Neural Networks 121 (2020) 396–408

Contents lists available at ScienceDirect

Neural Networks

journal homepage: www.elsevier.com/locate/neunet

Using a deep convolutional neural network to predict 2017 ozone concentrations, 24 hours in advance

Alqamah Sayeed, Yunsoo Choi *, Ebrahim Eslami, Yannic Lops, Anirban Roy, Jia Jung

Department of Earth and Atmospheric Sciences, University of Houston, TX 77004, United States of America

ARTICLE INFO

Article history:
Received 16 November 2018
Received in revised form 10 September 2019
Accepted 22 September 2019
Available online 28 September 2019

Keywords:
Convolutional
Neural network
Ozone prediction
Real-time forecasting
Deep learning

ABSTRACT

In this study, we use a deep convolutional neural network (CNN) to develop a model that predicts ozone concentrations 24 h in advance. We have evaluated the model for 21 continuous ambient monitoring stations (CAMS) across Texas. The inputs for the CNN model consist of meteorology (e.g., wind field, temperature) and air pollution concentrations (NO_x and ozone) from the previous day. The model is trained for predicting next-day, 24-hour ozone concentrations. We acquired meteorological and air pollution data from 2014 to 2017 from the Texas Commission on Environmental Quality (TCEQ). For 19 of the 21 stations in the study, results show that the yearly index of agreement (IOA) is above 0.85, confirming the acceptable accuracy of the CNN model. The results also show the model performed well, even for stations with varying monthly trends of ozone concentrations (specifically CAMS-012, located in El-Paso, and CAMS-013, located in Fort Worth, both with IOA=0.89). In addition, to ensure that the model was robust, we tested it on stations where fewer meteorological variables are monitored. Although these stations have fewer input features, their performance is similar to that of other stations. However, despite its success at capturing daily trends, the model mostly underpredicts the daily maximum ozone, which provides a direction for future study and improvement. As this model predicts ozone concentrations 24 h in advance with greater accuracy and computationally fewer resources, it can serve as an early warning system for individuals susceptible to ozone and those engaging in outdoor activities.

© 2019 Elsevier Ltd. All rights reserved.

1. Introduction

Ozone is a criteria pollutant generated by the photochemical reaction between nitrogen oxides (NO_x) and volatile organic compounds (VOCs) in the atmosphere (U.S. EPA, 2006). Depending on its concentrations, ozone can have serious health implications. In general, it is advisable that hourly ozone concentration limits should not exceed 80 ppbv and/or 50–60 ppbv for a maximum daily eight-hour average (MDA-8) (Ayres, Maynard, & Richards, 2006; Taylan, 2017). In urban areas, short-term exposure to ozone can cause headaches, chest pains, sore throats, and coughing and decreased lung function, depending on the concentration of ozone (usually >150–300 ppbv). At elevated concentrations of ozone above 100 ppbv, living beings are more susceptible to bacterial infections (Jacobson, 2005). According to Bell, Peng, and Dominici (2006), exposure to high concentrations of ozone and long-term exposure to low concentrations of ozone adversely affect human health. Similarly, Mills et al. (2007) claimed that long-term exposure to more than 40 ppbv of ozone can damage crops and ecosystems.

The severe implications of high tropospheric ozone concentrations to human health and the environment necessitate prior reporting of such concentrations as well as the time of the day they occur. Numerical air quality modeling, also referred to as chemical transport modeling (CTM), is often used for this purpose. Several efforts have been devoted to reproducing in-situ ozone concentrations. The most widely used air pollution chemical transport model is the Community Multiscale Air Quality (CMAQ) model, developed by USEPA (Byun & Schere, 2006). Byun et al. (2007) performed CMAQ simulation for base and HRVOC (a highly reactive volatile organic compound) emissions in the Houston-Galveston-Brazoria region (HGB), Texas, USA during a five-day summer period of TexAQS-2000 and found that HRVOC produced better results (slope, a = 0.81 ppbv; intercept, b = 5.48 ppbv, r^2 = 0.76). Misenis and Zhang (2010) used the WRF/Chem model during the same period and region. They ran simulations in various configurations and reported an hourly ozone prediction mean bias range between 3.8 ppbv and 11.0 ppbv. Using the CMAQ modeling system, the NAQFC (National Air Quality Forecast Capability) project (Chai et al., 2013) generated operational and experimental real-time ozone predictions for the

* Correspondence to: SR 426F, Department of Earth and Atmospheric Sciences, University of Houston, TX 77004, United States of America.
E-mail address: ychoi23@central.uh.edu (Y. Choi).

https://doi.org/10.1016/j.neunet.2019.09.033
0893-6080/© 2019 Elsevier Ltd. All rights reserved.

United States in 2010 and found that the model overestimates by 5.6 ppbv annually for the contiguous US (CONUS) and that the RMSE was 15.4 ppbv. They also found out that among the six regions of the study, the southeastern US had a maximum bias of 10.5 ppbv and the lower middle US (Texas, Oklahoma, Arkansas, Louisiana) had a minimum bias of 3.7 ppbv. Czader et al. (2015) developed another chemical transport model, STOPS (v1.0), a hybrid Eulerian–Lagrangian-based modeling tool, and compared its results with those of the CMAQ for August 25, 28, and 30, in the year 2000, for the Houston region. They found that the mean bias for surface ozone mixing ratios varied between −0.03 and −0.78 ppbv and the slope varied between 0.99 and 1.01 for several configurations. Another study by Pan, Choi, Roy, and Jeon (2017) simulated NO_x and ozone for the Houston region for September 5–14, 2013, with a spatial resolution of 4 km and 1 km and found correlations for ozone concentrations in the range of 0.79 to 0.87 and an index of agreement (IOA) of 0.74 to 0.86. The model, however, generated stronger correlations and IOA mean bias (MB): 10 to 14.9 ppbv and a mean absolute error (MAE) of 10.7 to 15.2 ppbv. Despite their greater domain coverage and reasonable results, the above models consume significant computational resources and time (e.g., Zhang, Bocquet, Mallet, Seigneur, & Baklanov, 2012). This necessitates the use of artificial intelligence techniques, such as deep neural networks, which are considerably more efficient and consume fewer computational resources and time (e.g., Fernando et al., 2012).

Generally, such model uses an artificial neural network (ANN) that can be trained from historical events that form the input and output set (perceived outcome based on given inputs). Since all atmospheric phenomena are interrelated, the ANN then predicts a future event(s) based on a given set of new inputs (or unseen inputs) (Bengio, 2009; Marr, 1977). Hoshyaripour et al. (2016) developed an FS–ANN model, compared it with WRF-Chem, and evaluated it in two stations in Sao Paulo (Brazil) between August 5 and 20, 2012. They found that while the daily mean IOA ranged from 0.39 to 0.59 (at various locations) from FS–ANN and 0.54 to 0.68 from WRF-Chem, the daily peak value was slightly higher (0.53 to 0.78 from FS–ANN and 0.63 to 0.67 from WRF-Chem). These results also showed a mean daily bias of between −3.49 and 1.60 ppbv from FS–ANN and between −8.05 and 4.25 ppbv from WRF-Chem. They concluded that WRF/Chem produces better results than FS–ANN, but the latter is computationally faster and cheaper.

Prasad et al. (2016) developed an adaptive neuro-fuzzy inference system (ANFIS) for Howrah City, India, for the years 2009 to 2011, reported an IOA of 0.81 and R^2 of 0.51 for one-day advance forecasting of ozone concentrations. Biancofiore et al. (2015) used a recurrent neural network (RNN) to predict one-, three-, six-, 12-, 24-, and 48-h ozone concentrations at an observation station in Pescara, Italy, in 2005. Although their model performed reasonably, with comparatively stronger correlation coefficients for the one-, three-, 24-, and 48-h concentrations (i.e., 0.97, 0.89, 0.86, and 0.83 respectively), it yielded poor correlation coefficients for the six- and 12-h concentrations (i.e., 0.78 and 0.77, respectively). The main reason for the poor six- and 12-h correlations is that both meteorological and pollutant criteria differ significantly (or reverse in the case of 12-h) from the 0th h while the stronger 24- and 48-h correlation coefficients occur at the same time of day. The other caveat in their model is that they predicted one specific hour of the day per iteration.

Although the above models are fast, they have either one or both of the following issues: (i) Most predict a single value (single hour of day, daily the mean, the daily maximum, or MDA-8) as in Biancofiore et al. (2015), and/or, (ii) their predictions are inaccurate, as in Hoshyaripour et al. (2016). A fast, stable, and accurate model can overcome these problems. To develop such a model, we have used the deep architecture of CNN (Krizhevsky, Sutskever, & Hinton, 2012; Lawrence, Gills, Tsoi, & Back, 1997; LeCun & Bengio, 1995). CNN is suitable for our study because of two qualities: (i) it is capable of understanding the complex features of input variables by applying a convolution using multiple filters and kernels (LeCun & Bengio, 1995); and (ii) since data (both meteorological and pollutant data) exhibit temporal coherence that CNN preserves by convolution on adjacent inputs only (Lawrence et al., 1997), it provides the desired accuracy in the results. Therefore, in this study, we begin by developing an artificial intelligence model based on a deep convolutional neural network to predict next-day 24-h ozone concentrations at a station and then evaluate the model based on daily, weekly, monthly and seasonal values for the year 2017. (Note: The model was trained for the years 2014 to 2016).

