Prediction of the number of asthma patients using environmental factors based on deep learning algorithms

Hyemin Hwang[1], Jae-Hyuk Jang[2], Eunyoung Lee[3], Hae-Sim Park[2] and Jae Young Lee[4*]

Abstract

Background   Air pollution, weather, pollen, and influenza are typical aggravating factors for asthma. Previous studies have identified risk factors using regression-based and ensemble models. However, studies that consider complex relationships and interactions among these factors have yet to be conducted. Although deep learning algorithms can address this problem, further research on modeling and interpreting the results is warranted.

Methods   In this study, from 2015 to 2019, information about air pollutants, weather conditions, pollen, and influenza were utilized to predict the number of emergency room patients and outpatients with asthma using recurrent neural network, long short-term memory (LSTM), and gated recurrent unit models. The relative importance of the environmental factors in asthma exacerbation was quantified through a feature importance analysis.

Results   We found that LSTM was the best algorithm for modeling patients with asthma. Our results demonstrated that influenza, temperature, PM_{10}, NO_2, CO, and pollen had a significant impact on asthma exacerbation. In addition, the week of the year and the number of holidays per week were an important factor to model the seasonality of the number of asthma patients and the effect of holiday clinic closures, respectively.

Conclusion   LSTM is an excellent algorithm for modeling complex epidemiological relationships, encompassing nonlinearity, lagged responses, and interactions. Our study findings can guide policymakers in their efforts to understand the environmental factors of asthma exacerbation.

Keywords   Recurrent neural network, Long short-term memory, Gated recurrent unit, Air pollution, Asthma, Influenza

*Correspondence:
Jae Young Lee
jaeylee@ajou.ac.kr
[1] Environmental Engineering Department, Ajou University, Suwon 16499, Korea
[2] Department of Allergy and Clinical Immunology, Ajou University School of Medicine, Suwon 16499, Korea
[3] Department of Neurology, McGovern Medical School, The University of Texas Health Science Center at Houston, Houston, TX 77030, USA
[4] Environmental and Safety Engineering Department, Ajou University, 206, World Cup-ro, Yeongtong-gu, Suwon 16499, Korea

Background

Asthma is one of the most prevalent respiratory diseases that have a significant public health burden. According to the Global Asthma Report 2022, 262 million people were affected by asthma, and 461 thousand people died from asthma worldwide in 2019 [1]. In addition, asthma is a chronic disease that seriously reduces patients' quality of life but has no definitive cure [2].

Due to the severity of asthma, many previous studies have attempted to understand the risk factors that exacerbate asthma, and various environmental factors such as air pollution, tobacco smoke, weather, allergens such as pollen, and pathogens such as influenza viruses have

Hwang *et al. Respiratory Research*   (2023) 24:302

Page 2 of 9

been identified as culprits for asthma exacerbation [3]. These studies mostly used regression-based statistical models such as the generalized linear model (GLM) [4], generalized additive model (GAM) [5], and distributed lag nonlinear model (DLNM) [6] and ensemble-based machine learning models such as the random forest (RF) [7] and gradient boosting machine (GBM) [8]. Cassino et al. analyzed tobacco use and O_3-associated emergency room visits for asthma in New York City based on a Poisson regression model [9]. Lee et al. studied the effects of air pollutants, pollen, weather conditions, and viruses on the number of emergency room patients with asthma in Seoul, South Korea, using DLNM [10]. Chen et al. studied the lagged nonlinear relationship between temperature and adult asthma hospitalizations in Beijing using the DLNM [2]. Sun et al. studied the association between pollen (trees, weeds, and grasses) and asthma in North Carolina using DLNM [11]. Jeddi et al. compared machine learning models for pediatric asthma diagnosis by considering environmental factors such as mites, cold air, strong odors, and mold [12]. Although previous studies have succeeded in identifying risk factors and modeling the risk of asthma using conventional statistical and machine learning algorithms, our understanding and modeling accuracy remain insufficient because of the complexities associated with nonlinearity, lagged relationships, interactions between factors, multicollinearity, and various confounders.

To model the relationship with higher accuracy, researchers have started to utilize state-of-the-art deep learning algorithms such as recurrent neural networks (RNNs) [13], long short-term memory (LSTM) [14], and gated recurrent units (GRUs) [15]. Woo et al. predicted the peak expiratory flow rate in children with asthma using real-time indoor air pollution data using RNN, GRU, and deep neural network [16]. Kim et al. studied the association between indoor particulate matter (PM) and asthma attacks in children in South Korea using the LSTM [17]. Chang and Ku used LSTM to predict the daily number of patients with asthma affected by weather and air pollution in Seoul, South Korea [18]. As research based on deep learning algorithms in the field of public health is still in its early stages, more research on modeling methodologies and epidemiological results from the models is necessary.

This study examined the association between the number of patients with asthma and 18 environmental factors in South Korea between 2015 and 2019 using the RNN, LSTM, and GRU algorithms. Eighteen environmental factors were categorized into air pollution, weather, pollen, and influenza. The accuracy of the model developed in this study was compared with that of conventional algorithms (GLM, GAM, RF, and GBM), and

permutation feature importance analysis was performed to identify the critical factors in asthma exacerbation and understand the interaction between various factors.

