentration (>40 µg/m³) of some of the peaks were underestimated.

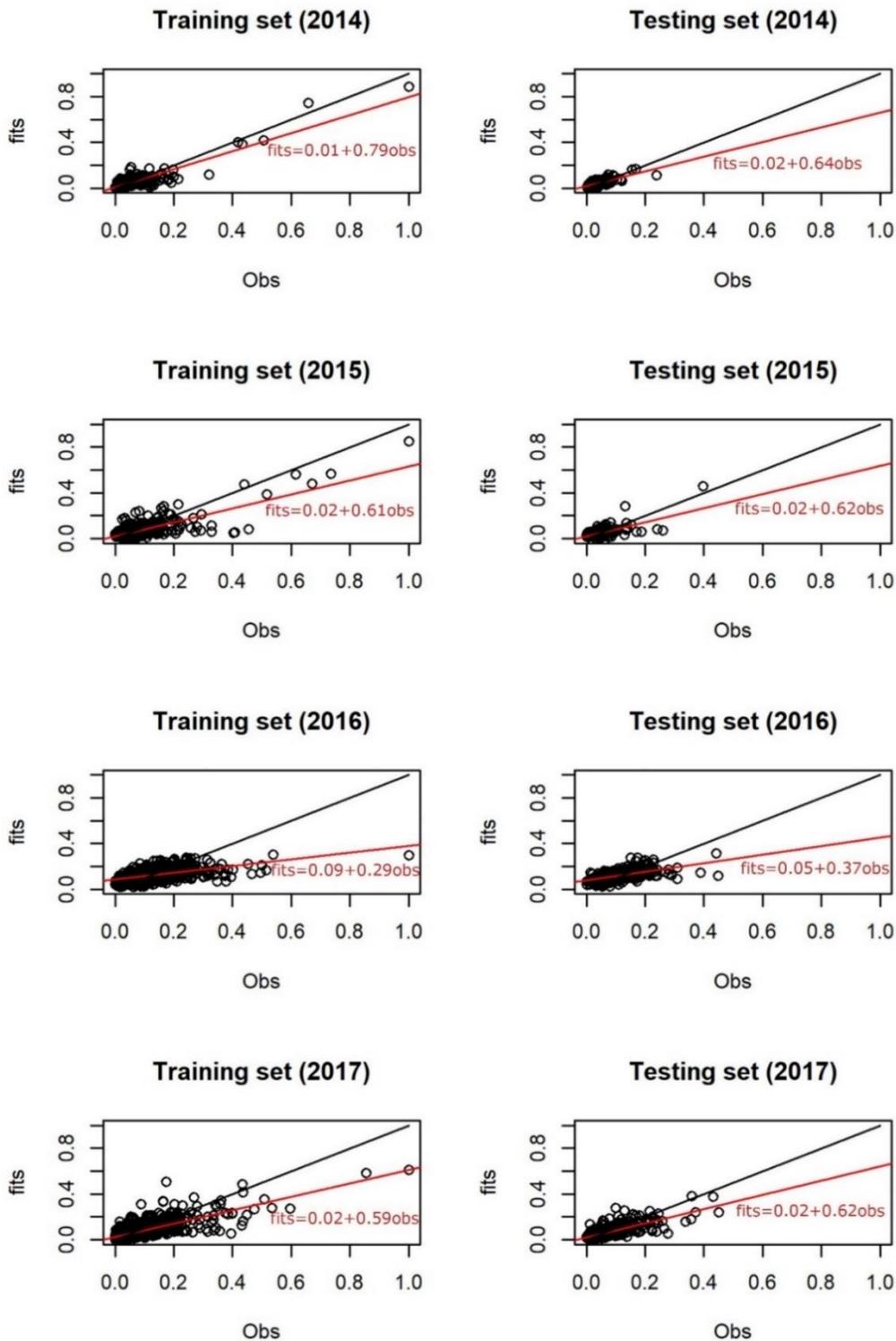


Figure 4. Scatter plots of measured versus predicted PM_{2.5} obtained by temporal ANN for 2014 to 2017 for test and training datasets separately.