2. Material and methods

We used a five-layer deep CNN architecture model to produce real-time 24-h predictions of ozone. The main purpose of using a multilayer CNN was to increase computational efficiency over traditional numerical models and various deep learning models. When compared with multilayer perceptron (MLP), it involves more thorough and complex calculations to preserve nonlinear characteristics of the input and output features. We performed several tests to determine the optimal number of layers (in this case, five) and achieve the lowest mean squared error (MSE). This technique involved site selection, observations, and model definition, training, and prediction. These steps are explained in the following subsection.

2.1. Site selection and observation

Anthropogenic precursors, along with meteorological conditions, significantly affect regional air quality (Baklanov, Korsholm, Mahura, Petersen, & Gross, 2008; Damo & Icka, 2012; Taylan, 2013a). Thus, we selected only stations (see Fig. 1 for station location and Table-A1 in the supplementary document for station details) in Houston with both ozone and NOx data available and at least three meteorology measurements monitored for at least four years (i.e., 2014 to 2017). In addition, we selected seven stations from other cities in Texas to evaluate the model performance outside of Houston. We obtained a four-year (2014–2017) observational data from the Texas Commission on Environmental Quality (TCEQ) website, which provides a variety of meteorological (e.g., wind speed, wind direction, temperature, pressure, precipitation, dew point temperature, relative humidity, solar radiation) and air pollutant (i.e., NO_x and ozone) data.

2.2. Model definition

The CNN, as introduced by LeCun and Bengio (1995), is generally used for image classification. However, since its inception, CNN has improved significantly and has been used in various applications from speech recognition to object detection (LeCun & Bengio, 1995). Deep CNN consists of several stacked neuron layers: (1) a convolutional layer, (2) a pooling layer, and (3) fully connected layers (Krizhevsky et al., 2012). These layers build a hierarchical and distributed network. The convolutional layer consists of neurons that are stacked together and respond only to the overlapping region instead of the whole signal. Mathematically, convolution is the integral measuring of the extent to which two functions overlap as one passes over the other. The pooling layer eliminates features with similar attributes, thus reducing the computational burden. In this layer, features are either averaged (i.e., average pooling) or maximum (i.e., MaxPooling) over

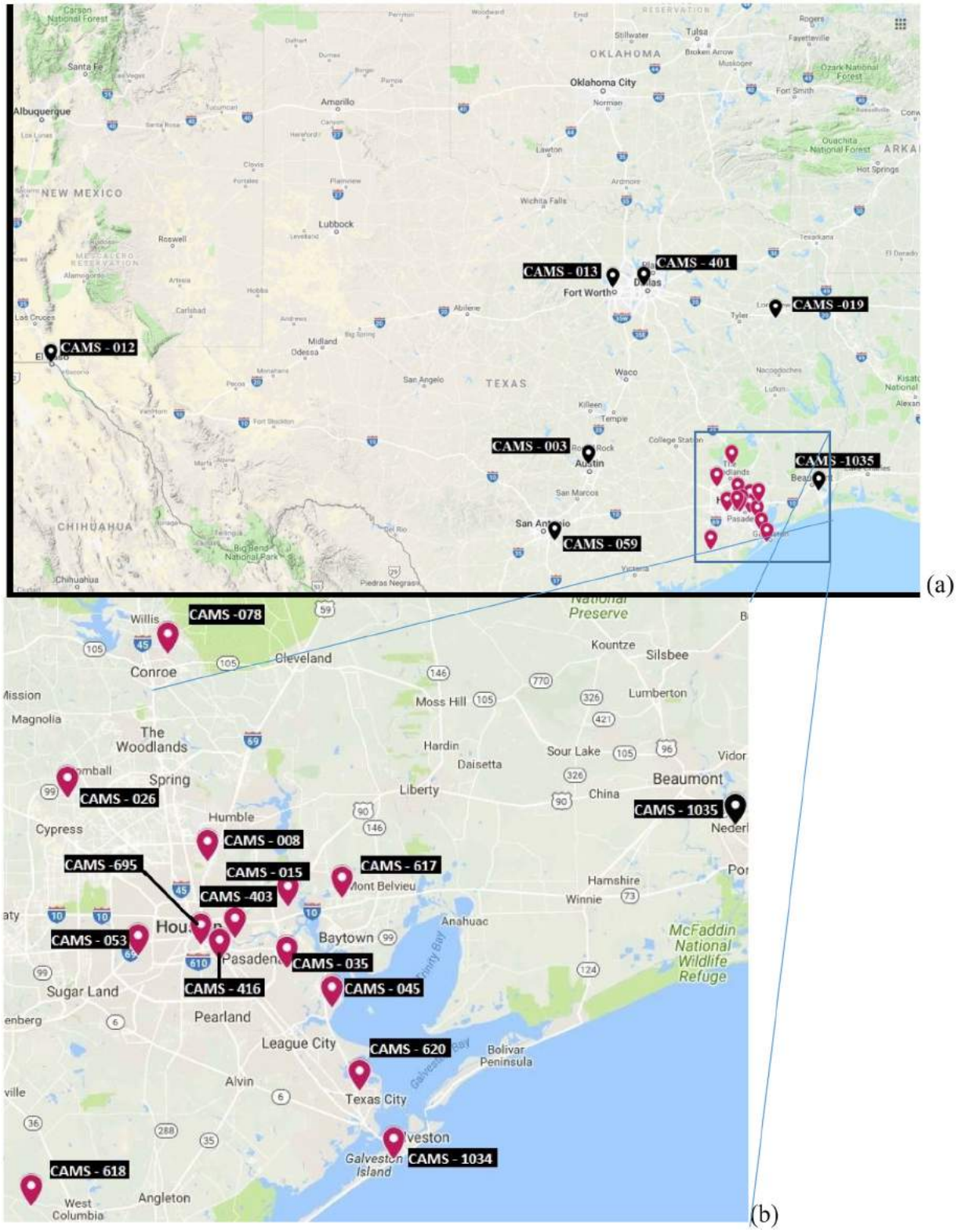

Fig. 1. Map of Texas (a) indicating locations of all stations used for the study. Map of Houston (b) shows the individual stations used in Houston for the study.

one section of a signal region to increase the robustness of the system by decreasing the number of features. The fully connected layer translates input features into output predictions.

Pooling can potentially remove important features that are sensitive to the output (e.g., a sudden change in peak due to the absence of favorable chemistry in the atmosphere) (Eslami, Choi, Lops and Sayeed, 2019; Eslami, Salman, Choi, Sayeed and Lops, 2019). Also, in contrast to an image or speech signal, the number of features (meteorology and air pollution) in this study is limited (e.g., Eslami, Choi et al., 2019; Eslami, Salman et al., 2019). The representation of each feature is essential; therefore, to avoid loss of information, we do not use the pooling layer (Hinton, 2014).

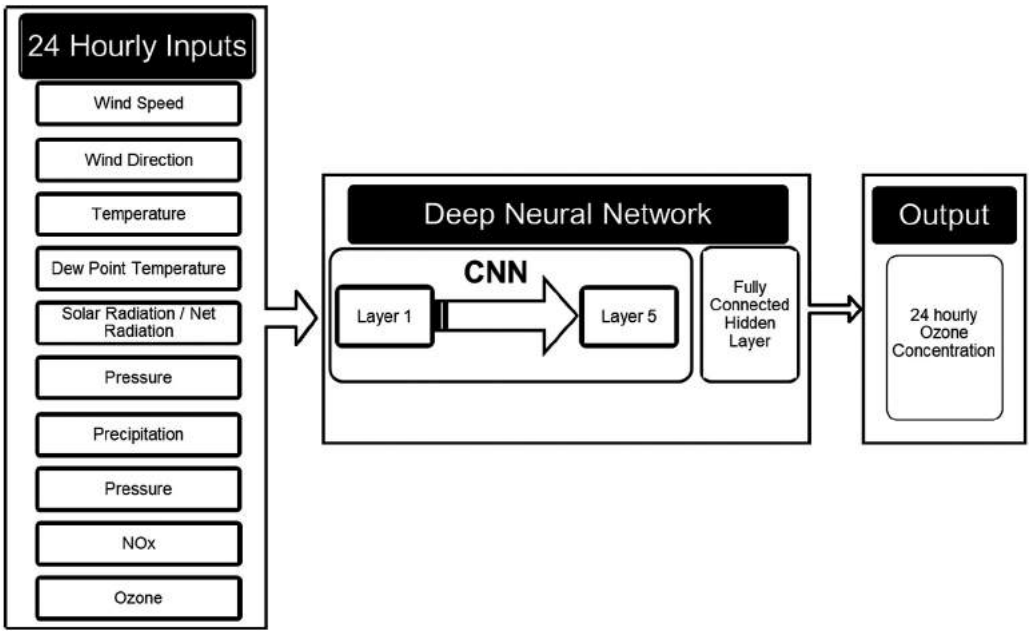

Fig. 2. Model Architecture: Detailed process flow of the deep CNN model.

This was confirmed during our test runs — where the addition of a pooling layer reduced the performance of the model.

Depth of any neural network depends on the type of problem it needs to solve. Ozone concentration is highly dependent on the availability of favorable meteorological conditions (e.g., sunlight) and the presence of certain chemical species (e.g., concentrations of volatile organic compounds and NO_x concentrations) in the atmosphere. Additionally, the transport of ozone from the surrounding region can lead to a change in its concentration. These factors add high non-linearity in time-series of ozone concentrations. Since the problem here is highly non-linear and dependent upon the external factors, we tested the architecture with more than one layer of CNN with various configurations (e.g., number of layers and kernel size). We then selected the best configuration with the least mean squared error on cross-validation set and maximum IOA on the test set (the year 2017). To build the deep CNN, we used five convolutional layers and one fully connected layer (Fig. 2). We applied convolution to the input features and the elements of the kernel (Fig. 3 and Figure B5 in the supplementary document). Since the kernel size (convolution window) here is 2×1, the convolution of two successive hours of input features takes place in the first layer. We then passed the results of the convolution operation to the activation function. The final features are the activation function (ReLU) applied to the output of the convolution (i.e., the three-dimensional tensor).