Methods

Data collection

Weekly counts of patients with asthma in South Korea from 2015 to 2019 were collected by the Health Insurance Review and Assessment Service. Patients with asthma were defined as those aged 17 years or older who visited a healthcare facility and were diagnosed with asthma (ICD-10 codes J45, J46, J820, and J828). The number of outpatients and emergency room (ER) patients with asthma was determined separately. Environmental data were collected from South Korea from 2015 to 2019. Daily air pollutant concentrations of CO, NO_2, O_3, PM_{10}, $PM_{2.5}$, and SO_2 were collected from 556 nationwide measurement stations by the Korea Environment Corporation. Daily meteorological data on mean temperature, minimum temperature, maximum temperature, diurnal temperature range, humidity, precipitation, solar radiation, and wind speed were collected from 100 measurement stations in South Korea by the Korea Meteorological Administration. The data collected from multiple measurement stations spread across the entire nation were averaged to obtain national air pollution and meteorological data. Any missing values, if present, from the measurement stations were excluded during the averaging. Data regarding the hazard index of the daily pollen concentration from oaks, pines, and grasses were obtained from the Korea Meteorological Administration. This index was designed to forecast pollen concentrations based on meteorological and environmental factors (see Additional file 1: Table S1 for a detailed description). Weekly numbers of influenza and Middle East respiratory syndrome (MERS) patients were collected by the Korea Centers for Disease Control and Prevention. Patients with influenza were defined as those diagnosed with influenza (ICD-10 codes J10.0–J11.8) or those who had an influenza-like illness (ILI). ILI is defined by WHO as a respiratory infection with onset within the past ten days and a fever of ≥ 38 °C and cough or sore throat. Daily and regional data were averaged and converted into weekly national data for South Korea.

Prediction of patients with asthma using environmental factors based on deep learning algorithms

To model the relationship between the number of patients with asthma (outpatients or ER patients) and environmental factors, we used 18 environmental factors and four potential confounders as input data. The 18 environmental factors included six air pollutant concentrations (CO, NO_2, O_3, PM_{10}, $PM_{2.5}$, and SO_2), three

Hwang *et al. Respiratory Research* (2023) 24:302

Page 3 of 9

pollen concentrations (pollen from oaks, pines, and grasses), eight meteorological conditions (mean temperature, minimum temperature, maximum temperature, diurnal temperature range, humidity, precipitation, solar radiation, and wind speed), and the number of patients with influenza. The four confounders were the week of the year, date, number of holidays per week, and number of patients with MERS. The week of the year is an indicator of where a particular week falls numerically within a year. The first week (week 1) of the year is defined as the week containing the first Wednesday of the year. For each year from 2015 to 2019, the week of the year is numbered from 1 to 52. The date is a number that inicidates a specific point in time, and it is calculated as the number of days that have passed since January 1, 1970. These were used to model the confounding effects of seasonality, long-term trends, holidays, and the 2015 MERS outbreak. All input factors were preprocessed with minimum–maximum normalization before modeling.

We used the RNN, LSTM, and GRU as the deep learning algorithms. The model consisted of four layers: an input layer, two hidden layers, and an output layer. Recurrent cells were used only in the input layer, whereas simple, fully connected neural network cells without recurrent connections were used in the hidden and output layers to simplify the model (see Additional file 1: Figure S1 for the topology). Dropout techniques were applied to all layers to prevent overfitting [19, 20]. The optimum model size and dropout rates were selected empirically by finding the best model among various candidates (see Additional file 1: Table S2 for the hyperparameter candidates). We used walk-forward expanding window cross-validation, where data from 2015–2016, 2015–2017, and 2015–2018 were used for training and data from 2017, 2018, and 2019 were used for testing, respectively. Walk-forward cross-validation is a well-known validation method for time-series data to remove the possibility of prediction leakage [21, 22]. The length of time steps in RNN, LSTM, and GRU was set as 5 weeks to model the long-term lagged effects of environmental factors on asthma, and the training learning rate was set as 0.004. During training, the mean squared error (MSE) for the test set was monitored, and training was stopped when the observed MSE did not improve after 50 epochs. The modeling and training were implemented using the Python packages "keras" and "tensorflow" [23, 24].

Comparison with conventional modeling methods

We compared the R^2 values of the neural network models with those of the GLM, GAM, RF, and GBM. The input and output variables used for modeling were identical to those used for the deep learning algorithms. The GLM and GAM were fitted using the maximum likelihood method under the assumption of a quasi-Poisson distribution. For the modeling and training of RF and GBM, the "sklearn" package of Python was used. The model hyperparameters for RF and GBM were optimized using "best_estimator_" of the "Grid method for the model's trainingSearchCV" function (see Additional file 1: Table S3 for the candidates) [25].

Permutation feature importance

After modeling, we evaluated the importance of all input features in the final model based on the permutation feature importance method [26], where feature importance is defined as the increase in the MSE when the values of a single feature are temporally shuffled. This method helps identify features with high contributions in predicting the output. In addition, we define the interaction between the two features as follows:

$$I = FI_{AB} - FI_A - FI_B \tag{1}$$

here, I is the interaction between two features, A and B, FI_{AB} is the increase in MSE with both A and B shuffled, FI_A is the increase in MSE with A shuffled, and FI_B is the increase in MSE with B shuffled.