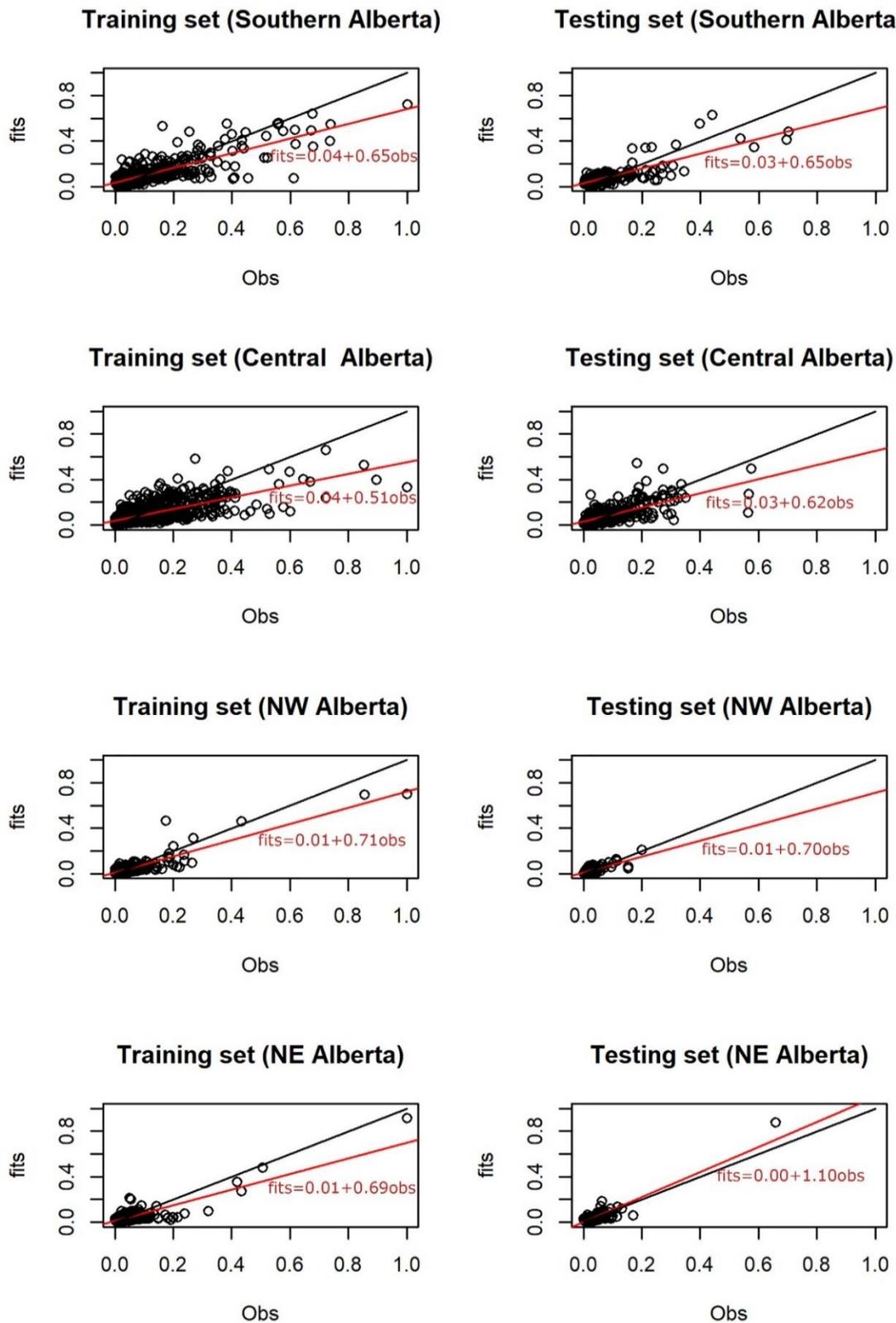


Figure 5. Scatter plots of measured versus predicted PM_{2.5} obtained by spatial ANN for all regions (Southern, Central, NW, and NE Alberta) for test and training datasets separately.

