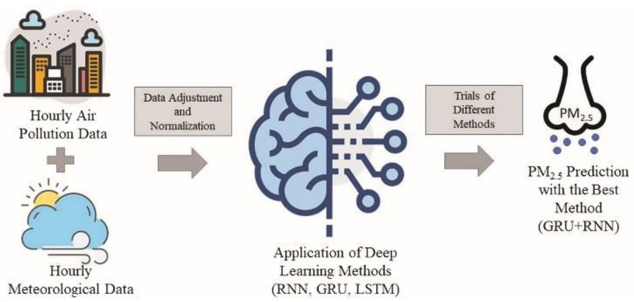
Global NEST Journal, Vol 22, No 1, pp 126-131

Short-term prediction of PM2.5 pollution with deep learning methods

Ayturan Y.A.[1], Ayturan Z.C.[2*], Altun H.O.[3], Kongoli C.[4,5], Tuncez F.D.[6], Dursun S.[2] and Ozturk A.[7]
[1]Department of Electrical and Computer Engineering, Graduate School of Natural and Applied Science, Karatay University, Konya, Turkey
[2]Department of Environmental Engineering, Engineering and Natural Science Faculty, Konya Technical University, Konya, Turkey
[3]Department of Electrical and Electronic Engineering, Engineering Faculty, Karatay University, Konya, Turkey
[4]Earth System Science Interdisciplinary Center (ESSIC), University of Maryland College Park, USA
[5]NOAA/National Environmental Satellite and Data Information Service (NESDIS), College Park, MD, USA
[6]Department of Human Resources and Social Security, Graduate School of Social Sciences, Karatay University, Konya, Turkey
[7]Department of Computer Engineering, Engineering Faculty, Karatay University, Konya, Turkey
Received: 27/08/2019, Accepted: 10/03/2020, Available online: 12/03/2020
*to whom all correspondence should be addressed: e-mail: zcayturan@ktun.edu.tr
https://doi.org/10.30955/gnj.003208

Graphical abstract



Abstract

Particulate matter (PM), classified according to aerodynamic diameter, is one of the harmful pollutants causing health damaging effects. It is considered as cancerogenic by the World Health Organization (WHO) because of the substances found in the chemical composition of PM. In this study, short-term prediction of $PM_{2.5}$ pollution at 1, 2 and 3 hours was modelled using deep learning methods. Three deep learning algorithms and the combination thereof were evaluated: Long-short term memory units (LSTM), recurrent neural networks (RNN) and gated recurrent unit (GRU). Air Quality Monitoring Stations of the Ministry of Environment and Urbanization of Turkey were utilized to obtain the data. Specifically, meteorological and air pollution data were obtained from a monitoring station located in Keçiören District of Ankara. Several trials were conducted using different combinations of RNN, GRU and LSTM models. Pollutant concentrations and meteorological factors were integrated into the model as input parameters to predict $PM_{2.5}$ concentration for 1, 2 and 3 hours. Best results with R^2 of 0.83, 0.7 and 0.63 for 1-, 2-, and 3-hour predictions, respectively, were obtained by using a combination of GRU and RNN models. The results of this study are promising for explaining the effect of different deep learning models on prediction performance.

Keywords: Air pollution, particulate matter, deep learning, prediction, GRU, RNN.

1. Introduction

The chemical compounds that lower the air quality are usually referred to as air pollutants. These compounds may be found in the air in two major forms: in a gaseous form and in a solid form (suspended in air), the latter referred to as Particular Matter (PM). Especially the pollutants which are not originally found in the atmosphere such as dust, gaseous, smell, smoke and fume may affect the health of all living creatures negatively (Güngör *et al.*, 2013). Although there are several air pollutants such as sulphur oxides (SO_x), nitrogen oxides (NO), carbon monoxide (CO) in the atmosphere which are problematic for all ecosystem, particulate matter (PM) is one of the mostly important air pollutants (Vesilind *et al.*, 2010; Boubel *et al.*, 1994).