For any kind of neural network to achieve efficient optimization within a weight matrix while preserving nonlinearity, it needs to have an activation function (Nair & Hinton, 2010). For this model, we used a ReLU as the activation function, defined by Eq. (1) as follows:

$$f(x) = \max(0, x). \tag{1}$$

The final feature map obtained at the end of the first layer of the CNN acts as input for the second layer. Similarly, the output feature map of the second layer is the input for the third layer, and so on. In this way, the model has a five-layer CNN, each layer having 32 filters (activation by ReLU), each with two kernels randomly initialized with some value for the first iteration. After determining the feature maps in the last convolutional layer, a fully connected hidden layer with 264 nodes gives the final output of the model. We implemented the algorithm in the Keras environment with a TensorFlow backend (Abadi et al., 2016; Chollet, 2015).

2.3. Model training and prediction

Once we defined the model architecture, it required a training set consisting of input and output features from the previous day. For example: If we wish to predict the (n+1)th day, we train the model until the nth day with the input feature of the (n−1)th day and the output target of the nth day. Therefore, to predict the (n+1)th day, the model has n training examples (for details see Section C — Experimental set-up in the supplementary document). The model is then trained by the greedy layer-wise algorithm (Bengio, Lamblin, Popovivi, & Larochelle, 2007). Instead of optimizing the model in a single step, it is divided into a number of stages. The algorithm trains each model layer by layer. Initially, all the layers are frozen, and only the first layer is trained. Then the model assigns a weight to each input feature and computes the output, which is then compared to the actual observations, and the MSE is calculated. Depending on the MSE, the model changes the weight of the input feature and computes a new output. Again, this output is compared with actual observations, and the MSE is calculated. Once the MSE is minimized, the model is said to be trained. Once the first layer is trained, the second layer is added, and weights are preserved from the previous layer. The entire process is repeated once again on the two-layer network, which is then applied to the third layer, again preserving the weights from the two-layer model, and so on. This process optimizes the weights in a computationally efficient manner.

3. Results and discussion

We took the observed values of the 21 stations from the TCEQ CAMS stations for the year 2014–2017. We then trained the model for 2014–2016 (keeping 20% data for cross-validation for each deep learning and machine learning model). We applied the model to make predictions for the entire year, updating each day (i.e., adding the previous day in the training set). For example,

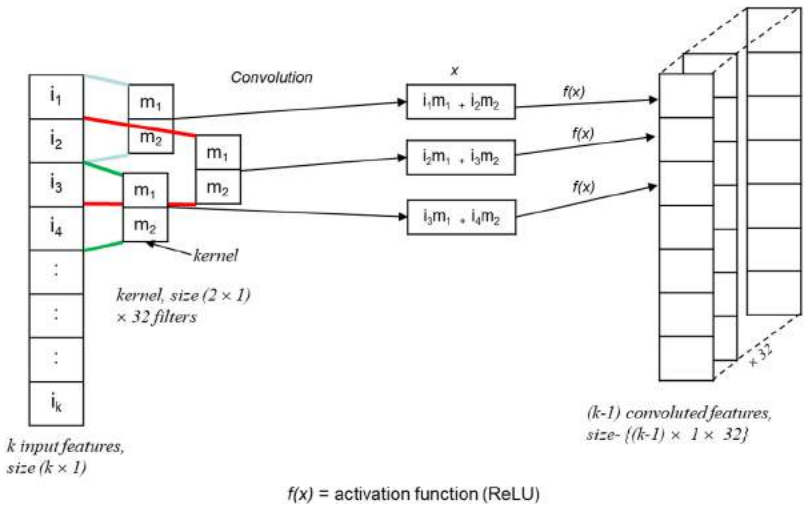

Fig. 3. Expanded view of the first CNN layer (the architecture has four more layers with a similar structure): Operations performed in a convolutional layer of CNN. (k is the total number of input features, whereas "i" is the input feature. m1 and m2 are the elements of each kernel that are randomly initialized).

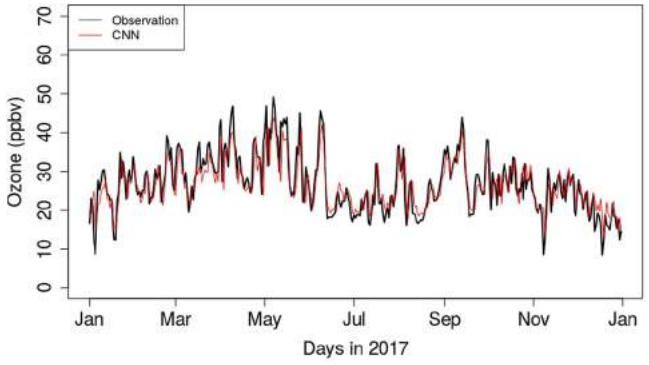

Fig. 4. Model-measurement comparisons for a daily mean of concentrations averaged over all stations mean for the year 2017. *Y*-axis represents daily mean ozone concentration, and the *x*-axis represents days of the year. Black solid lines are observed ozone concentration (in ppb), and red lines are predicted ozone concentration (in ppb). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

we initially trained the model until December 31, 2016, and predicted ozone concentrations for January 1, 2017. Then we added observations from January 1, 2017 to the training set and again trained the model and then predicted ozone concentrations for January 2, 2017 from the model. We repeated this process for 365 days. Once we had predicted ozone concentrations for all of 2017, we evaluated each station from the hourly and maximum daily eight-hour average (MDA-8) values.

3.1. Evaluation based on hourly values

Figs. 4 and 5 show the daily mean of the ozone concentrations over all stations. The results indicate the model slightly under-predicted ozone concentrations for most of the year. The maximum under- and over-prediction of the model were −15.5 ppbv and 12.9 ppbv, respectively. The monthly mean biases were in a range of −2.43 to 2.10 ppbv. The model underpredicted monthly mean biases in January, February, March, April, May, and September but overpredicted them in the remaining months. The IOA (Willmott, 1981) of the ozone concentration was 0.89, and the correlation was 0.81 of all stations.

The warmer months of the year (June, July, and August) showed minimum prediction bias and a median of ozone concentration close to zero (Fig. 5). Fig. 6 shows the monthly mean bias followed a positive trend to the observed ozone concentration, indicating that from January to May when the ozone concentration was increasing, the model bias (under-prediction) also increased. In June, July, and August when the observed ozone concentration showed a decreasing trend, the model bias increased (over-predicted). This trend of under- and over-prediction also occurred in subsequent months. In winter, the intensity of sunlight (due to clouds and solar zenith angle) reaching troposphere is comparatively lower, which leads to little variability in ozone formation. During the warm months, the meteorological conditions are stable, which leads to efficient formation of ozone (Eslami, Choi et al., 2019; Eslami, Salman et al., 2019). Due to this reason, the hourly ozone concentration had a more uniform daily trend in the summer months compared to the winter months. Since the trend is more consistent in summer, the model was more effective at predicting ozone concentrations during summer.

In general, the monthly mean ozone concentration decreased during the summer months (June, July, and August) with the lowest in July (after a yearly high in May) (Supplement Figure B1 in the supplementary document shows this monthly trend). This trend occurred in all stations except CAMS-012 (El Paso) and CAMS-013 (Fort Worth). At station CAMS-012, instead of decreasing after May, the ozone level peaked in June and then decreased until December. Since our model is station-specific, it was able to recognize general trends in ozone concentrations and predict them with an IOA of 0.89 and a correlation of 0.81. (see Tables A2 and A3 in the supplementary document for monthly IOA and correlation values).

Table 1 lists the annual averaged ozone mean bias and RMSE for 2017 at each station. The maximum mean bias, which was greater than 1 ppbv, occurred at stations CAMS-620 and CAMS-403. The bias stemmed from large variations in ozone concentrations because of the proximity of these stations to petrochemical plants, which increased their sensitivity to NO_x (Pan et al., 2017). As a result, ozone concentrations varied from day to night as well as from weekday to weekday. The other stations showed a bias of 0.5 and −1 ppbv. The RMSE of stations CAMS-012, 403, 008, 015, 301, 695, 053, 045, and 013 is between 9–10 ppbv.

A detailed analysis of each station suggests under-predictions for January to May and September at most of the stations (Fig. 7). As ozone concentrations steadily increased from January to May, the model tried to follow it, and thus lagged, leading to under-prediction during these months. In June, July, and August, ozone

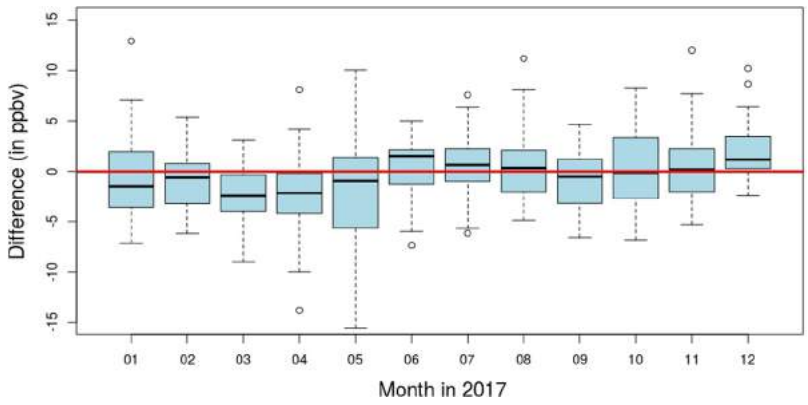

Fig. 5. Box and whisker plot for differences between the daily mean of predicted and observed ozone concentrations averages over all stations for 2017. The *x*-axis represents the month of the year; *y*-axis the difference between predicted and observed daily mean ozone concentrations in ppb. Upper and lower ends of the box indicate 25th and 75th percentiles respectively; the center of the box represents median and whiskers represent the maximum and minimum. Dot circle represents outliers or a single value.