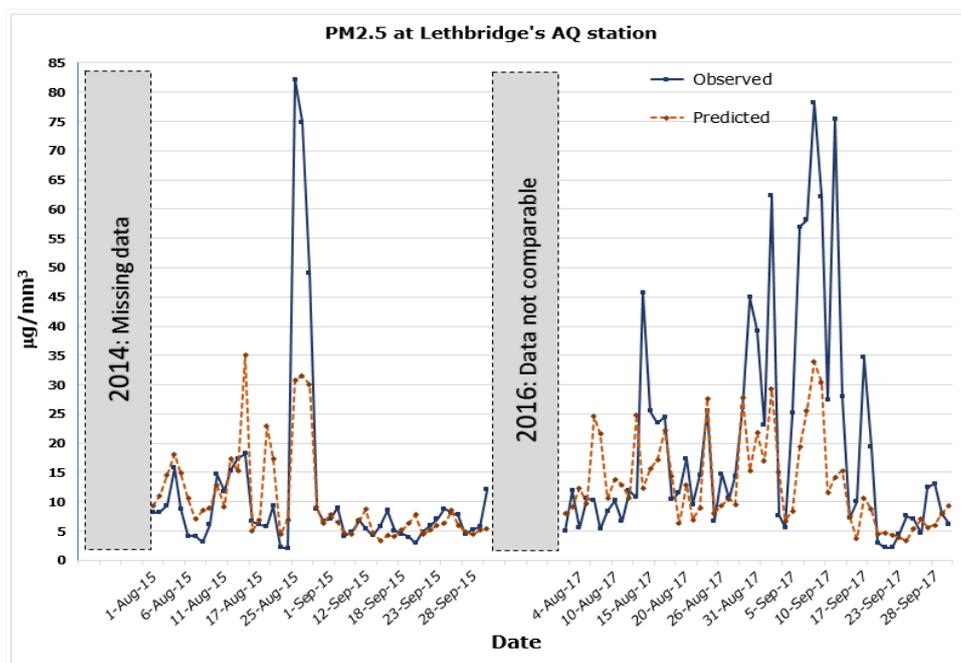


Figure 6. Observed and predicted PM_{2.5} using the southern Alberta ANN model at Lethbridge's air quality (AQ) station.

4. Discussion

The daily averaged PM_{2.5} concentration was proven to be more significant to exposed population health compared to hourly measurements [21]. Therefore, the present study performed different ANNs and multiple linear regression models to predict daily PM_{2.5} concentration in several wildfire episodes of 2014 to 2017 over the province of Alberta, Canada. Remote sensing data have been widely used to predict PM_{2.5} through regression models. However, regression models are more suitable for predicting mean values than the episodic events that usually lead to overestimations and underestimations of low and high concentration, respectively [21]. Therefore, to demonstrate the prediction improvement of ANN and, also, to show the constraints of linear regression models in predicting short-term PM_{2.5} during severe fire episodes, yearly MLRs were also developed, and the results were compared to ANN models.

Comparing the neural network models with MLRs indicated the robust performance of the machine learning approach for quantifying the spatiotemporal smoke-generated PM_{2.5}. The considerable increase in the ANN R² compared to MLR models (17% to 113%) illustrates the crucial capability of neural networks; however, the prediction results still need to be improved by applying finer resolution data, denser AQ stations, and also more relevant and informative predictors. The order of the best to the worst prediction model remained fixed for both approaches; the 2014 ANN and regression models resulted in the best prediction, and 2016 models presented the weakest ones, which could indicate the influence of quantity and distribution of sample data in prediction accuracy.

4.1. Temporal and Spatial Prediction

Figures 4 and 5 presented the predicted values versus observed values obtained by temporal and spatial models for both training and test dataset.

Dividing the dataset temporally based on each year also helped to improve the result, except for in the 2016 dataset. The ANNs developed by the three other datasets when fire smoke was present (2014, 2015, and 2017) significantly had an increased prediction accuracy over the pooled dataset ANN: the models' correlation coefficient increased from 0.41 for the pooled dataset to 0.6 for 2014 and 2017 and to 0.45 for 2015, i.e., an increase of 10% to 46%. The relatively poor performance of both the MLR and ANN models for the

2016 dataset was due to low $PM_{2.5}$ concentration, as discussed earlier. Excluding the 2016 dataset from the spatial Southern dataset also increased the R^2 of the test model by 22%.

The province of Alberta is relatively large and all AAZs do not experience fire smoke on the same day, therefore regional ANNs were performed to improve the accuracy of the ANNs by dividing the study area into four regions and training the ANN for each region individually. The higher R^2 and lower RMSE of the spatial models for the test dataset (Figure 5) indicated that all spatial ANNs (regional) improved compared to by the pooled ANN (provincial) shown in Table 3. The R^2 showed an increment of 29%, 24%, 46%, and 92% for the southern, central, NW, and NE regions, respectively. Therefore, having enough numbers of the training dataset (i.e., ground measurements) enables the development of regional models where the spatial and meteorological criteria in the location of the ground stations are closer to each other and can result in better prediction. Regarding the number of AQ stations in the present study, performing separate ANNs for each AAZ was not possible.