Results

General analysis of the number of patients with asthma

Figure 1 shows the weekly numbers of outpatients and ER patients with asthma in South Korea for each year from 2015 to 2019 (see Additional file 1: Figure S2 for the 5-year curves and Additional file 1: Table S4 for descriptive statistics). From 2015 to 2019, the number of outpatients with asthma decreased, whereas the number of ER patients with asthma increased. The number of patients with asthma showed seasonal variability (Fig. 1), with the lowest in summer (July and August) and the highest in winter (December and January) and spring (March and April). Additional file 1: Figures S3-S8 show the time-series curves of various independent variables, and Additional file 1: Tables S5 and S6 show the descriptive statistics for air pollutant concentrations and climate conditions in South Korea during the study period.

The areas shaded in green in Fig. 1 (between Weeks 1 and 4, and between Weeks 48 and 52) show the weeks when the number of influenza patients surged, especially in 2018 (see Additional file 1: Figure S3 to find the surge in influenza patients). As shown in the figure, the number of patients with asthma increased during this period. The areas shaded in blue show the weeks in which the two biggest holidays in Korea, Lunar New Year's Day (between Weeks 4 and 8) and Korean Thanksgiving Day (between Weeks 36 and 41), were located. During the holidays, the number of outpatients shows downward spikes due to the holiday clinic
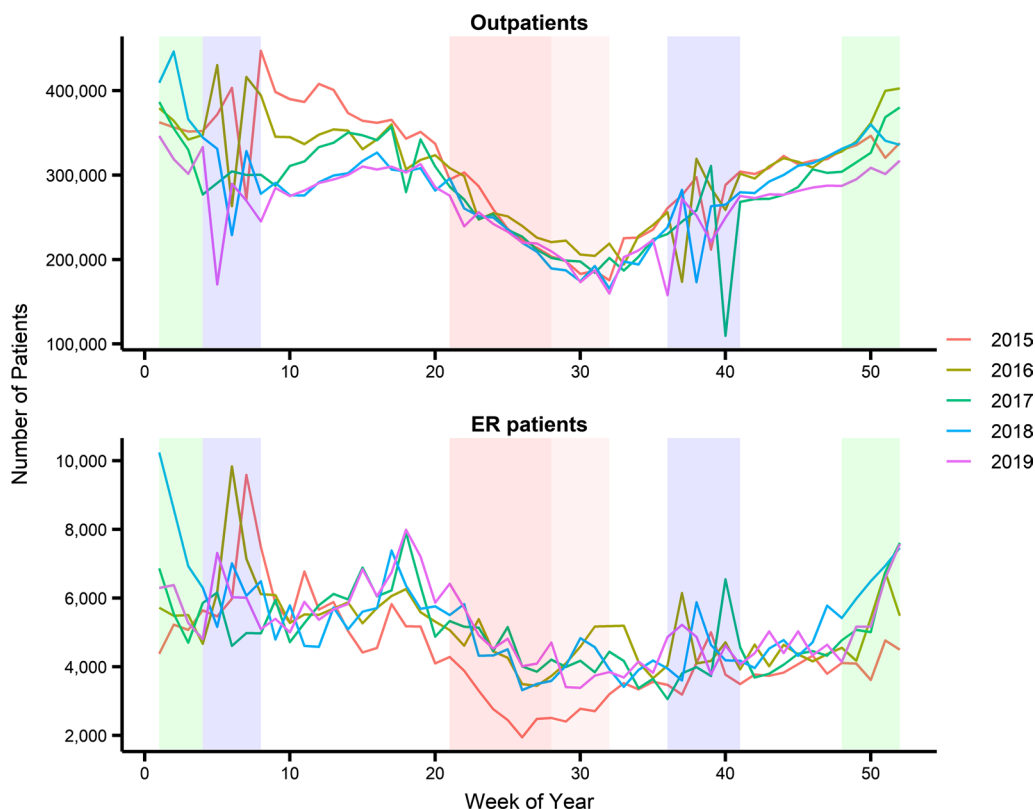
Hwang *et al. Respiratory Research*   (2023) 24:302

Page 4 of 9

Fig. 1 The number of outpatients and ER patients with asthma in South Korea each year from 2015 to 2019. The areas shaded in green, red, and blue highlighted the weeks when the numbers of influenza patients, patients with MERS, and holidays surged, respectively. The peaks in the green and blue periods correspond to the peaks in the number of influenza patients and the number of holidays in a week. The smaller number of ER visits during the red period in 2015 compared to 2016–2019 can be attributed to the MERS outbreak that occurred in 2015

closures. In contrast, the number of ER patients shows upward spikes due to the "balloon effect" of the holiday clinic closures. This study used the number of holidays per week (Additional file 1: Figure S3) to model the confounding effects of holidays. The area shaded in darker red in Fig. 1 (between weeks 21 and 28) represents the weeks in 2015 when the MERS outbreak occurred in South Korea (see Additional file 1: Figure S3 for the number of patients with MERS). MERS is one of the 16 diseases classified as a "Class 1 infectious disease," a term used for diseases of significant health importance owing to its high mortality rates; it was the only Class 1 infectious disease that caught public attention in South Korea during the study period. The area in lighter red (between weeks 28 and 32) shows the weeks between the last identification of patients with MERS and South Korea's official declaration of a "de facto end" to MERS (July 28, 2015). During this period, the number of ER visits in many tertiary hospitals where patients with MERS were admitted was significantly reduced in 2015 because people were afraid of being infected. Since the MERS outbreak only occurred in 2015, the number of

ER visits between weeks 21 and 32 in 2015 was apparently smaller than those during 2016–2019 as shown in Fig. 1.