PM is especially dangerous because of the negative effects on respiratory and nervous system (Krzyzanowsk and Schwela, 1999). PM pollutants are mainly classified according to their aerodynamic diameter. Especially PM_{10} (particles having aerodynamic diameter between 10 µm and 2.5 µm) and $PM_{2.5}$ (particles having aerodynamic diameter smaller than 2.5 µm) have a great importance on assessment of particulate matter in the air (Schnelle *et al.*, 2015). $PM_{2.5}$ may stay suspended in the air during months while PM_{10} may settle in a few hours (World Health Organization, 2005). Moreover, PM_{10} may be filtered in upper respiratory tract while $PM_{2.5}$ may reach to bronchus and create more serious health problems such as heart attack, asthma, premature birth, decrease in lung functions and even death (Karakas, 2015; Wang *et al.*, 2016). In addition to physical features, chemical features of PM pollution are also important. The cancerogenic and toxic substances may be carried on them. Given the health hazards of air pollution, it is important to monitor and predict its level in the atmosphere. This study focuses on the short-term prediction of $PM_{2.5}$ pollution using deep learning architectures.

Deep learning is one of the emerging fields of artificial intelligence. Artificial neural networks, which is a machine learning class, have been widely used to solve complex world problems (Basheer and Hajmeer, 2000; Ayturan *et al.*, 2018). Unfortunately, their prediction performance has been not so promising because of the problems in training of large data sets and disappearance of gradian (Goh *et al.*, 2017). Deep learning is a sub-class of machine learning and it carries machine learning one step beyond. Deep learning may solve problems by using more layers and bigger data sets and processing all layers simultaneously in order to get more accurate results (LeCun *et al.*, 2015). Most of the deep learning models have been developed with respect to the application of steps such as input and output vector determination, transfer function determination, network structure selection, hidden layer determination, weight features and learning algorithm determination (Wang, 2003).

All these positive properties of deep learning make it suitable for modeling and prediction of air pollution. A wide variety of models can be used for this purpose such as long-short term memory units (LSTM), recurrent neural networks (RNN), air quality estimation method based on deep learning (STDL), deep air learning (DAL), convolutional neural networks (CNN) and gated recurrent unit (GRU). There have been several studies on air pollution modelling using deep learning methods. Li *et al.* (2016) estimated PM2.5 using an STDL based model. They used batch auto coders and were able to obtain highly efficient results. In another study, Zhang *et al.* (2016) modelled $PM_{2.5}$, and PM_{10} using CNN and were also able to obtain low average error values. In addition, DAL method for the Beijing city of China was modeled by placing air pollutants and meteorological data from different stations in each divided section using city grid method. Interpolation and property analysis were also added to the models and highly efficient models were developed (Qui *et al.*, 2018). According to a study conducted in South Korea, meteorological data obtained from different stations were used to predict $PM_{2.5}$ by LSTM method. The long-term prediction results of these models were promising (RMSE of 12.41 for 8-hour prediction and 13.54 for 24-hour prediction) (Bui *et al.*, 2018). Moreover, Kok *et al.* (2017) studied O_3 and NO_2 prediction for Aarhus and Brasov cities using LSTM and they were able to obtain successful prediction results with low RMSE values (3.26 for O_3 and 3.79 for NO_2). Another study was conducted for Beijing city of China using LSTM based model. PM2.5 concentration was predicted for 5, 10 and 120 hours and the results were promising (RMSE of 44.15 for 5-hour and 108.4 for 120-hour prediction, respectively) (Reddy *et al.*, 2017). Athira *et al.* (2018) used the LSTM, RNN and GRU to predict $PM_{2.5}$ pollution, their results showed that GRU based models performed better relative to other models.

This study explores the use of RNN, GRU and LSTM models and their combinations to determine the optimal strategy for short-term prediction of $PM_{2.5}$ pollution in Keçiören District of Ankara. The layout of the paper is as follows: Section 2 describes in detail the study area, the meteorological and pollution data used and the methodology. It includes a description of the RNN, GRU and LSTM model application and selection procedures. Section 3 provides a discussion of prediction results, and finally section 4 summarizes the conclusions of the study.