The results also show higher improvements in R^2 and RMSE by spatial ANNs compared to temporal ANNs. That demonstrates the geographic dispersion as well as the temporal dispersion of fire smoke, implying that more complex spatiotemporal ANNs would perform better; however, it would require more data. The obtained results of the 2017 ANN were also compared to our previous study [16], where a linear mixed effect (LME) model was applied by developing nested period-zone specific random effects of the AOD- $PM_{2.5}$ relationship to predict daily $PM_{2.5}$ concentration for the same study area and period. To be able to compare the ANN model with the developed LME model, the same number of $PM_{2.5}$ measurements were used, and most of the predictors remained fixed. However, due to the different model's requirements and to train the ANN model, a few more predictors had to be added to the neural network. The R^2 of the nested LME model was 0.74 compared to 0.57 for the ANN model. Comparing the obtained results of both models demonstrated that the nested LME model performed better to predict daily fire-generated $PM_{2.5}$. However, the ANN prediction performance was also relatively good. Furthermore, ANNs are relatively more straightforward to implement and need fewer hypotheses related to data distribution.

We also need to mention that when the $PM_{2.5}$ concentration was very low and under the background level (e.g., in the 2016 dataset), the ANN model, like the linear regression model, poorly performed in modelling the AOD- $PM_{2.5}$ relationship. Therefore, it seems that AOD may not present a reasonable prediction at lower $PM_{2.5}$ concentrations, because the predictive ability of the AOD- $PM_{2.5}$ model will decrease.

4.2. Limitations

Based on the results, a relatively weak correlation between AOD and $PM_{2.5}$ was demonstrated in the low correlation coefficient of the models, even when adding ancillary variables to the AOD- $PM_{2.5}$ relationship did not help improve the predictive accuracy of any of the MLR models. Despite the numerous advantages of remote sensing products, there are also some weaknesses in applying satellite data (Kumar 2010).

Associations between AOD values and $PM_{2.5}$ measurements were also affected by different factors, such as the removal of dense smoke plumes generated by wildfires by the MODIS AOD algorithm, or mistaking cloud as the aerosol, and, also, it is highly influenced by the vertical distribution of AOD. Another possibility for a relatively poor AOD- $PM_{2.5}$ relationship can be due to high relative humidity that leads to an increase in the size of some hygroscopic particles (ammonium sulphate and ammonium nitrate). Under this situation, the inflated particles will affect the light extinction efficiencies, and results in a higher value of AOD while the $PM_{2.5}$ concentration does not change [21].

The present study suffers from these limitations too. The weak relationship between AOD and $PM_{2.5}$ was probably due to the uncertainty of the satellite products. Relatively coarse resolution of the AOD images and the spatial and temporal mismatch between the $PM_{2.5}$ measurements and AOD values are two of the possible limitations. The $PM_{2.5}$

measurements and other meteorological predictors used in this study are averages of 24-h recorded values, while the AOD for each pixel was a single value obtained at the Aqua/Terra overpass time. Spatial mismatching could also be another source of uncertainty in the AOD–PM_{2.5} relationship; each AOD pixel in our study covered a 10-by-10 km area, while PM_{2.5} data were point-based measurements aggregated across 24 h. However, the ANN models were able to improve the results significantly and could reduce the influence of these limitations.

5. Conclusions

Quantitative spatial information of PM_{2.5} concentrations during elevated smoke periods due to wildfire is very useful for regulating air quality and monitoring fire smoke-related problems of exposed, vulnerable populations. Satellite-based data provide an informative tool for monitoring PM_{2.5} with a relatively high spatial and temporal resolution. AOD derived from satellite images is widely used to predict surface PM_{2.5} concentrations. In the present study, ANNs have been utilized to predict daily PM_{2.5} concentrations generated from wildfires over the province of Alberta, Canada, during the fire seasons of 2014 to 2017.

Different ANNs were developed for the different datasets, including: (i) a pooled data set combining all four years of data for the whole province; (ii) yearly datasets for the entire study area; and (iii) regional four-year datasets for the southern, central, northwest, and northeast regions of the province. To improve the accuracy of the models, and also to investigate the effects of the low level of PM_{2.5} concentrations in ANN models, the 2016 dataset, which had the lowest level of PM_{2.5} concentration, were excluded from the southern Alberta dataset, and the model was trained again. The spatial resolution of the current study was 10 × 10 km, which was determined by the spatial resolution of MODIS AOD images.