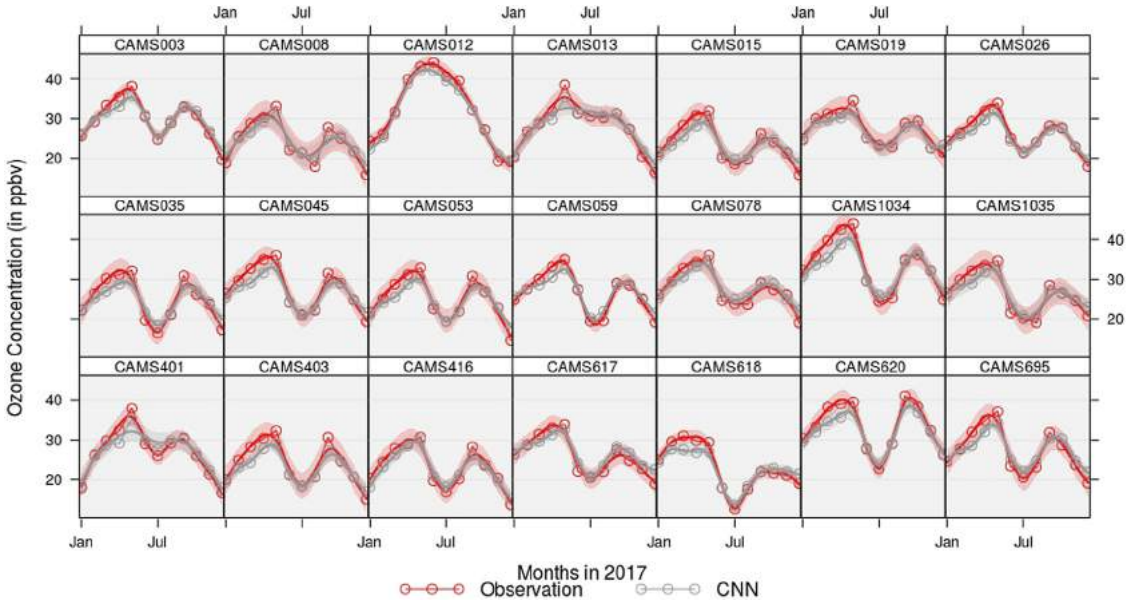

Fig. 6. The smooth trend of model and observation: Circles indicate monthly mean with the shaded region showing a 95% confidence interval.

concentrations steadily decreased, and the delay in model response to the decreasing trend led to over predictions. In September and August, observed ozone concentrations again increased, and the model again lagged, thus under-predicting. This behavior could be attributed to a lack of examples that could explain these variations for an efficient training process. In addition, the input variables for ozone prediction were inadequate; that is, the measurements for several important ozone predictors such as the cloud fraction and PBL height were not available for the study area, causing the unexplained bias in ozone prediction on some days, which weakened the overall performance of the model. Fig. 7 shows that the bias in the Houston area varied more than it did in other regions in Texas. In Houston, the formation of ozone, a secondary pollutant, is triggered by reactions between primary pollutants emitted by the oil and gas industry, automobiles, and biogenic sources at various locations in and around the city (Pan et al., 2017). In addition, the influence of the Gulf of Mexico on the meteorological condition of Houston results in large hourly and monthly variations of these input variables of the model. Figs. 8 and 9 plot variations in daily maxima of predictions and observations. The trends were similar to daily mean bias. The maximum range of the bias for daily maximum ozone was $\sim\pm$10 ppbv (barring a few outliers) while the IOA and correlation for a daily maximum of ozone concentration (averaged over all stations) were 0.87 and 0.81 respectively.

Model accuracy, represented by the IOA, is shown in Fig. 10 (refer to Table A2 in the supplementary document for values). The IOAs of 19 stations were in the range of 0.85–0.90. Station CAMS-078 had the highest IOA of 0.90. Only two stations (CAMS-045 and CAMS-620) had an IOA below 0.85. Both stations are in the Galveston Bay area and positioned downwind of petrochemical refineries in Houston. One explanation for the low IOA at these stations can be attributed to frequent changes in hourly observed ozone concentrations resulting from the combined effect of the land–sea breeze and the presence of petrochemical refineries (NO_x sensitive region) (Pan et al., 2017). The month-wise analysis suggests summer months (June, July, August, and September) had maximum IOAs, and winter (DJF) months had the lowest at almost every station. This trend is similar to the mean monthly trend, suggesting less overprediction and more underprediction of ozone concentrations.

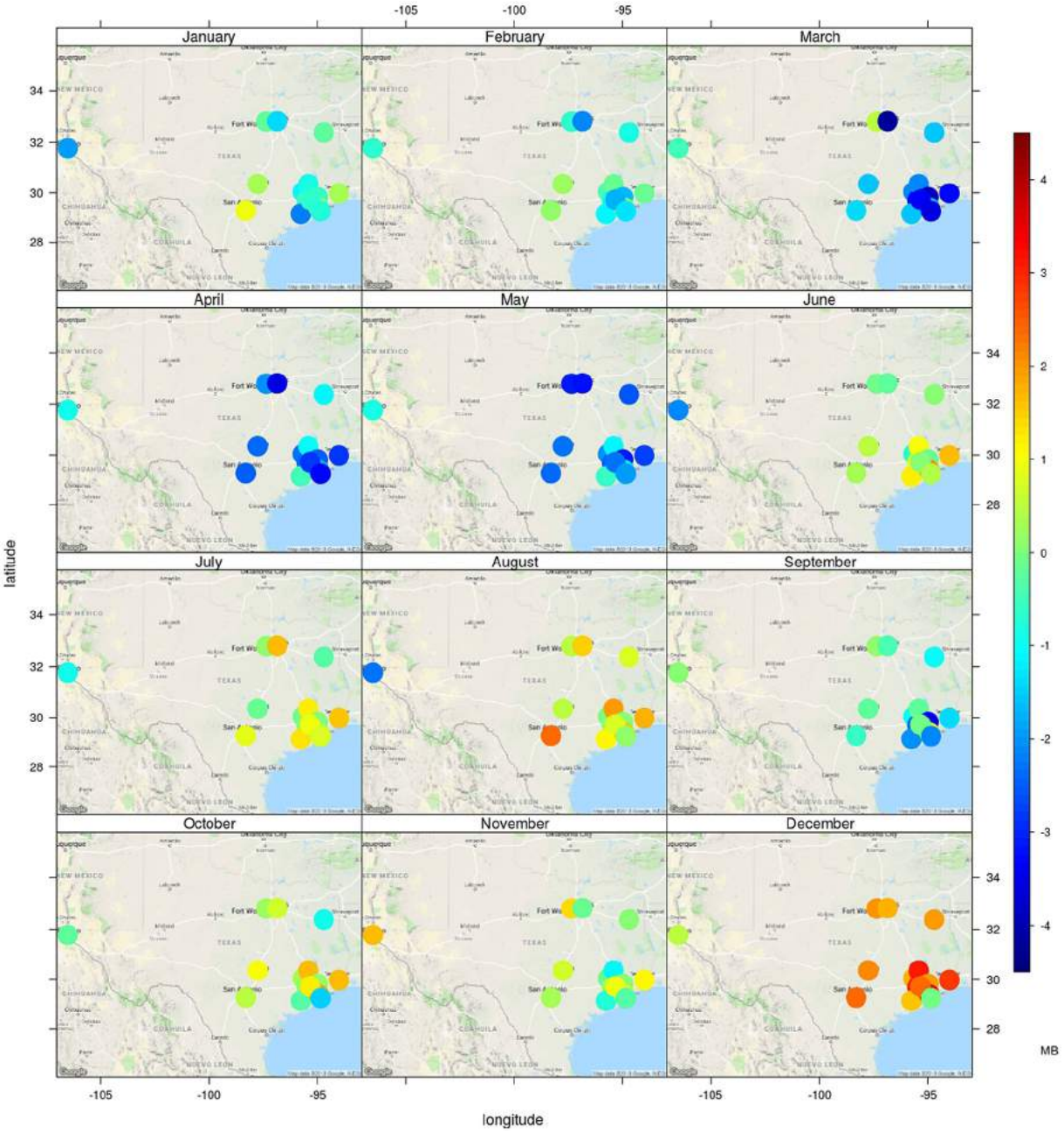

Fig. 7. Mean Bias (Monthly MB of each Station): Each dot shows location-wise MB for each month of the Year 2017.