2. Material and methods

2.1. Study area

Keçiören district of Ankara province of Turkey was selected as the study area. Keçiören is the second largest district of Ankara with respect to population. It has a surface area of 189 km^2 and altitude of 950 m. According to the population census in 2016, the population of the district was determined as 903 565 people. It is also one of the most crowded districts of Turkey and its population is more than several cities in Turkey. Because of the high population density, low income and excess usage of coal for heating purposes, it was selected as study area. Figure 1 shows a map of Turkey, Ankara, Keçiören and the location of air quality monitoring station. This station belongs to the Ministry of Environment and Urbanization so the monitoring and maintaining of the measurement devices are suitable for the accurate measurements.

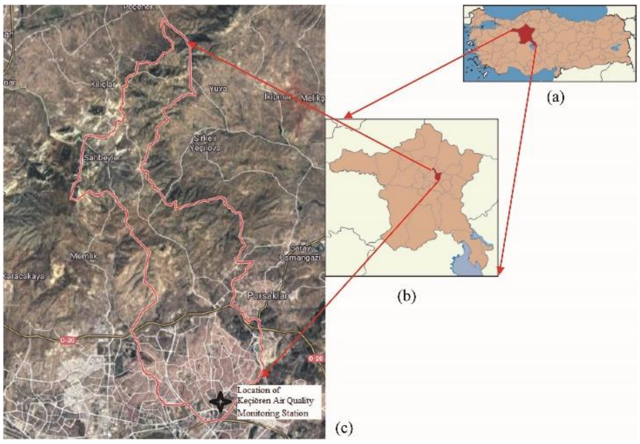


Figure 1. (a) Map of Turkey, (b) Map of Ankara, (c) Map of Keçiören and location of air quality monitoring station

2.2. Data

Data were obtained from a publicly available data sharing system of Turkey Ministry of Environment and Urbanization (MEU, 2019). The hourly data of two years' time period (2017–2018) was taken from this system. The dataset was arranged in cloud environment provided by Google Colab. This data set only represents the meteorological and air pollution parameters measured in that area. There were some missing parameters present in this data together with the extreme values in some parameters (i.e. high wind speeds in the storms, high or low temperature degrees, high pollutant concentrations because of the traffic, construction etc.). Therefore, firstly unnecessary information and missing variables were removed. Next, the data was normalized with max-min normalization before integration into the model so as to standardize values and increase the model performance. A total of 17 parameters used as inputs including meteorological and air pollution variables for the prior hours and 1 output air pollution parameter for the future (forecast) hours. Table 1 presents the input and output parameters used in this study.

Table 1. Input and output parameters and time frames used in models

Input parameters	Time frame of input parameters	Output parameter	Time frame of output parameter
$PM_{2.5}$ concentration (μg/m³)	Previous hours (12 h)	$PM_{2.5}$ concentration (μg/m³)	Next hours (1, 2 and 3 h)
NO concentration (μg/m³)			
NO_2 concentration (μg/m³)			
Cabin temperature (°C)			
Relative humidity (%)			
Cabin humidity (%)			
PM_{10} concentration (μg/m³)			
Sun radiation (W/m²)			
SO_2 concentration (μg/m³)			
NOx concentration (μg/m³)			
Air temperature (°C)			
Air pressure (mbar)			
Wind speed (m/s)			
O_3 concentration (μg/m³)			
Wind direction (degree)			
UVB Radiation (W/m²),			
UVA Radiation (W/m²)			

2.3. Methodology

In order to determine most appropriate model for the data set, short-term prediction models were developed using several combinations of three different deep learning methods: RNN, GRU and LSTM. Figure 2 gives the basic block system of the three models used.