The models were validated by dividing the data into test and training datasets and running each network 100 times. We also validated the temporal ANN using one of the AQ stations (Lethbridge), which was excluded from the training process. Spatial and temporal ANNs were trained by grouping the data based on each year (temporal ANN) and by dividing the entire study area based on neighbouring AAZs (spatial ANN). The temporal and spatial ANNs have improved the prediction accuracy of the pooled one.

The ANNs provided acceptable predictions of PM_{2.5}, in spite of the sparse distribution of ground AQ monitoring, gaps and uncertainty in AOD images, and also the complexity of the problem due to the fact that PM_{2.5} during the study period was generated from different sources. ANNs significantly improved linear regression models' performance in predicting PM_{2.5} concentration using AOD images, temporally and spatially.

Author Contributions: Conceptualization, M.M., S.B, I.C., and B.F.; Data curation, M.M. and I.C.; Formal analysis, M.M., I.C. and S.B.; Investigation M.M., S.B, I.C., B.F.; Methodology M.M., S.B., I.C., and B.F.; Software, M.M. and I.C.; Supervision, S.B.; Validation, M.M., S.B., I.C. and B.F.; Writing—original draft, M.M.; Writing—review & editing, M.M., S.B., I.C., and B.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Open access data have been used.

Acknowledgments: We would like to acknowledge Roland Ngom and Vineet Saini for their support and contribution to the project. M.M. wishes to thank Eyes High Doctoral Recruitment Scholarship for supporting her doctoral work. S.B. wishes to thank the Canadian Institutes for Health Research (CIHR) Institute for Population and Public Health, as well as the O'Brien Institute for Public Health for funding her research on air pollution and public health. We are grateful to our colleagues and members of the Geography of Health research group of the O'Brien Institute for Public Health for their advice and insightful discussions. Finally, we would like to acknowledge the anonymous reviewers for their constructive criticism and truly helpful comments.