This is a confounding effect of Class 1 infectious diseases, such as MERS, and we attempted to model this effect using the number of patients as an independent variable. Note that there is a four-week lag between the last identification of the patient with MERS (Week 28) and getting back to normal with the declaration of "de facto end" (Week 32). LSTM and GRU are suitable for modeling such long lags using their long-term memory.

Modeling patients with asthma based on deep learning algorithms

To determine the best model for predicting the number of patients with asthma, we generated and trained 648 models of RNN, LSTM, and GRU (216 models each) and evaluated them with R^2 based on a walk-forward cross-validation framework. Figure 2 shows the R^2 histograms of 216 RNN, LSTM, and GRU models. The LSTM and GRU models performed better than the RNN models in predicting the number of patients with asthma. This
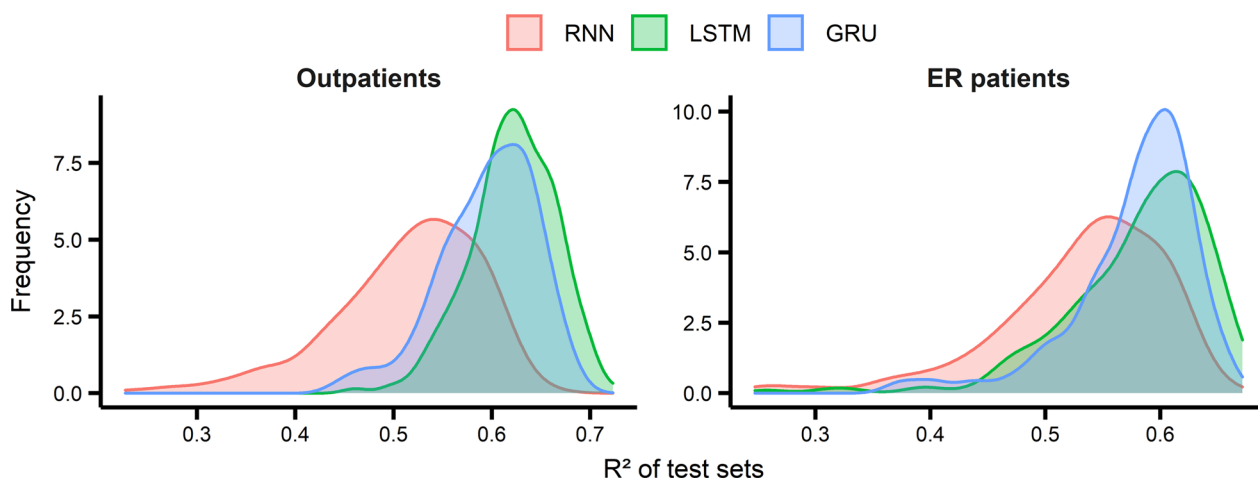
Fig. 2 Performance (R^2) histograms of 216 RNN, LSTM, and GRU models

may be due to the internal gates that solve the vanishing gradient problem in the RNN. Our results show that, in general, LSTM models perform slightly better than GRU models. This may be due to the higher number of gates in the LSTM model than in the GRU model (three in LSTM and two in GRU), providing more flexibility in modeling.

Additional file 1: Figure S9 shows the performance (R^2) scatter plot for the 648 RNN, LSTM, and GRU models. The top 10% of the models are located in Area 1. Additional file 1: Table S7 lists all hyperparameter values of the models in Area 1. For the final model, we selected the model that provided the best average performance for outpatients and ER patients (R^2 values of 0.723 for outpatients and 0.650 for ER patients).

Performance comparison of asthma patient predictive models

Table 1 compares the performances of the final RNN, LSTM, and GRU models with those of conventional algorithms. The hyperparameters used in the final model are listed in Additional file 1: Table S7. Based on our results, LSTM was the best model for outpatients and

ER patients, with R^2 of 0.723 and 0.650, respectively. This algorithm performed better than the other algorithms investigated in this study. The ensemble-based models (RF and GBM) performed the worst for both outpatients and ER patients (R^2 of − 0.321–0.583), whereas the regression-based models (GLM and GAM) ranked in the middle (R^2 of 0.631–0.706).

The R^2 gap between the training and test sets was the smallest for the GLM, indicating the least overfitting among all algorithms. This is likely the result of linear modeling, which has a lower chance of overfitting than nonlinear modeling. The R^2 gap was most significant for RF and GBM, indicating considerable overfitting. The gaps for RNN, LSTM, and GRU were smaller than those for GAM, RF, and GBM, despite the complexity and flexibility of the models. This results from the dropout and early stopping techniques implemented in training.

Feature importance analysis

Figure 3 shows the results of the feature importance analysis for outpatients and patients in the ER. In Fig. 3, the blue dashed line shows the model's baseline MSE, and

Table 1 Performance (R^2) comparison among various algorithms for training and test sets

Model	Training		Test		Gap (Training–Test)	
	Outpatients	ER patients	Outpatients	ER patients	Outpatients	ER patients
GLM	0.870	0.792	0.672	0.637	0.198	0.155
GAM	0.943	0.888	0.706	0.631	0.237	0.257
RF	0.857	0.508	0.540	-0.321	0.317	0.829
GBM	0.957	0.907	0.583	-0.128	0.374	1.035
RNN	0.894	0.803	0.625	0.606	0.269	0.197
LSTM	0.886	0.861	0.723	0.650	0.163	0.211
GRU	0.905	0.827	0.673	0.651	0.232	0.176