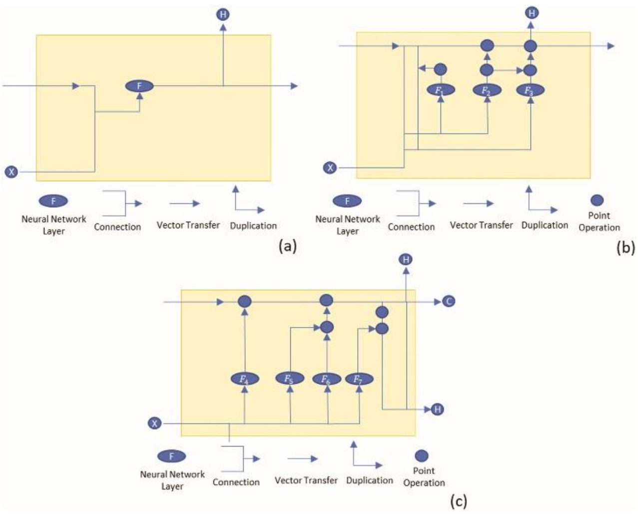


Figure 2. Block system of (a) RNN, (b) GRU and (c) LSTM models (Rathor, 2018)

RNN model mainly consists of one input (X), one output (H) and previous output (Ht-1) and it has no gates present. In this model input and previous output multiplied with each other in the presence of activation function (tanh) (Rathor, 2018). The mathematical equation followed by RNN is given in Equation 1.

$$H = \tanh(H_{t-1} * X) \qquad (1)$$

GRU model is similar with RNN model except with the presence of update gate. GRU also consists of same input (X) and previous output (Ht-1) like RNN. The update gate is used to decide whether previous output will affect final output. New mathematical operations like subtraction, summation, multiplication with new set of weights affect final output (Rathor, 2018). The mathematical equations followed by GRU are given in Equation 2, 3, 4 and 5.

$$F_1 = \sigma\left(W_1 * \left[H_{t-1}, X\right]\right) \qquad (2)$$

$$F_2 = \sigma\left(W_2 * \left[H_{t-1}, X\right]\right) \qquad (3)$$

$$F_3 = tanh\left(W_3 * \left[F_1 * H_{t-1}, X\right]\right) \qquad (4)$$

$$H = \left(1 - F_2\right) * H_{t-1} + \left(F_2 * F_3\right) \qquad (5)$$

LSTM model has two more gates: forget gate and output gate. This means that there is addition of two extra mathematical operations and two extra weight sets in the system. In LSTM, two previous output enters to the system and produce return cycle from output to input (Ct-1 and Ht-1). With the help of LSTM final output was obtained with several control operation (Rathor, 2018). LSTM is the most complicated method within three of them and appropriate for big data sets. The mathematical equations followed by LSTM are given in Equation 6, 7, 8, 9, 10 and 11.

$$F_4 = \sigma\left(W_4 * \left[H_{t-1}, X\right] + b_4\right) \qquad (6)$$

$$F_5 = \sigma\left(W_5 * \left[H_{t-1}, X\right] + b_5\right) \qquad (7)$$

$$F_6 = tanh\left(W_6 * \left[H_{t-1}, X\right] + b_6\right) \qquad (8)$$

$$C = F_4 * C_{t-1} + F_5 * F_6 \tag{9}$$

$$F_7 = \sigma\left(W_7 * \left[H_{t-1}, X\right] + b_7\right) \tag{10}$$

$$H = F_7 * \tanh(C) \tag{11}$$

2.3.1. Model selection

In order to determine the most appropriate algorithm fitted to the data, a number of parameters were kept constant such as block number used in models, prediction period, and programming language. Model block numbers were selected as 50 in all trials. For all trials previous 12 hours of data were integrated into the models and the model predictions 1 hour later were evaluated (Table 2). All models were developed in python programing language. The results of different models were evaluated according to R-squared (R^2), root mean squared logarithmic error (RMSLE), root mean square error (RMSE), mean absolute error (MAE) and standard deviation (STD) values. Several combinations of the three models such as GRU, RNN, LSTM, LSTM+LSTM, RNN+RNN, GRU+GRU, GRU+RNN, LSTM+RNN, LSTM+GRU, GRU+LSTM, RNN+LSTM, RNN+GRU were explored to determine the optimal model. For each model selection, five trials were conducted, and the average values were determined to make selection suitable.