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

References

1. Youssouf, H.C.; Roblou, L.L.; Assamoi, E.M.; Salonen, R.O.; Maesano, C. Quantifying Wild Fire Exposure for Investigating Health-Related Effects. *Atmos. Environ.* **2014**, *97*, 239–251. [CrossRef]
2. World Health Organization. Wildfire Smoke a Guide for Public Health Officials. 2016. Available online: https://www3.epa.gov/airnow/wildfire_may2016.pdf (accessed on 6 January 2021).
3. Black, C.; Tesfaigzi, Y.; Bassein, J.A.; Miller, L.A. Wild Fire Smoke Exposure and Human Health: Significant Gaps in Research for a Growing Public Health Issue. *Environ. Toxicol. Pharmacol.* **2017**, *55*, 186–195. [CrossRef]
4. Cascio, W.E. Wildland Fire Smoke and Human Health. *Sci. Total Environ.* **2018**, *624*, 586–595. [CrossRef]
5. Liu, J.C.; Pereira, G.; Uhl, S.A.; Bravo, M.A.; Bell, M.L. A Systematic Review of the Physical Health Impacts from Non-Occupational Exposure to Wild Fire Smoke. *Environ. Res.* **2015**, *136*, 120–132. [CrossRef]
6. Lv, B.; Hu, Y.; Chang, H.H.; Russell, A.G.; Cai, J.; Xu, B.; Bai, Y. Daily Estimation of Ground-Level PM_{2.5} Concentrations at 4 km Resolution over Beijing-Tianjin-Hebei by Fusing MODIS AOD and Ground Observations. *Sci. Total Environ.* **2017**, *580*, 235–244. [CrossRef]
7. Zheng, Y.; Zhang, Q.; Liu, Y.; Geng, G.; He, K. Estimating Ground-Level PM_{2.5} concentrations over Three Megalopolises in China Using Satellite-Derived Aerosol Optical Depth Measurements. *Atmos. Environ.* **2016**, *124*, 232–242. [CrossRef]
8. Xie, Y.; Wang, Y.; Zhang, K.; Dong, W.; Lv, B.; Bai, Y. Daily Estimation of Ground-Level PM_{2.5} Concentrations over Beijing Using 3 Km Resolution MODIS AOD. *Environ. Sci. Technol.* **2015**, *49*, 12280–12288. [CrossRef]
9. Chudnovsky, A.A.; Koutrakis, P.; Kloog, I.; Melly, S.; Nordio, F.; Lyapustin, A.; Wang, Y.; Schwartz, J. Fine Particulate Matter Predictions Using High Resolution Aerosol Optical Depth (AOD) Retrievals. *Atmos. Environ.* **2014**, *89*, 189–198. [CrossRef]
10. Chang, H.H.; Hu, X.; Liu, Y. Calibrating MODIS Aerosol Optical Depth for Predicting Daily PM_{2.5} Concentrations via Statistical Downscaling. *J. Expo. Sci. Environ. Epidemiol.* **2014**, *24*, 398–404. [CrossRef]
11. Chen, B.B.; Sverdluk, L.G.; Imashev, S.A.; Solomon, P.A.; Lantz, J.; Schauer, J.J.; Shafer, M.M.; Artamonova, M.S. Empirical Relationship between Particulate Matter and Aerosol Optical Depth over Northern Tien-Shan, Central Asia. *Air Quality. Atmos. Health* **2013**, *6*, 385–396. [CrossRef]
12. Kloog, I.; Nordio, F.; Coull, B.A.; Schwartz, J. Incorporating Local Land Use Regression and Satellite Aerosol Optical Depth in a Hybrid Model of Spatiotemporal PM_{2.5} Exposures in the Mid-Atlantic States. *Environ. Sci. Technol.* **2012**, *46*, 11913–11921. [CrossRef]
13. Lee, H.J.; Coull, B.A.; Bell, M.L.; Koutrakis, P. Use of Satellite-Based Aerosol Optical Depth and Spatial Clustering to Predict Ambient PM_{2.5} Concentrations. *Environ. Res.* **2012**, *118*, 8–15. [CrossRef]
14. Chudnovsky, A.A.; Lee, H.J.; Kostinski, A.; Kotlov, T.; Koutrakis, P. Prediction of Daily Fine Particulate Matter Concentrations Using Aerosol Optical Depth Retrievals from the Geostationary Operational Environmental Satellite (GOES). *J. Air Waste Manag. Assoc.* **2012**, *62*, 1022–1031. [CrossRef]
15. Lee, H.J.; Liu, Y.; Coull, B.A.; Schwartz, J.; Koutrakis, P. A Novel Calibration Approach of MODIS AOD Data to Predict PM_{2.5} Concentrations. *Atmos. Chem. Phys.* **2011**, *11*, 7991–8002. [CrossRef]
16. Mirzaei, M.; Bertazzon, S.; Couloigner, I. Estimation of local daily PM_{2.5} concentration during wildfire episodes: Integrating MODIS AOD with multivariate linear mixed effect (LME) models. *Air Quality. Atmos. Health* **2020**, *13*, 173–185. [CrossRef]
17. Geng, G.; Murray, N.L.; Tong, D.; Joshua, S.; Fu, X.H.; Lee, P.; Meng, X.; Chang, H.H.; Liu, Y. Satellite-Based Daily PM_{2.5} Estimates during Fire Seasons in Colorado. *J. Geophys. Res.* **2018**, *123*, 8159–8171. [CrossRef]
18. Mirzaei, M.; Bertazzon, S.; Couloigner, I. Modeling Wildfire Smoke Pollution by Integrating Land Use Regression and Remote Sensing Data: Regional Multi-Temporal Estimates for Public Health and Exposure Models. *Atmosphere* **2018**, *9*, 335. [CrossRef]
19. Reid, C.E.; Jerrett, M.; Petersen, M.L.; Gabriele, G.P.; More, P.E.; Tager, I.B.; Ra, S.M.; Balmes, J.R. Spatiotemporal Prediction of Fine Particulate Matter During the 2008 Northern California Wild Fire Using Machine Learning. *Environ. Sci. Technol.* **2015**. [CrossRef] [PubMed]
20. Gardner, M.W.; Dorling, S.R. Artificial neural networks (the multilayer perceptron)—A review of applications in the atmospheric sciences. *Atmos. Environ.* **1998**, *32*, 2627–2636. [CrossRef]
21. Gupta, P.; Christopher, S.A. Particulate Matter Air Quality Assessment Using Integrated Surface, Satellite, and Meteorological Products: 2. A Neural Network Approach. *J. Geophys. Res.* **2009**, *114*, 1–14. [CrossRef]
22. Hu, X.; Belle, J.H.; Meng, X.; Wildani, A.; Waller, L.A.; Strickland, M.J.; Liu, Y. Estimating PM_{2.5} Concentrations in the Conterminous United States Using the Random Forest Approach. *Environ. Sci. Technol.* **2017**, *51*, 6936–6944. [CrossRef] [PubMed]
23. Tu, J.V. Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *J. Clin. Epidemiol.* **1996**, *49*, 1225–1231. [CrossRef]
24. Grivas, G.; Chaloulakou, A. Artificial neural network models for prediction of PM₁₀ hourly concentrations, in the Greater Area of Athens, Greece. *Atmos. Environ.* **2006**, *40*, 1216–1229. [CrossRef]
25. Hebb, D.O. *The Organization of Behavior: A Neuropsychological Theory*; John Wiley: New York, NY, USA, 1949.
26. Shanmuganathan, S. Artificial Neural Network Modelling. In *Studies in Computational Intelligence*; Springer International Publishing: Cham, Switzerland, 2016; Volume 628. [CrossRef]
27. Perez, P.; Reyes, J. An Integrated Neural Network Model for PM₁₀ Forecasting. *Atmos. Environ.* **2006**, *40*, 2845–2851. [CrossRef]
28. Ordieres, J.B.; Vergara, E.P.; Capuz, R.S.; Salazar, R.E. Neural Network Prediction Model for Fine Particulate Matter (PM_{2.5}) on the US-Mexico Border in Reza (Chihuahua), El Paso (Texas) and Ciudad Juarez. *Environ. Model. Softw.* **2005**, *20*, 547–559. [CrossRef]