Hwang *et al. Respiratory Research*   (2023) 24:302
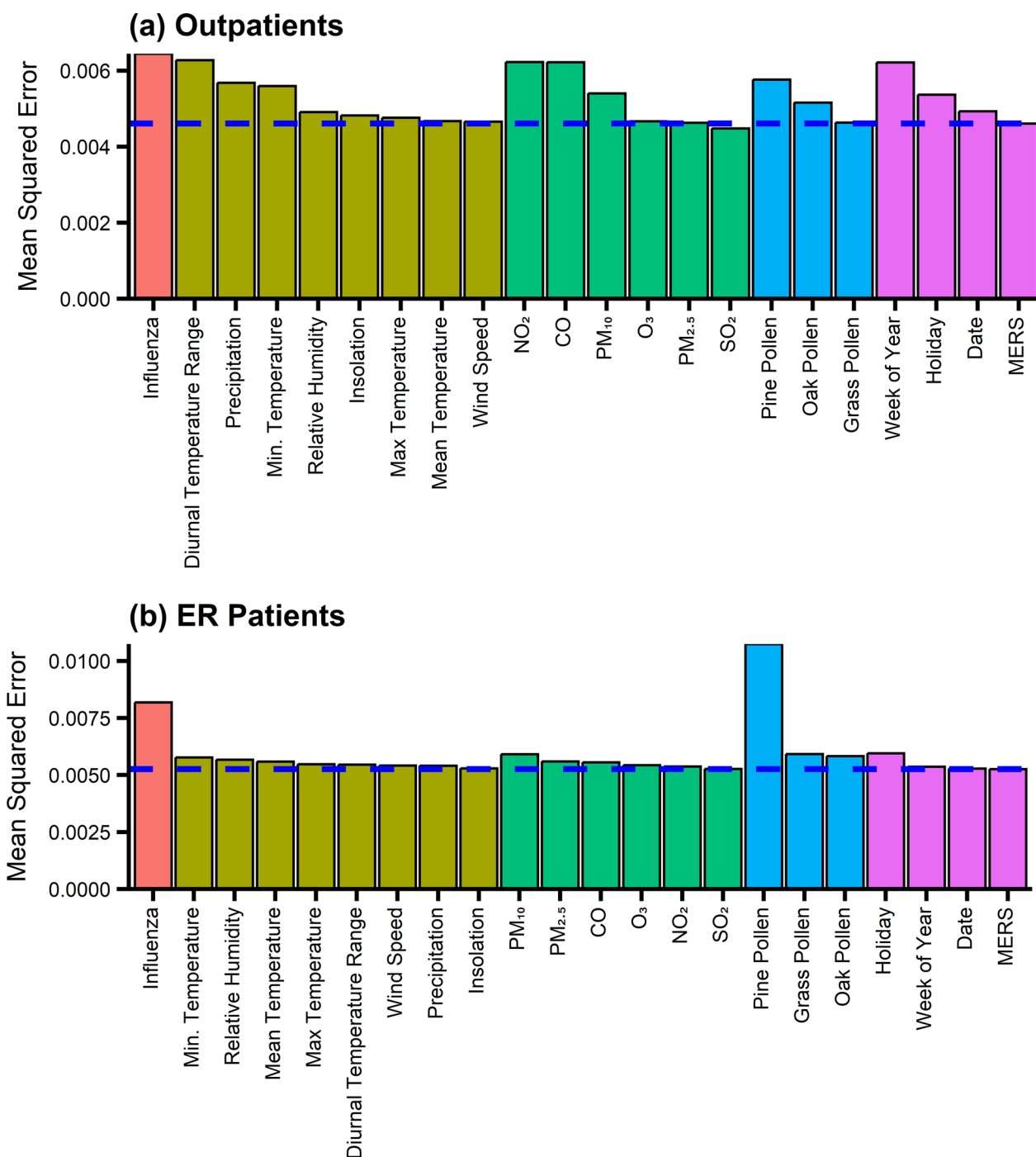
Page 6 of 9

Fig. 3 Feature importance of (**a**) outpatients and (**b**) ER patients

any feature that yielded a higher MSE than the baseline after shuffling was considered significant (the higher the value, the more important it is). Influenza was one of the most important factors for outpatients and ER patients among the various environmental factors. Influenza can cause airway swelling, trigger asthma attacks, and

exacerbate symptoms [27]. Temperature substantially impacted the number of outpatient visits and had little effect on the number of ER visits. This can be interpreted as the temperature being associated more with the gradual development of asthma than the acute development of asthma, which causes patients to visit non-emergency

Hwang *et al. Respiratory Research*   (2023) 24:302

Page 7 of 9

care facilities. Among four temperature-related factors (mean, maximum, minimum, and diurnal temperature range), the diurnal temperature range and minimum temperature were the two most critical factors affecting outpatients with asthma. Air pollutants such as NO_2, CO, and PM_{10} had a significant impact on outpatients while they had little impact on ER patients. This may indicate that asthma exacerbation attributable to NO_2, CO, and PM_{10} is not severe enough for patients to visit ER. Pine pollen had a substantial impact on the number of ER patients with asthma, whereas it had a relatively smaller impact on the number of outpatients with asthma. The association between pine pollen and ER patients with asthma can be observed in Additional file 1: Figure S8 during the pine pollen season (between weeks 16 and 20). The week of the year and the number of holidays per week were also important in modeling the seasonal variability and confounding effect of holiday clinic closure, respectively. Additional file 1: Figure S10 shows the two-dimensional feature importance analysis. Our result indicates that the simultaneous exposure to both NO_2 and CO has a synergic effect on asthma exacerbation. However, there is no significant interaction effect among other environmental factors.