2.3.2. Model application

First, the data set consisting of 500 days over a two-year period (2017–2018) was divided into training and testing subsets. 350 days of data were used for training and the remaining 150 days of data were used for testing. Secondly, the training and testing subsets were divided into input and output parameters. The flow chart of the applied model is given in Figure 3.

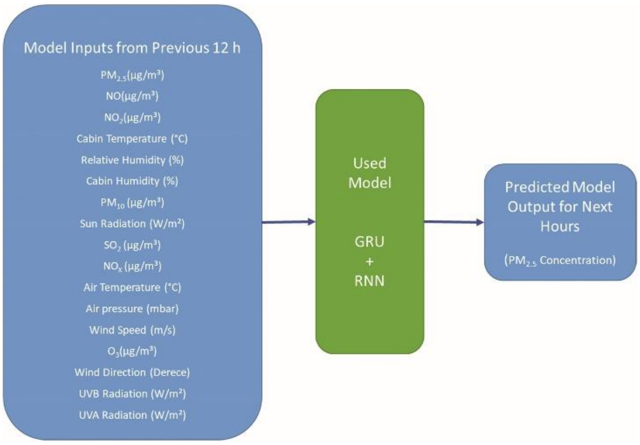


Figure 3. Flowchart of GRU+RNN model

Next, the shape of training and input-output parameters in the test set was generated. This provides the model applicable and definable features. For $PM_{2.5}$ prediction, layers for GRU and RNN were defined in order.

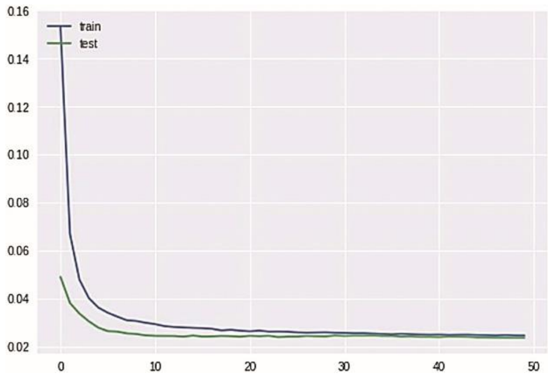


Figure 4. Train and test losses of the model (The x-axis represents the number of epochs while the y-axis represents the losses)

In this model, mean absolute error (MAE) was used as a loss function and Adam Optimization Algorithm was used for the most appropriate alterative for stochastic gradian slope. Adam algorithm is preferred using deep learning in solution of several problems, owing to its' unique features such as low memory requirement and working well with the hyperparameters. In this algorithm exponential weight of past gradients average of the squares of them were calculated and stored. Then, all parameters were updated with respect to the direction of information combined in the memory (Rizwan, 2018). Finally, data argument confirmed by the fit function was arranged and the train and test losses were monitored (Figure 4). The generated model structure is given in Table 2. As seen in Table 2, the time laps used in GRU model is 12 hours, and block number 50 in both GRU and RNN. The number of parameters used in GRU is 10200 and RNN is 5050. Dense layer represents the output layers of 1, 2, 3 hours later. As a result of model application, all test data set could be predicted. At this point, all predicted results were combined with the test data set and the normalization were reversed.

Table 2. GRU+RNN model structure

Layer (type)	Output Shape	Param #
gru_1 (GRU)	(None, 12, 50)	10200
dropout_1 (Dropout)	(None, 12, 50)	0
simple_rnn_1 (SimpleRNN)	(None, 50)	5050
dropout_2 (Dropout)	(None, 50)	0
dense_1 (Dense)	(None, 3)	153

Total params: 15,403
Trainable params: 15,403
Non-trainable params: 0

None

3. Results and discussion

Table 3 shows error statistics of the model selections. The GRU and RNN model combination gives the highest correlation and lowest error statistics: a R^2 of 0.832, RMSLE of 0.398, RMSE of 6.282, MAE of 4.211 and STD of 4.661.

Therefore, this combination (GRU+RNN) was selected as the best method.