3.2. Based on MDA-8 (Categorical statistics)

Categorical statistics, which are based on a threshold, determine the likelihood of occurrence of an event (Zhang et al., 2012). They are more applicable to the evaluation of a forecasting model. To evaluate our model, we use the following categorical statistics: HIT (the Hit Rate), CSI (the Critical Success Index), FAR (False Alarm Rate), ETS (Equitable Threat Score), and POC (the Proportion of Correct) (as in Chai et al., 2013; Eder, Kang, Mathur, Yu, & Schere, 2006). Based on the threshold value, four different cases arise:

(a) N_a = Number of times a prediction is above, and an observation is below the threshold.
(b) N_b = Number of times a prediction and an observation are above the threshold.
(c) N_c = Number of times a prediction and an observation are below the threshold.
(d) N_d = Number of times a prediction is below, and an observation is above the threshold.

From these cases, we defined the following quantities to evaluate the model:

HIT represents the fraction of instances in which we predict an extreme event (events above the threshold) from all actual occurrences of extreme events.

$$HIT = \frac{N_b}{N_b + N_d} \tag{2}$$

CSI is the fraction of instances in which we correctly predict an extreme event out of all events after removing correctly predicted

Table 1
Discrete statistics of AI model: MB-Mean Bias, NMB-Normalized Mean Bias, RMSE-Root Mean Square Error and NME-Normalized Mean Bias of the average of ozone concentration of all 21 stations.

Station name	Region	Discreet			
		MB (ppbv)	NMB (%)	RMSE (ppbv)	NME (%)
CAMS-003	Austin-Round Rock	−0.146	−0.49	7.831	20.02
CAMS-008	Houston-Galveston-Brazoria	−0.028	−0.12	9.428	29.17
CAMS-012	El Paso-Juarez	−0.769	−2.38	9.647	23.19
CAMS-013	Dallas-Fort Worth	−0.131	−0.47	9.082	25.05
CAMS-015	Houston-Galveston-Brazoria	−0.355	−1.51	9.411	30.22
CAMS-019	Tyler-Longview-Marshall	−0.554	−2.04	8.468	24.22
CAMS-026	Houston-Galveston-Brazoria	−0.719	−2.76	8.605	24.62
CAMS-035	Houston-Galveston-Brazoria	−0.358	−1.45	8.531	26.10
CAMS-045	Houston-Galveston-Brazoria	−0.945	−3.42	9.152	24.85
CAMS-053	Houston-Galveston-Brazoria	−0.651	−2.62	9.303	28.02
CAMS-059	San Antonio	0.052	0.20	7.996	22.41
CAMS-078	Houston-Galveston-Brazoria	0.159	0.57	8.504	23.48
CAMS-401	Dallas-Fort Worth	0.002	0.01	9.372	26.68
CAMS-403	Houston-Galveston-Brazoria	−1.143	−4.78	9.524	29.64
CAMS-416	Houston-Galveston-Brazoria	−0.385	−1.68	8.684	27.99
CAMS-617	Houston-Galveston-Brazoria	0.523	2.01	8.501	24.57
CAMS-618	Houston-Galveston-Brazoria	−0.430	−1.87	8.111	26.23
CAMS-620	Houston-Galveston-Brazoria	−1.188	−3.58	8.822	19.98
CAMS-695	Houston-Galveston-Brazoria	0.005	0.02	9.350	25.62
CAMS-1034	Houston-Galveston-Brazoria	−0.837	−2.50	8.751	19.22
CAMS-1035	Beaumont-Port Arthur	−0.213	−0.80	7.877	22.38
Average of ALL	–	−0.39	−1.14	8.81	27.94

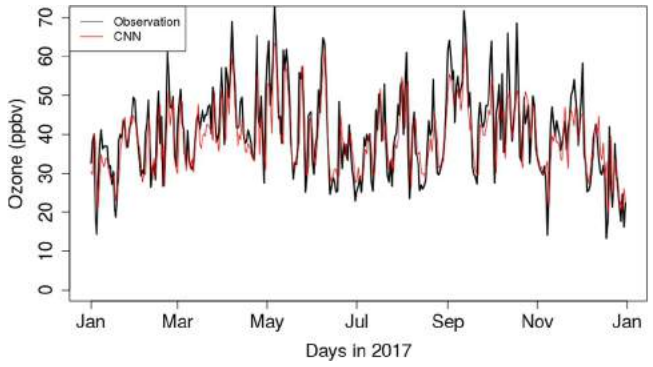

Fig. 8. Model-measurement comparisons for daily maximum ozone concentrations for 2017. Black lines represent observations and red model predictions . (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

below the threshold instances.

$$CSI = \frac{N_b}{N_a + N_b + N_d} \tag{3}$$

FAR is the fraction of instances in which we wrongly predict an extreme event of all predictions of an extreme event.

$$FAR = \frac{N_a}{N_a + N_b} \tag{4}$$

ETS is similar to CSI but is more accurate for measuring performance skill of a model. It ranges between −1 (poor model) to 1 (perfect model) Schaefer, 1990. A positive value of ETS means a skillful model.

$$ETS = \frac{N_b - N_r}{N_a + N_b + N_d - N_r} \tag{5}$$

$$\text{where } N_r = \frac{(N_a + N_b) \times (N_b + N_d)}{N_a + N_b + N_c + N_d} \tag{6}$$

POC is the fraction of instances in which we accurately predict an exceedance or a non-exceedance.

$$POC = \frac{N_b + N_c}{N_a + N_b + N_c + N_d} \tag{7}$$

To set the threshold, we selected two parameters, (i) effects on human health; (ii) the fraction of observations above the threshold that make sense of the categorical statistics. Considering the severe health repercussions of ozone, the Environmental Protection Agency (EPA), on October 1, 2015, strengthened

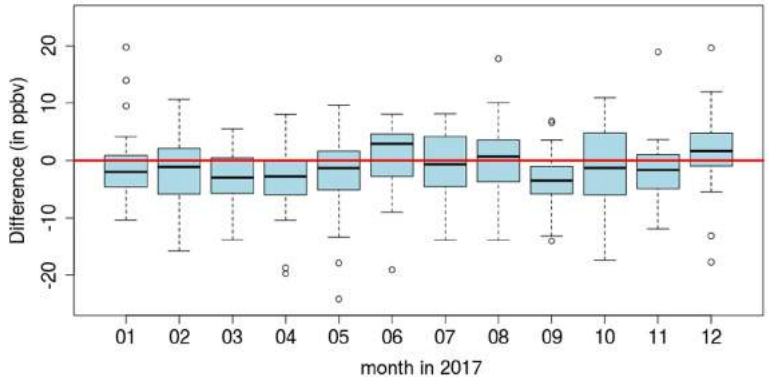

Fig. 9. Box and whisker difference plot for a daily maximum of model predicted and measured ozone concentrations.

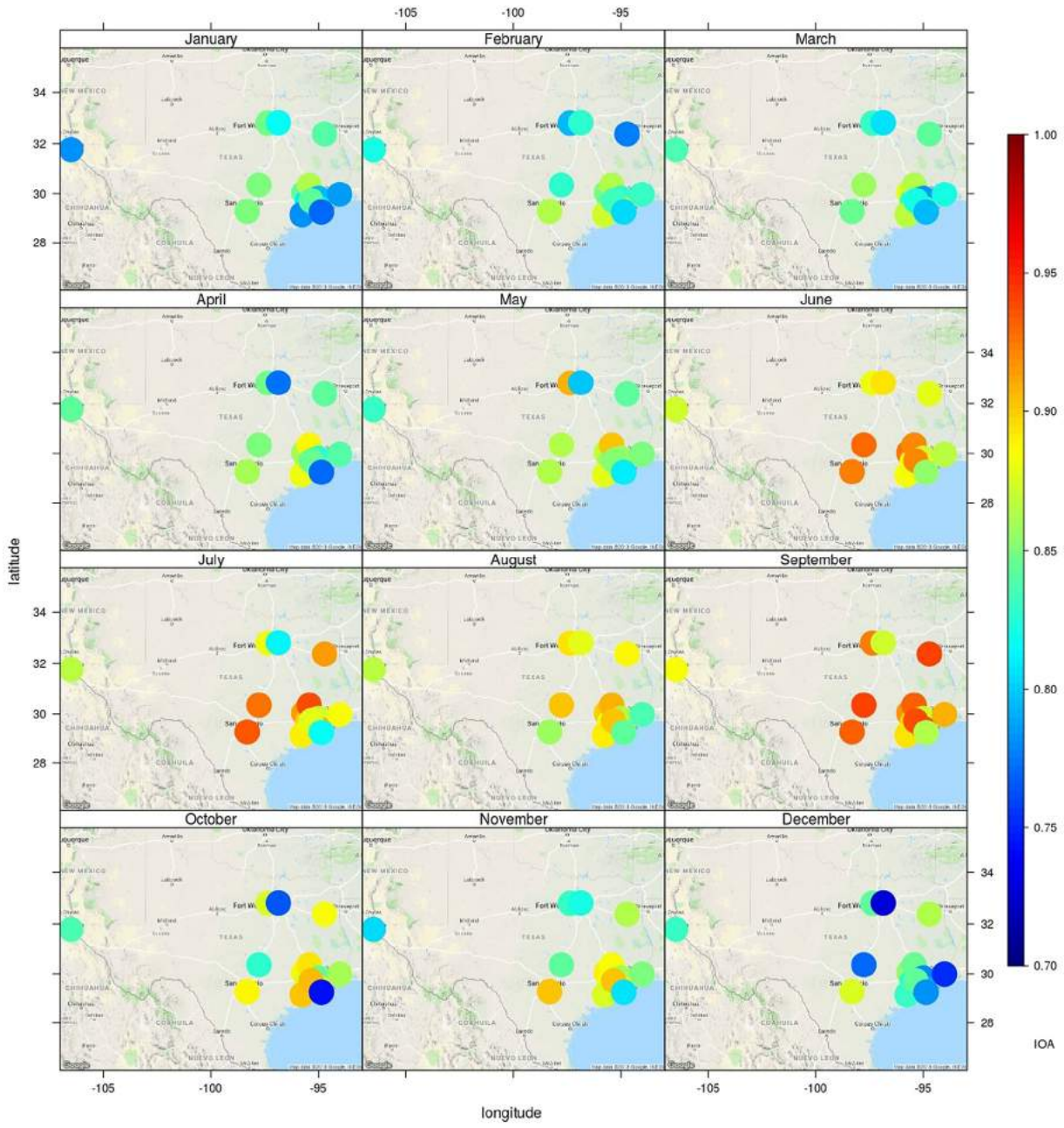

Fig. 10. Monthly Index of Agreement for all stations in Texas for the year 2017. Each circle represents the location and corresponding value.