Discussion

The climate is known for impacting asthma directly through the airway response to climate or indirectly through altered exposure to air pollutants, allergens, and pathogens [28]. For example, high temperatures increase the ambient ozone through a photochemical reaction [29]. In addition, abrupt changes in weather and an extensive diurnal temperature range can increase the risk of asthma by affecting inflammatory responses and immunity [30–32]. In this study, we examined the association between outpatients with asthma and climate, and we showed that diurnal temperature range and minimum temperature are important factors in modeling. This result agrees with the results of previous studies. Chen et al. showed that both low and high temperatures were associated with an increased risk of asthma, whereas the majority of the burden was attributable to moderate cold exposures [2]. Xu et al. and Kim et al. studied the relationship between the diurnal temperature difference and asthma [33, 34].

Air pollutants are associated with asthma through both direct and indirect mechanisms [35]. The infiltration of air pollutants can directly trigger inflammation and increase oxidative stress, which may lead to cell and tissue damage in airways [36–38]. Air pollutants can also be involved in indirect mechanisms interacting with inhaled pathogens and allergens, thereby increasing the risk of infections and allergic reactions [35]. The influence of NO_2 and CO on asthma patients was observed in this study, which agrees with a previous study demonstrating the impact of exposure to traffic emissions such as NO_2 and CO on asthma [39]. Additionally, PMs may act as a container for allergens and pathogens such as pollen, fungal spores, and viruses, delivering them deep into the airways [40, 41]. In this study, the risk of asthma exacerbation with PM_{10} was higher than that with $PM_{2.5}$, which is consistent with the results of a previous study by Tecer et al. [42].

Influenza infection can trigger an immune response by releasing cytokines and increasing susceptibility to asthma [27, 43]. Several studies have reported that influenza is associated with asthma exacerbation in adults [10, 44, 45].

This study had a few limitations. First, public environmental data was measured at official stations, instead of personal exposure data. This may have resulted in the underestimation of the actual impact of exposure to asthma-exacerbating environments. Second, demographic factors (such as age, gender, and sex) and socioeconomic conditions (such as occupation, income level, and educational attainment) were not considered in this study because such information was unavailable. Additional studies are warranted to consider such factors in modeling to increase the accuracy.

Conclusions

This study is the first to analyze the association between outpatients and ER patients with asthma and 18 environmental factors, including air pollutants, weather conditions, pollen, and influenza, in South Korea. Additionally, it proved the relative and quantitative importance of all 18 factors in terms of asthma exacerbation. Models with various hyperparameter values were evaluated to optimize the deep learning algorithm. With the optimal hyperparameters, we found that LSTM was the best model for predicting patients with asthma among the eight algorithms studied. It can model nonlinear lagged relationships with interactions between features without causing multicollinearity and overfitting problems. From feature importance analysis, we found out that influenza and pine pollen were the two most important factors exacerbating asthma in outpatients and ER patients.

Abbreviations

GLM	Generalized linear model
GAM	Generalized additive model
DLNM	Distributed lag nonlinear model
RF	Random forest
GBM	Gradient boosting machine
RNN	Recurrent neural network
LSTM	Long short-term memory
GRU	Gated recurrent unit
PM	Particulate matter

Hwang *et al. Respiratory Research*   (2023) 24:302

Page 8 of 9

ER	Emergency room
ILI	Influenza-like illness
MERS	Middle East respiratory syndrome
MSE	Mean squared error
R^2	Coefficient of determination

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12931-023-02616-x.

Additional file 1. Table S1. Information on pollen. **Table S2.** Hyperparameter candidates for the neural network models. **Table S3.** Hyperparameters considered for RF and GBM. **Table S4.** Descriptive statistics for the number of asthma patients in South Korea between 2015 and 2019. **Table S5.** Descriptive statistics for the air pollutant concentrations in South Korea between 2015 and 2019. **Table S6.** Descriptive statistics for the climate conditions in South Korea between 2015 and 2019. **Table S7.** The hyperparameter values of models in Area 1 of Figure S9 and the hyperparameter values selected for the final model. **Figure S1.** Topology of deep learning algorithms. **Figure S2.** The number of asthma outpatients and ER patients in South Korea between 2015 and 2019. The black line shows the linear regression fit of the patient data. **Figure S3.** The numbers of asthma patients, influenza patients, MERS patients, holidays in a week in South Korea in each year from 2015 to 2019. The areas shaded in green, red, and blue highlighted the weeks, when the numbers of influenza patients, MERS patients and holidays surged, respectively. **Figure S4.** The number of asthma patients and pollutant concentrations (CO, O_3, and SO_2) in South Korea in each year from 2015 to 2019. **Figure S5.** The number of asthma patients and pollutant concentrations (NO_2, PM_{10}, and $PM_{2.5}$) in South Korea in each year from 2015 to 2019. **Figure S6.** The number of asthma patients, temperature, and relative humidity in South Korea in each year from 2015 to 2019. **Figure S7.** The number of asthma patients, precipitation, wind speed, and insolation in South Korea in each year from 2015 to 2019. **Figure S8.** The number of asthma patients and pollen hazard index in South Korea in each year from 2015 to 2019. **Figure S9.** Performance (R^2) scatter plot of modeling outpatients (x axis) and ER patients (y axis) for 648 models of RNN, LSTM, and GRU. The red dashed lines indicate the 90 percentile of R^2 for outpatients and ER patients. The four areas that were divided with the red dashed lines were denoted as Areas 1 to 4. **Figure S10.** Observed and predicted number of asthma patients for outpatient and ER patients.