Table 3. Summary of the results of different model trials

Model	Evaluation criteria				
	R^2	RMSLE	RMSE	MAE	STD
GRU	0.817	0.414	6.897	4.492	5.233
RNN	0.827	0.410	7.138	4.703	5.370
LSTM	0.801	0.429	7.507	4.664	5.882
LSTM+LSTM	0.815	0.430	6.968	4.630	5.207
RNN+RNN	0.826	0.408	7.203	4.697	5.461
GRU+GRU	0.820	0.405	6.573	4.354	4.923
GRU+RNN	0.832	0.398	6.282	4.211	4.661
LSTM+RNN	0.825	0.414	6.392	4.297	4.732
LSTM+GRU	0.816	0.439	6.933	4.604	5.184
GRU+LSTM	0.813	0.428	7.373	4.799	5.598
RNN+LSTM	0.802	0.528	7.885	5.765	5.379
RNN+GRU	0.823	0.480	6.859	4.853	4.848

3.1. One-hour prediction

Figure 5 shows the regression curve of actual and predicted values (a) and the time evolution of actual and predicted values (b) for the GRU+RNN model applied to the test data.

R^2 = 0.832, RMSLE 0.404, RMSE 6.272 and MAE 4.211 were obtained by one-hour prediction of GRU+RNN model.

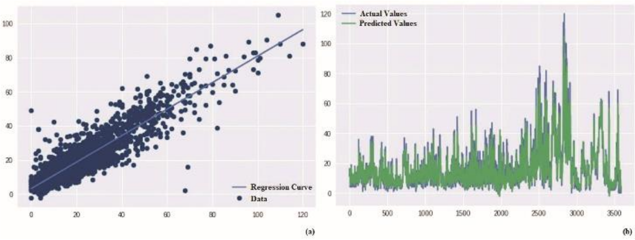


Figure 5. (a) Regression curve of actual (y-axis) and predicted (x-axis) values for one-hour prediction (b) comparison of actual and predicted results (y-axis) with respect to time in seconds (x-axis) for one-hour prediction

3.2. Two-hour prediction

Figure 6 shows similar results as Figure 5 but for a two-hour prediction.

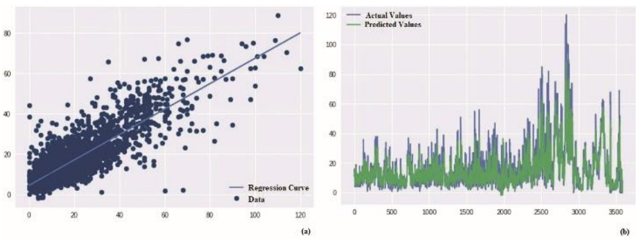


Figure 6. (a) Regression curve of actual (y-axis) and predicted (x-axis) values for two-hour prediction (b) comparison of actual and predicted results (y-axis) with respect to time in seconds (x-axis) for two-hour prediction

R^2 0.709, RMSLE 0.507, RMSE 8.451 and MAE 5.696 were obtained for two-hour prediction of GRU+RNN model.

3.3. Three-hour prediction

Figure 7 shows similar results as Figures 6 and 5 for a three-hour prediction.

R^2 0.611, RMSLE 0.576, RMSE 9.789 and MAE 6.554 were obtained by three-hour prediction of GRU+RNN model.

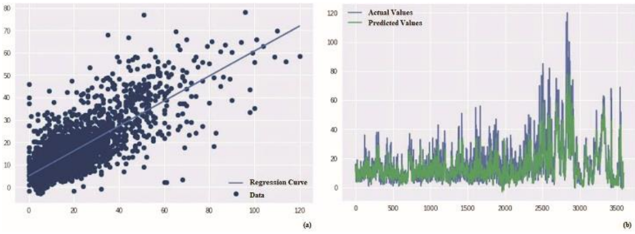


Figure 7. (a) Regression curve of actual (y-axis) and predicted (x-axis) values for three-hour prediction (b) comparison of actual and predicted results (y-axis) with respect to time in seconds (x-axis) for three-hour prediction

The results of this study indicates that the prediction performance of the model is comparable with results reported by previous studies (Bui *et al.*, 2018; Kok *et al.*, 2017; Reddy *et al.*, 2017; Athira *et al.*, 2018). Moreover, the prediction performance of GRU and RNN models is higher for the short-term prediction as also reported by other studies (Athira *et al.*, 2018; Chung, 2014). Data is the most important factor affecting the system performance. In this study, the too high or low (extreme) values in some parameters were kept in the data. Despite the normalization of the data, these extreme values may affect system performance. Furthermore, the number of data was limited because of the missing values in the data set.