the ground-level maximum 8-h averaged ozone standard from 75 ppbv to 70 ppbv (National Ambient Air Quality Standards for Ozone-2015). US-EPA guideline for ambient air quality states that an ozone concentration of 70 ppbv is unhealthy for individuals who are sensitive to ozone (https://cfpub.epa.gov/airnow/index.cfm?action=pubs.aqiguideozone). For all 21 stations in this study, the number of occurrences above the 70 ppbv threshold was 95/7665 (i.e., only 1.23%). Thus, using the 70 ppbv threshold will not produce meaningful value for the parameters of categorical statistics. The next level, 55 ppbv, is unhealthy for individuals who are exceptionally sensitive to ozone. The number of occurrences above this threshold was 734 (i.e., 9.57%). Therefore, we set the threshold limit at 55 ppbv for the evaluation of the categorical statistics.

Table 2 lists the categorical statistics obtained from a comparison between MDA-8 observations and predictions (Table A4 in the supplementary document represents categorical statistics from Chai et al. (2013) for comparison). The model has an overall ETS of 0.26, a HIT of 0.34, and a FAR of only 0.32. The accuracy, or POC, of the model is also adequate, with a score of 0.92. Except for CAMS-019, all stations have an ETS higher than 0.10. Since the model mostly under predicts, the HIT ranges from 0.04 to 0.53, with one station (CAMS-019) at 0.04 and the other stations above 0.15. More than half of the stations exhibit a HIT higher than 0.30. The FAR is between 0.13 and 0.71 with only four stations reporting a FAR greater than 0.50 and ten stations reporting a FAR as low as 0.20 or below.

3.3. Case study

3.3.1. Stations with fewer meteorological input features

To evaluate the effectiveness of the model, we selected two stations (CAMS-003 and CAMS-059) that monitor three meteorological input features and six stations (CAMS-045, 053, 078, 617,

Table 2
Daily MDA8 ozone categorical statistics for 2017 with 55 ppbv threshold. See text for details.

	N_a	N_b	N_c	N_d	N_r	HIT	CSI	FAR	POC	ETS
CAMS003	6	18	325	16	2.24	0.53	0.45	0.25	0.94	0.42
CAMS008	10	21	314	20	3.48	0.51	0.41	0.32	0.92	0.37
CAMS012	19	33	268	45	11.11	0.42	0.34	0.37	0.82	0.25
CAMS013	10	19	305	31	3.97	0.38	0.32	0.34	0.89	0.27
CAMS015	2	7	342	14	0.52	0.33	0.30	0.22	0.96	0.29
CAMS019	2	1	338	24	0.21	0.04	0.04	0.67	0.93	0.03
CAMS026	4	11	321	29	1.64	0.28	0.25	0.27	0.91	0.22
CAMS035	1	7	336	21	0.61	0.25	0.24	0.13	0.94	0.22
CAMS045	4	4	334	23	0.59	0.15	0.13	0.50	0.93	0.11
CAMS053	3	14	320	28	1.96	0.33	0.31	0.18	0.92	0.28
CAMS059	4	5	331	25	0.74	0.17	0.15	0.44	0.92	0.13
CAMS078	9	16	316	24	2.74	0.40	0.33	0.36	0.91	0.29
CAMS401	6	10	322	27	1.62	0.27	0.23	0.38	0.91	0.20
CAMS403	2	9	335	19	0.84	0.32	0.30	0.18	0.94	0.28
CAMS416	10	4	332	19	0.88	0.17	0.12	0.71	0.92	0.10
CAMS617	5	8	331	21	1.03	0.28	0.24	0.38	0.93	0.21
CAMS618	4	3	345	13	0.31	0.19	0.15	0.57	0.95	0.14
CAMS620	4	16	329	16	1.75	0.50	0.44	0.20	0.95	0.42
CAMS695	7	19	319	20	2.78	0.49	0.41	0.27	0.93	0.38
CAMS1034	4	19	309	33	3.28	0.37	0.34	0.17	0.90	0.30
CAMS1035	2	5	341	17	0.42	0.23	0.21	0.29	0.95	0.19
Overall	**118**	**249**	**6813**	**485**	**35.14**	**0.34**	**0.29**	**0.32**	**0.92**	**0.26**

618, and 620) that monitor four meteorological input features. Although these stations monitor few meteorological variables (i.e., wind speed and direction, temperature, and solar radiation), the performance of these stations is on par with other stations. The IOAs for CAMS-003 and 059 were greater than 0.89. CAMS-003 had the highest ETS of 0.42 while CAMS-059 had an ETS of 0.13. While stations CAMS-053, 078, 617, and 618 performed well with IOAs greater than 0.88 and correlations greater than 0.79; stations CAMS-045 and 620 had IOAs of 0.84 (an explanation of their weak performance and solution for these stations are discussed in the next section). An analysis of the importance of input features to predicted ozone concentrations suggested that the results stemmed from the dependence of the model on previous day ozone concentrations. The model assigns higher weights to the previous day ozone than to the previous day meteorology, indicating that the model considers ozone concentration (previous day) the most critical input variable. Thus, even with only three or four meteorological input variables, the model is relatively accurate at forecasting ozone concentrations.

3.3.2. Station CAMS-045

In this study, we found that the stations with the weakest performance were CAMS-045, CAMS-620, and CAMS-1034. These stations are uniquely located near large bodies of water, which can initiate a cooling effect that results in the lowering of the height of the PBL (planetary boundary layer). As a result, observed ozone concentrations varied considerably hour by hour (Chai et al., 2013). During the training phase of the model, these variations created noise that may have led to lower performance metrics. To mitigate the problem of noise, we fed more input data into the model for training purposes. We trained station CMAS-045 with ten years (2007–2016) of data instead of three years (2014–2016) and observed improvement in the results: The IOA increased by 2.5% and the correlation by 3.6%. Table 3 shows a month-by-month comparison of the effect of additional training examples.

Table 3
Comparison of IOA and correlation for station CAMS045 based on 10 and 3-year training.

Months	IOA		Correlation	
	10 years	3 years	10 years	3 years
Jan	77.74	74.87	65.46	61.09
Feb	82.90	80.57	73.36	68.84
Mar	83.26	79.28	74.84	72.18
Apr	80.72	77.16	68.13	64.10
May	85.41	83.71	76.35	73.17
Jun	86.77	86.98	77.35	78.16
Jul	85.62	84.30	75.52	73.84
Aug	86.98	84.41	77.60	74.01
Sep	87.02	84.59	78.22	76.51
Oct	83.72	82.44	73.22	71.29
Nov	80.80	79.87	69.20	67.61
Dec	79.14	75.82	65.71	59.73
Over all	86.49	84.38	77.43	74.76

3.4. Model comparison

To test the robustness of the model used for this study (i.e., CNN model), we compared various machine learning models that are commonly used in predicting time-series, particularly those with high nonlinearity (e.g., Eslami, Choi et al., 2019; Eslami, Salman et al., 2019). They include multilayer perceptron (MLP) structure (Glorot & Yoshua, 2010; Hinton, 1989), deep neural networks (DNN) formed by having densely-connected layers (Figure B6 in supplementary document), recurrent neural networks (RNN) with Gated Recurrent Unit (GRU) (Cho et al., 2014), Lasso Regression (Tibshirani, 1996), and Ridge Regression (Tikhonov, Leonov, & Yagola, 1998). These models, including CNN, were trained on the same dataset as discussed above. Table 4 shows the model configuration and performance comparisons based on the computational time required to train each model. Once the model was trained, it was used to predict the whole year of 2017. In addition to these models we also tested linear auto-regression models with and without exogenous variable; and time-delayed neural network (TDNN) details about which can be found in Section E and F of the supplementary document.

For comparing the performance of different models, we compared the IOA of 24-h time series (Fig. 11) and IOA of the daily maximum of ozone concentration (Fig. 12) for all stations. The CNN model performed better than all other models while DNN was the close second. CNN model had the highest mean IOA (0.89) for the 24-hourly time series. When comparing the daily maximum of ozone concentration (Figs. 12 and 13), we found that the CNN model (IOA = 0.78) performed better than the other models. Furthermore, Figs. 12 and 13 show all deep learning models (CNN, DNN, and RNN) performed notably better than other machine learning model (MLP) and regression model (Lasso

Table 4
Specifications of neural network models compared in this study.