29. Jiang, D.; Zhang, Y.; Hu, X.; Zeng, Y.; Tan, J.; Shao, D. Progress in Developing an ANN Model for Air Pollution Index Forecast. *Atmos. Environ.* **2004**, *38*, 7055–7064. [[CrossRef](#)]
30. Dorling, S.R.; Foxall, R.J.; Mandic, D.P.; Cawley, G.C. Maximum Likelihood Cost Functions for Neural Network Models of Air Quality Data. *Atmos. Environ.* **2003**, *37*, 3435–3443. [[CrossRef](#)]
31. Perez, P.; Trier, A.; Reyes, J. Prediction of PM Concentrations Several Hours in Advance Using Neural Networks in Santiago, Chile. *Atmos. Environ.* **2000**, *34*, 1189–1196. [[CrossRef](#)]
32. Comrie, A.C. Comparing Neural Networks and Regression Models for Ozone Forecasting. *J. Air Waste Manag. Assoc.* **1997**, *47*, 653–663. [[CrossRef](#)]
33. AAC Alberta Airsheds Council. 2006. Available online: <https://www.albertaairshedsCouncil.ca/> (accessed on 6 January 2021).
34. ACIS. Agroclimatic Atlas Map for Wind Speed, Data Provided by Alberta Agriculture and Forestry. 2019. Available online: <https://agriculture.alberta.ca/acis> (accessed on 29 December 2020).
35. NEO. NEO, Nasa Earth Observations. 2017. Available online: https://neo.sci.gsfc.nasa.gov/view.php?datasetId=MODAL2_D_AER_OD&date=2018-08-01 (accessed on 6 January 2021).
36. DMTI. The Gold Standard Canada's Most Complete and Accurate Mapping Data, DMTI Spatial. 2010. Available online: <https://www.dmtispatial.com/> (accessed on 22 August 2018).
37. Lary, D.J.; Faruque, F.S.; Malakar, N.; Moore, A.; Roscoe, B.; Adams, Z.L.; Egelston, Y. Estimating the global abundance of ground level presence of particulate matter (PM_{2.5}). *Geospat Health* **2014**, *8*, 611–630. [[CrossRef](#)]
38. May, R.; Dandy, G.; Maier, H. Review of Input Variable Selection Methods for Artificial Neural Networks. In *Artificial Neural Networks—Methodological Advances and Biomedical Applications*; InTech: London, UK, 2011; pp. 19–44. [[CrossRef](#)]
39. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2019.
40. Riedmiller, M.A.; Braun, H. A direct adaptive method for faster backpropagation learning: The RPROP algorithm. In Proceedings of the IEEE International Conference on Neural Networks, San Francisco, CA, USA, 28 March–1 April 1993; Volume 1, pp. 586–591.