4. Conclusions

This study explored the use of RNN, GRU and LSTM deep learning models and their combinations to determine the optimal model selection for short-term prediction of $PM_{2.5}$ pollution in Keçiören District of Ankara. Meteorological and air pollution data were obtained from a monitoring station of Ministry of Environment and Urbanization. The monitoring station used in this study was chosen with respect to the standard deviation (SD) of the data form the all stations found in the Ankara. The data of station which has lowest SD value was chosen as the study station, so the results of this data set are expected to be the best.

The model performance was evaluated based on statistical indices like RMSLE, RMSE, MAE and R^2. The best model was the GRU and RNN combination. Next, the selected model was used to predict $PM_{2.5}$ pollution for 1, 2 and 3 hours.

• Best results were obtained for 1-hour prediction of the selected model with R^2 of 0.832, RMSLE of 0.404, RMSE of 6.272 and MAE of 4.21.

• Model predication performance decreased when the time period for prediction was increased. (R^2 of 0.709 for 2-hour prediction and R^2 of 0.611 for 3-hour prediction)

The results of this study are promising for explaining the effect of different deep learning models on the prediction of air pollution concentrations depending on other pollutants in the air and meteorological factors. However, the predictive performance of the model may be influenced by the presence of the extreme values contained in the data. With the elimination of these extreme values, model performance may be improved. As well as data sets with different meteorological factors and pollution concentrations, a total of 17 input parameters were inserted into the model. The effect of some of these parameters on $PM_{2.5}$ pollution may be less than others. The model performance may be improved with the detection and the elimination of these parameters. Although there are many studies focusing on $PM_{2.5}$ modeling, $PM_{2.5}$ modeling using the deep learning method is a relatively new topic. There are limited number of studies in this area and this study may constitute the basis for further research. With the help of this study, similar models with longer-term prediction performance can be developed with better data.

By means of $PM_{2.5}$ prediction, determination of future concentrations, preparation of control law and regulations, determination of possible pollutant sources, control of sudden pollution episodes and taking preventive precautions are possible.

References

Athira V. Geetha P. Vinayakumar R., Soman.K.P. (2018), Deepairnet: Applying recurrent networks for air quality prediction, *Procedia Computer Science*, **132**, 1394–1403.

Ayturan Y.A., Ayturan Z.C. and Altun H.O. (2018), Air Pollution Modelling with Deep Learning: A Review, *International Journal of Environmental Pollution and Environmental Modelling*, **1**(3), 58–62.

Basheer I.A. and Hajmeer M. (2000), Artificial neural networks: fundamentals, computing, design, and application, *Journal of Microbiological Methods*, **43**(1), 3–31.

Boubel W.R., Fox D.L., Turner D.B. and Stern A.C. (1994), Fundamentals of Air Pollution, Third Edition, Elsevier, USA, pp. 3–153.

Bui T.-C., Le V.-D., Cha S.-K. (2018), A Deep Learning Approach for Forecasting Air Pollution in South Korea Using LSTM, arXiv:1804.07891v3.

Chung J., Gulcehre C., Cho K.H. Bengio Y. (2014), Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. Web Page: https://arxiv.org/abs/1412.3555, Retrieval Date: 20.05.2019.

Goh G.B., Hodas N.O. and Vishnu A. (2017), Deep learning for computational chemistry, *Journal of Computational Chemistry*, **38**(16), 1291–1307, arXiv:1701.04503, DOI: 10.1002/jcc.24764. PMID 28272810.