Model	Hidden/convolutional layer(s) structure[a]	Number of epochs[a]	Optimizer[b]	Computational time (in s)
CNN	5 layer of CONV1D/264[c]	100	Adam	16.67
RNN (GRU)	128/64/32	400	Adam	1179.23
DNN	128/64/32	100	Adam	4.90
MLP repressor	–	100	SGD	1.8–4.3
Ridge regression	–	–		<1
Lasso regression	–	–		<1

[a]Optimized using trial and error tests.
[b]Both Adam and stochastic gradient descent (SGD) were explained in Kingma and Ba (2014).
[c]Convolutional layers with filter size 32 and kernel size 2, following with a fully connected hidden layer with size 264.

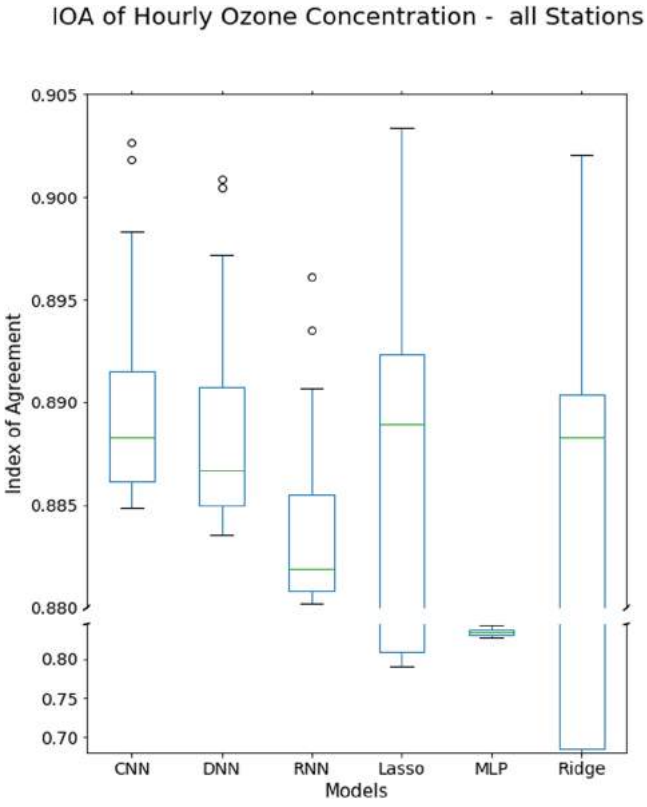

Fig. 11. Box and whisker plot of the index of agreement (IOA) based on hourly ozone concentration for the year 2017 of all the stations. *Y*-axis represents IOA, and *X*-axis represents the models in the study.

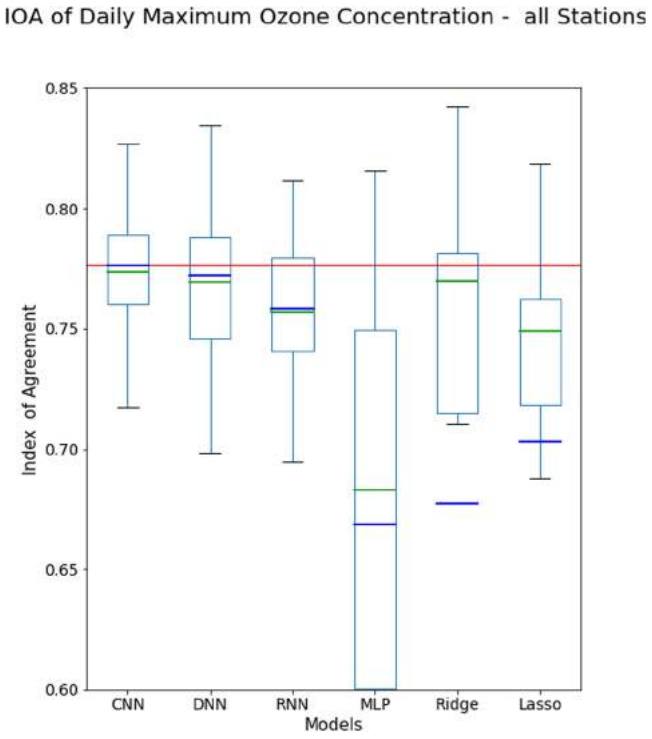

Fig. 12. Box and whisker plot for the index of agreement based on daily maximum ozone concentration of all stations. *Y*-axis represents IOA, and *X*-axis represents the models in the study.

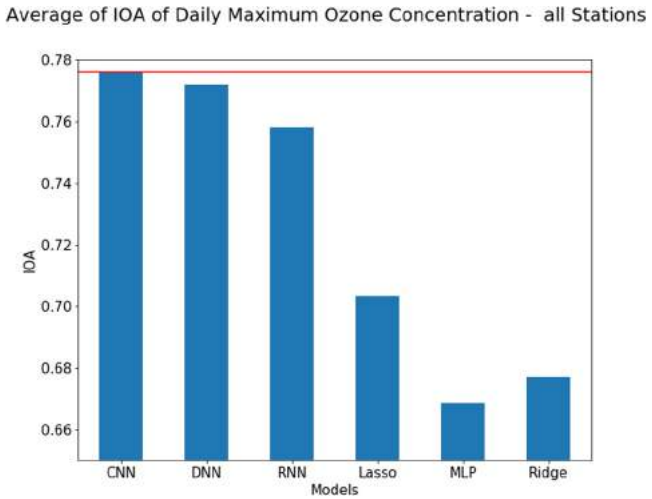

Fig. 13. Bar plot average of IOA for all stations. IOA is based on hourly ozone concentration for the year 2017. *Y*-axis represents IOA, and *X*-axis represents the models in the study.

and Ridge). It means that these models understood the topology within the daily ozone time series (the relationship between different hours during a day), even though they were trained with the same amount of training samples. This can also be seen in the categorical analysis. CNN model has the highest mean POC (0.92), CSI (0.273) and ETS (0.244) among all the models (Figure B4 in the supplementary document). CNN was also better than other models in predicting an extreme event with the highest mean HIT rate (0.313). To summarize the abovementioned comparisons, our CNN model performed statistically better than the other machine learning and regression models.

Even though Ridge and Lasso regression shows a better computational efficiency than the other models, they were unable to predict the outliers (high ozone peaks) because of the L2 (Ridge regression) and L1 (Lasso regression) regularizations. The regularization process has superb computational efficiency; however, it negatively affected the accuracy performance of real-time prediction compared with deep learning models. Also, these regression techniques have high variance in IOA for hourly time-series as

well as in the IOA of daily maximum concentration across all stations. Although the MLP model was about three times faster than CNN model used for this study, the accuracy (in terms of IOA) was 2%–8% less than CNN model for all stations. Both DNN and RNN models show comparable accuracy but have performance lower than that of CNN. Additionally, RNN was ∼70 times slower in training as compared to CNN for this study. It means that, even though the RNN model can reach a comparable level of prediction accuracy as other two deep learning models, it needs notably higher computational time. From the above discussion, we can conclude that the CNN model was better than the other models tested for this study across various evaluation parameters for real-time prediction of hourly ozone concentrations.

One commonly-used practice in improving machine learning models, especially in predictive regression modeling problems, is feature importance analysis. It is the automatic selection of features in the data that are most relevant to the predictive modeling problem. In feature importance analysis, the method selects the most critical features presented in the data without changing their values, and uses these selections in the training process of the predictive model (e.g., a CNN model). In addition, using irrelevant features can negatively impact the performance of our machine learning model in predicting hourly ozone time-series by making the model learn based on unimportant features. In this way, we can test the robustness of our CNN model by presenting inputs with or without the knowledge from feature importance analysis. Here, we performed the feature selection using a random forest (RF) model (Breiman, 2001). We choose RF due to its generally good predictive performance, low overfitting, and easy interpretability compared with other machine learning models. First, we trained an RF model to determine the importance of input features in predicting hourly ozone concentrations. Once a list of the features with their importance was obtained, CNN models were trained with best 24-set (24, 48, 72, etc.) of features. We then compared each model against the standard CNN model used in this study (Table A5 in the supplementary document). Results indicated that the averaged improvement in IOA was around 0.5% higher after applying the RF feature selection compared to the standard CNN model. This shows that our CNN model was able to extract enough information needed to make a proper prediction and understanding the importance of each input feature without applying a prior feature selection method. One of the major caveats in using RF for feature selection is that every station has a different number of features on which their performance improved and it would become cumbersome for the user to run multiple models for an optimum result. Thus, to have a generalized model for all stations, we used the standard CNN.

4. Conclusions

In this paper, we developed, discussed, and evaluated the real-time 24-h ozone prediction model based on a deep convolutional neural network. After designing the model architecture, we trained it on examples from 2014 to 2016. We predicted the concentrations of ozone for each day of 2017 by training the model with examples until the last day and evaluated the prediction data using discrete and categorical statistics.

In general, the observed ozone concentrations at all stations increased from winter to spring and then decreased during the summer months. During the fall, concentrations again steadily rose until September and then declined to their lowest levels in December. Although the model was able to capture these patterns, the response of the model to these changes was slow. For example, when the concentration of observed ozone increased, the model generally underpredicted (for both daily and monthly variations), and when observed ozone concentrations declined, it generally overpredicted. The change in the trends of prediction was largely the result of the delay of the model in responding to changes in the concentrations of observed ozone.