Author contributions

HH: Data curation, Formal analysis, Methodology, Software, Visualization, Roles/Writing—original draft. J-HJ: Validation. EL: Data curation. H-SP: Conceptualization, Validation. JYL: Conceptualization, Supervision, Validation, Funding acquisition, Methodology, Writing—review & editing.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Network TGA. The global asthma report 2022. Int J Tuberc Lung Dis. 2022;26:1–104.
2. Chen Y, Kong D, Fu J, Zhang Y, Zhao Y, Liu Y, Chang Z, Liu Y, Liu X, Xu K, Jiang C, Fan Z. Associations between ambient temperature and adult asthma hospitalizations in Beijing, China: a time-stratified case-crossover study. Respir Res. 2022;23:1–2.
3. D'Amato G, Holgate ST, Pawankar R, Ledford DK, Cecchi L, Al-Ahmad M, Al-Enezi F, Al-Muhsen S, Ansotegui I, Baena-Cagnani CE, Baker DJ, Bayram H, Bergmann KC, Boulet LP, Buters JT, D'Amato M, Dorsano S, Douwes J, Finlay SE, Garrasi D, Gómez M, Haahtela T, Halwani R, Hassani Y, Mahboub B, Marks G, Michelozzi P, Montagni M, Nunes C, Oh JJ, Popov TA, Portnoy J, Ridolo E, Rosário N, Rottem M, Sánchez-Borges M, Sibanda E, Sienra-Monge JJ, Vitale C, Annesi-Maesano I. Meteorological conditions, climate change, new emerging factors, and asthma and related allergic disorders. A statement of the World Allergy Organization. World Allergy Organ J. 2015;8:1–52.
4. McCullagh P, Nelder JA. Generalized linear models. USA: Springer; 1983.
5. Wood SN. Generalized additive models: an Introduction with R. 2nd ed. Boca Raton: CRC Press; 2017.
6. Gasparrini A. Distributed lag linear and non-linear models in R: the package dlnm. J Stat Softw. 2011;43:1–20.
7. Ayyadevara VK. Random forest. In: Pro machine learning algorithms: a hands-on approach to implementing algorithms in Python and R. Berkeley, CA: Apress; 2018. p. 105–16.
8. Prettenhofer P, Louppe G. Gradient boosted regression trees in scikit-learn. InPyData 2014;2014.
9. Cassino C, Ito K, Bader IR, Ciotoli C, Thurston G, Reibman JO. Cigarette smoking and ozone-associated emergency department use for asthma by adults in New York City. Am J Respir Crit Care Med. 1999;159:1773–9.
10. Lee SW, Yon DK, James CC, Lee S, Koh HY, Sheen YH, Oh JW, Han MY, Sugihara G. Short-term effects of multiple outdoor environmental factors on risk of asthma exacerbations: age-stratified time-series analysis. J Allergy Clin Immunol. 2019;144:1542-50.e1.
11. Sun X, Waller A, Yeatts KB, Thie L. Pollen concentration and asthma exacerbations in Wake County, North Carolina, 2006–2012. Sci Total Environ. 2016;544:185–91.
12. Jeddi Z, Gryech I, Ghogho M, El Hammoumi M, Mahraoui C. Machine learning for predicting the risk for childhood asthma using prenatal, perinatal, postnatal and environmental factors. Healthcare (Basel). 2021;9:1464.
13. Medsker LR, Jain LC, editors. Recurrent neural networks: design and applications. Boca Raton: CRC Press; 1999.
14. Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput. 1997;9:1735–80.
15. Cho K, van Merrienboer B, Bahdanau D, Bengio Y. On the properties of neural machine translation: Encoder-decoder approaches. arXiv preprint arXiv:1409.1259; 2014.
16. Woo J, Lee JH, Kim Y, Rudasingwa G, Lim DH, Kim S. Forecasting the effects of real-time indoor PM2.5 on peak expiratory flow rates (PEFR) of asthmatic children in Korea: a deep learning approach. IEEE Access. 2022;10:19391–400.
17. Kim D, Cho S, Tamil L, Song DJ, Seo S. Predicting asthma attacks: effects of indoor PM concentrations on peak expiratory flow rates of asthmatic children. IEEE Access. 2019;8:8791–7.
18. Chang M, Ku Y. LSTM model for predicting the daily number of asthma patients in Seoul, South Korea, using meteorological and air pollution data. Environ Sci Pollut Res Int. 2023;30:37440–8.
19. Billa J. Dropout approaches for LSTM based speech recognition systems. In: 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE; 2018. p. 5879–83.

Hwang *et al. Respiratory Research*   (2023) 24:302

Page 9 of 9

20. Pham V, Bluche T, Kermorvant C, Louradour J. Dropout improves recurrent neural networks for handwriting recognition. In: 2014 14th International Conference on Frontiers in Handwriting Recognition. IEEE; 2014. p. 285–90.