Güngör A. (2013), Isparta İlindeki Atmosferde Bulunan Kükürt dioksit (SO_2) ve Partikül Madde (PM) Konsantrasyonunun Çoklu Doğrusal Regresyon Yöntemi İle Modellenmesi, Master Thesis, Graduate School of Natural and Applied Sciences, Süleyman Demirel University, Turkey.

Karakas B. (2015), İç ve dış hava ortamlarında partiküler madde (PM10, PM2. 5 ve PM1) konsantrasyonlarının değerlendirilmesi, Master Thesis, Graduate School of Natural and Applied Sciences Hacettepe University, Turkey.

Kok I., Simsek M.U. and Ozdemir S. (2017), A deep learning model for air quality prediction in smart cities, *2017 IEEE International Conference on Big Data (BIGDATA)*, 1983–1990.

Krzyzanowsk M. and Schwela D. (1999), 7-Patterns of Air Pollution in Developing Countries, *Air Pollution and Health*, 105–113, https://doi.org/10.1016/B978-012352335-8/50082-X.

LeCun Y., Bengio Y., Hinton G. (2015), Deep learning. *Nature*, **521**, 436–444, Macmillan Publishers Limited, DOI: 10.1038/ nature14539.

Li X., Peng L., Hu Y. Shao J. Chi T. (2016), Deep learning architecture for air quality predictions, *Environmental Science and Pollution Research*, **23**(22), 22408–22417.

MEU. (2019), Ministry of Environment and Urbanization (MEU)/National Air Quality Monitoring Station, Web page: https://www.havaizleme.gov.tr/Services/AirQuality, Retrieval Date: 05.01.2019.

Qi Z., Wang T., Song G., Hu W., Li X. and Zhang Z.M. (2018), Deep air learning: Interpolation, prediction, and feature analysis of fine-grained air quality, *IEEE Transactions on Knowledge and Data Engineering*, **30**(12), 2285–2297.

Rathor S. (2018), Simple RNN vs GRU vs LSTM - Difference lies in More Flexible control. Web page: https://medium.com/ @saurabh.rathor092/simple-rnn-vs-gru-vs-lstm-difference-lies-in-more-flexible-control-5f33e07b1e57, Retrieval Date: 12.01.2019.

Reddy V., Yedavalli P., Mohanty S. and Nakhat U. (2017), Deep air: Forecasting air pollution in Beijing, China, Web page: https://www.semanticscholar.org/paper/Deep-Air-%3A-Forecasting-Air-Pollution-in-Beijing-%2C-Reddy-Mohanty/ 271279939e545be172f7b754b67d6811c11c8fae, Retrieval Date: 10.04.2019.

Rizwan M. (2018), Adam Optimization Algorithm, Web page: https://engmrk.com/adam-optimization-algorithm/, Retrieval Date: 16.04.2019.

Schnelle K.B., Dunn R.F. and Ternes M.E. (2015), Air pollution control technology handbook, CRC Press, USA.

Vesilind P.A., Morgan S.M. and Heine L.G. (2010), Introduction to Environmental Engineering-SI Version, Cengage Learning, USA.

Wang S.-C. (2003), Artificial neural network, In Interdisciplinary computing in java programming, Springer, pp. 81–100.

Wang X.-D., Zhang X.-M., Zhuang S.-W., Luo Y., Kang S. and Liu Y.-L. (2016), Short-term effects of air pollution on acute myocardial infarctions in Shanghai, China, 2013–2014, *Journal of Geriatric Cardiology*, **13**(2), 132–137, DOI: 10.11909/j.issn.1671-5411.2016.02.005.

World Health Organization. (2005), WHO Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide, Global update 2005, Summary of risk assessment. Web page: https://apps.who.int/iris/bitstream/handle/106 65/69477/WHO_SDE_PHE_OEH_06.02_eng.pdf;jsessionidC9 2A31EB7B3B80C7FF31A5CD92E78B0A?sequence1, Retrieval Date: 25.05.2019.

Zhang C., Yan J., Li C., Rui X., Liu L. and Bie R. (2016), On estimating air pollution from photos using convolutional neural network. In Proceedings of the 2016 ACM on Multimedia Conference, 297–301.