Because of the location of CAMS-012 (El-Paso) 1140 m above sea-level, it exhibited a meteorological condition that differed markedly from those of other stations in this study. As a result, this station did not follow the general trend mentioned above. Here, the monthly mean ozone concentration steadily increased until May–June and then decreased until December before increasing again. The probable cause for this increase could be the presence of comparatively high temperatures and sunlight during the daytime, which can lead to the high production of ozone. Another explanation could be the accumulation of ozone from the reversal of the wind direction from westerly to easterly (Figure B2 in the supplementary document) during the summer and the presence of the Rocky Mountains in the West. Even though this station did not exhibit the general trend, our model was able to capture the trend in the observed ozone concentration with an IOA of 0.89 and a correlation of 0.81. The performance of the model at this station suggests that it is able to comprehend the station-specific trends and produce satisfactory results.

At a few stations (e.g., CAMS-045), the model did not perform satisfactorily because of more frequent hourly variations in observed ozone concentrations, either due to emissions of NO_x or the lowering of the PBL triggered by the cooling effect from the sea breeze. However, with the addition of more years of training data, the performance metrics of these stations improved. We trained the model for seven more years for station CAMS-045, and its IOA improved from 0.84 to 0.86, and its correlation improved from 0.74 to 0.77. In addition, even though the model is capable of predicting daily trends with reasonable accuracy, it mostly underpredicts daily maxima (also evident in the scatter plot Figure B3 Supplementary Document). This drawback can be overcome with the addition of more days of training data as well as more inputs (i.e., more meteorological inputs like PBL, cloud fraction or air pollutants), which can be obtained from a numerical model such as WRF or CMAQ.

The deep CNN model that we developed to forecast ozone concentrations 24 h in advance showed significant improvement over both numerical and statistical models. Furthermore, when compared with other neural networks (MLP, RNN, and DNN) and regression models (Lasso and Ridge), our CNN model was better in predicting both daily time-series and daily maximum of ozone concentrations. Additionally, it was able to generate the same level of accuracy compared with a model with feature importance analysis.

In addition, the model successfully predicted ozone concentrations with IOAs greater than 0.85 for 19 of the 21 stations. We could considerably improve the performance metrics of the stations by adding more years of training examples, as demonstrated in the second case study. The model also proved its effectiveness on stations with fewer meteorological inputs (i.e., CAMS-003, CAMS-059). In addition, it took only a few minutes to predict a single day so that it could be employed at any station. The benefit of this model is that it could be a useful, efficient tool for forecasting air quality and issuing health advisories in advance, thus reducing the serious effects of ozone on human health.

Acknowledgment

This study was supported by the High Priority Area Research Seed Grant of the University of Houston.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.neunet.2019.09.033.

References

Abadi, M., et al. (2016). Tensorflow: a system for large-scale machine learning. In *OSDI*, 16, 265-283.

Ayres, J., Maynard, R., & Richards, R. (2006). *Air pollution and health*. London: Imperial College Press.

Baklanov, A., Korsholm, U., Mahura, A., Petersen, C., & Gross, A. (2008). ENVIRO-HIRLAM: on-line coupled modeling of urban meteorology and air pollution. *Advances in Science and Research*, *2*, 41–46.

Bell, M. L., Peng, R. D., & Dominici, F. (2006). The exposure-response curve for ozone and risk of mortality and the adequacy of current ozone regulations. *Environmental Health Perspectives*, *114*, 532–536.

Bengio, Y. (2009). Learning deep architectures for AI. *Foundations and Trends in Machine Learning*, *2*, 1–127.

Bengio, Y., Lamblin, P., Popovivi, D., & Larochelle, H. (2007). Greedy layer-wise training of deep networks. *Advances in Neural Information Processing Systems*, *19*.

Biancofiore, F., et al. (2015). Analysis of surface ozone using a recurrent neural network. *Science of the Total Environment*, *514*, 379–387.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32.

Byun, D., & Schere, K. L. (2006). Review of the governing equations, computational algorithms, and other components of the models-3 community multiscale air quality (CMAQ) modeling system. *Applied Mechanics Reviews*, *59*, 51–77.

Byun, D., et al. (2007). Evaluation of air quality models for the simulation of a high ozone episode in the houston metropolitan area. *Atmospheric Environment*, *41*(2007), 837–853.

Chai, T., et al. (2013). Evaluation of the United States national air quality forecast capability experimental real-time predictions in 2010 using air quality system ozone and NO2 measurements. *Geoscientific Model Development*, *6*, 1831–1850.

Cho, et al. (2014). On the Properties of Neural Machine Translation: Encoder-Decoder Approaches, arXiv:1409.1259.

Chollet, F. (2015). keras. https://keras.io, accessed on 10.07.18.

Czader, B. H., et al. (2015). Development and evaluation of the screening trajectory ozone prediction system (STOPS, version 1.0). *Geoscientific Model Development*, *8*, 1383–1394.

Damo, R., & Icka, P. (2012). Assessment of air pollution using the global pollution index for city of Korça, Albania. *Journal of International Environmental Application and Science*, *7*, 858–863.

Eder, B., Kang, D., Mathur, R., Yu, S., & Schere, K. (2006). An operational evaluation of the Eta-CMAQ air quality forecast model. *Atmospheric Environment*, *40*, 4894–4905.

Eslami, E., Choi, Y., Lops, Y., & Sayeed, A. (2019). A real-time hourly ozone prediction system using deep convolutional neural network. *Neural Computing and Applications (2019)*, http://dx.doi.org/10.1007/s00521-019-04282-x.

Eslami, E., Salman, A. K., Choi, Y., Sayeed, A., & Lops, Y. (2019). A data ensemble approach for real-time air quality forecasting using extremely randomized trees and deep neural networks. *Neural Computing and Applications*, http://dx.doi.org/10.1007/s00521-019-04287-6.

Fernando, H. J. S., Mammarella, M. C., Grandoni, C., Fedele, P., Di Marco, R., Dimitrova, R., et al. (2012). Forecasting PM10 in metropolitan areas: efficacy of neural networks. *Environmental Pollution*, *163*, 62–67.

Glorot, X., & Yoshua, B. (2010). Understanding the difficulty of training deep feedforward neural networks, In *International conference on artificial intelligence and statistics*.

Hinton, G. E. (1989). Connectionist learning procedures. *Artificial Intelligence*, *40*(1), 185–234.

Hinton, G. E. (2014). What's wrong with convolutional nets? MIT Brain and Cognitive Sciences – Fall Colloquium Series, URL http://techtv.mit.edu/collections/bcs/videos/30698-what-s-wrong-with-convolutional-nets.

Hoshyaripour, G., et al. (2016). Prediction of ground-level ozone concentration in Sao Paulo, Brazil: Deterministic versus statistic models. *Atmospheric Enviroment*, *145*, 365–375.

Jacobson, M. Z. (2005). *Fundamentals of atmospheric modelling, Vol. 2* (pp. 357–419). Cambridge University Press.

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization, arXiv preprint arXiv:1412.6980.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, *60*(6), 84–90.

Lawrence, S., Gills, L., Tsoi, A. H., & Back, A. D. (1997). Face recognition: A convolutional neural-network approach. *IEEE Transactions on Neural Networks*, *8*(1).

LeCun, Y., & Bengio, Y. (1995). Convolutional networks for images, speech, and time series, The handbook of brain theory and neural networks, ISBN:0-262-51102-9, (pp. 255-258).

Marr, D. (1977). Artificial intelligence – A personal view. *Artificial Intelligence*, *9*, 37–48.

Mills, G., Buse, A., Gimeno, B., Bermejo, V., Holland, M., Emberson, L., et al. (2007). A synthesis of AOT40-based response functions and critical levels of ozone for agricultural and horticultural crops. *Atmospheric Environment*, *41*, 2630–2643.

Misenis, C., & Zhang, Y. (2010). An examination of sensitivity of WRF/Chem predictions to physical parameterizations, horizontal grid spacing, and nesting options. *Atmospheric Research*, *97*, 315–334.

Nair, V., & Hinton, G. E. (2010). Rectified Linear Units Improve Restricted Boltzmann Machines, In *International conference on machine learning (ICML-10)*.

Pan, S., Choi, Y., Roy, A., & Jeon, W. (2017). Allocating emissions to 4 km and 1 km horizontal spatial resolutions and its impact on simulated NOx and O3 in houston, TX. *Atmospheric Environment*, *164*, 398–415.

Prasad, K., et al. (2016). Development of ANFIS models for air quality forecasting and input optimization for reducing the computational cost and time. *Atmospheric Enviroment*, *128*, 246–262.

Schaefer, J. T. (1990). The critical success index as an indicator of forecasting skill. *Wea. Forecasting*, *5*, 570–575.

Taylan, O. (2013a). Assessing air quality in Jeddah by modeling suspended PM10 concentration. *J. Int. Environ. Appl. Sci*, *8*(3), 326–335.

Taylan, O. (2017). Modeling and analysis of ozone concentration by artificial intelligent techniques for estimating air quality. *Atmospheric Enviroment*, *150*, 356–365.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (methodological)*, *58*(1), 267–288, Wiley.

Tikhonov, A. N., Leonov, A. S., & Yagola, A. G. (1998). *Nonlinear ill-posed problems*. London: Chapman & Hall, ISBN: 0412786605, Retrieved 90818.

U. S. EPA Air Quality Criteria For LEAD (Final Report, 2006). U.S. Environmental Protection Agency, Washington, DC, EPA/600/R-05/144aF-bF, 2006..

Willmott, C. J. (1981). On the validation of models. *Physical Geography*, *2*(2), 184–194.

Zhang, Y., Bocquet, M., Mallet, V., Seigneur, C., & Baklanov, A. (2012). Real-time air quality forecasting, part i: history, techniques, and current status. *Atmospheric Environment*, *60*, 632–655.