21. Yoon H. Time series data analysis using wavenet and walk forward validation. J Korea Soc Simul. 2021;30:1–8.

22. Tran TN, Phuc DT. Grid search of multilayer perceptron based on the walk-forward validation methodology. Int J Electr Comput Eng. 2021;11:1742.

23. Gulli A, Pal S. Deep learning with Keras. Birmingham: Packt Publishing Ltd; 2017.

24. Raschka S, Mirjalili V. Python machine learning: machine learning and deep learning with Python, scikit-learn, and TensorFlow 2. Birmingham: Packt Publishing Ltd.; 2019.

25. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay E. Scikit-learn: machine learning in Python. J Mach Learn Res. 2011;12:2825–30.

26. Altmann A, Toloşi L, Sander O, Lengauer T. Permutation importance: a corrected feature importance measure. Bioinformatics. 2010;26:1340–7.

27. Xiong X, Wei Y, Lam HCY, Wong CKH, Lau SYF, Zhao S, Ran J, Li C, Jiang X, Yue Q, Cheng W, Wang H, Wang Y, Chong KC. Association between cold weather, influenza infection, and asthma exacerbation in adults in Hong Kong. Sci Total Environ. 2023;857: 159362.

28. Hales S, Lewis S, Slater T, Crane J, Pearce N. Prevalence of adult asthma symptoms in relation to climate in New Zealand. Environ Health Perspect. 1998;106:607–10.

29. Bloomer BJ, Stehr JW, Piety CA, Salawitch RJ, Dickerson RR. Observed relationships of ozone air pollution with temperature and emissions. Geophys Res Lett 2009;36.

30. Bull GM. The weather and deaths from pneumonia. Lancet. 1980;1:1405–8.

31. Graudenz GS, Landgraf RG, Jancar S, Tribess A, Fonseca SG, Faé KC, Kalil J. The role of allergic rhinitis in nasal responses to sudden temperature changes. J Allergy Clin Immunol. 2006;118:1126–32.

32. Togias AG, Naclerio RM, Proud D, Fish JE, Adkinson NF Jr, Kagey-Sobotka A, Norman PS, Lichtenstein LM. Nasal challenge with cold, dry air results in release of inflammatory mediators. Possible mast cell involvement. J Clin Invest. 1985;76:1375–81.

33. Xu Z, Huang C, Su H, Turner LR, Qiao Z, Tong S. Diurnal temperature range and childhood asthma: a time-series study. Environ Health. 2013;12:12.

34. Kim J, Lim Y, Kim H. Outdoor temperature changes and emergency department visits for asthma in Seoul, Korea: a time-series study. Environ Res. 2014;135:15–20.

35. Bronte-Moreno O, González-Barcala FJ, Muñoz-Gall X, Pueyo-Bastida A, Ramos-González J, Urrutia-Landa I. Impact of air pollution on asthma: a scoping review. Open Respir Arch. 2023;5: 100229.

36. Cohen AJ, Brauer M, Burnett R, Anderson HR, Frostad J, Estep K, Balakrishnan K, Brunekreef B, Dandona L, Dandona R, Feigin V, Freedman G, Hubbell B, Jobling A, Kan H, Knibbs L, Liu Y, Martin R, Morawska L, Pope CA 3rd, Shin H, Straif K, Shaddick G, Thomas M, van Dingenen R, van Donkelaar A, Vos T, Murray CJL, Forouzanfar MH. Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. Lancet. 2017;389:1907–18.

37. Kelly FJ. Oxidative stress: its role in air pollution and adverse health effects. Occup Environ Med. 2003;60:612–6.

38. Li XY, Gilmour PS, Donaldson K, Macnee W. In vivo and in vitro proinflammatory effects of particulate air pollution (PM10). Environ Health Perspect. 1997;105(suppl 5):1279–83.

39. Villeneuve PJ, Chen L, Rowe BH, Coates F. Outdoor air pollution and emergency department visits for asthma among children and adults: a case-crossover study in northern Alberta. Canada Environ Health. 2007;6:1–15.

40. Delfino RJ, Zeiger RS, Seltzer JM, Street DH, Matteucci RM, Anderson PR, Koutrakis P. The effect of outdoor fungal spore concentrations on daily asthma severity. Environ Health Perspect. 1997;105:622–35.

41. Ostro B, Lipsett M, Mann J, Braxton-Owens H, White M. Air pollution and exacerbation of asthma in African–American children in Los Angeles. Epidemiology. 2001;12:200–8.

42. Tecer LH, Alagha O, Karaca F, Tuncel G, Eldes N. Particulate matter (PM(2.5), PM(10–2.5), and PM(10)) and children's hospital admissions for asthma and respiratory diseases: a bidirectional case-crossover study. J Toxicol Environ Health A. 2008;71:512–20.

43. Chang YJ, Kim HY, Albacker LA, Baumgarth N, McKenzie AN, Smith DE, Dekruyff RH, Umetsu DT. Innate lymphoid cells mediate influenza-induced airway hyper-reactivity independently of adaptive immunity. Nat Immunol. 2011;12:631–8.

44. Feldman LY, Zhu J, To T. Estimating age-specific influenza-associated asthma morbidity in Ontario, Canada. Respir Med. 2019;155:104–12.

45. Gerke AK, Yang M, Tang F, Foster ED, Cavanaugh JE, Polgreen PM. Association of hospitalizations for asthma with seasonal and pandemic influenza. Respirology. 2014;19:116–21.

Publisher's Note