

## Research Article

# Error Correction of Meteorological Data Obtained with Mini-AWSs Based on Machine Learning

Ji-Hun Ha,<sup>1</sup> Yong-Hyuk Kim<sup>2</sup>, Hyo-Hyuc Im,<sup>1</sup> Na-Young Kim,<sup>1</sup> Sangjin Sim,<sup>1</sup> and Yourim Yoon<sup>1</sup>

<sup>1</sup>Korea Oceanic and Atmospheric System Technology, No. 1503, 90 Gyeongin-ro 53-gil, Guro-gu, Seoul 08215, Republic of Korea

<sup>2</sup>School of Software, Kwangwoon University, 20 Kwangwoon-ro, Nowon-gu, Seoul 01897, Republic of Korea

<sup>3</sup>Department of Computer Engineering, College of Information Technology, Gachon University, 1342 Seongnam-daero, Sujeong-gu, Seongnam-si, Gyeonggi-do 13120, Republic of Korea

Correspondence should be addressed to Yong-Hyuk Kim; [yhdfly@kw.ac.kr](mailto:yhdfly@kw.ac.kr)

Received 18 December 2017; Accepted 25 March 2018; Published 30 April 2018

Academic Editor: Herminia García Mozo

Copyright © 2018 Ji-Hun Ha et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Severe weather events occur more frequently due to climate change; therefore, accurate weather forecasts are necessary, in addition to the development of numerical weather prediction (NWP) of the past several decades. A method to improve the accuracy of weather forecasts based on NWP is the collection of more meteorological data by reducing the observation interval. However, in many areas, it is economically and locally difficult to collect observation data by installing automatic weather stations (AWSs). We developed a Mini-AWS, much smaller than AWSs, to complement the shortcomings of AWSs. The installation and maintenance costs of Mini-AWSs are lower than those of AWSs; Mini-AWSs have fewer spatial constraints with respect to the installation than AWSs. However, it is necessary to correct the data collected with Mini-AWSs because they might be affected by the external environment depending on the installation area. In this paper, we propose a novel error correction of atmospheric pressure data observed with a Mini-AWS based on machine learning. Using the proposed method, we obtained corrected atmospheric pressure data, reaching the standard of the World Meteorological Organization (WMO;  $\pm 0.1$  hPa), and confirmed the potential of corrected atmospheric pressure data as an auxiliary resource for AWSs.

## 1. Introduction

Numerical weather prediction (NWP) refers to a method of weather forecasting based on the numerical analysis of current meteorological conditions using physical and mechanical principles of atmospheric processes. Nowadays, NWP accounts for a considerable proportion of weather forecasts worldwide [1]. Since Richardson [2] claimed in 1922 that it is possible to predict weather using a numerical process, advances in the relevant fields of science and technology in the past decades have brought NWP to a level of objective and quantitative estimation. However, there is a growing need for more precise and accurate meteorological information due to the increasing frequency and magnitude of extreme weather events such as local heavy precipitation, typhoons, and droughts [3–5].

Because NWP-based weather forecasting uses observed data, its prediction accuracy can be improved by reducing the observation interval. However, installing a sufficient number of automatic weather stations (AWSs) to ensure an adequate spacing is difficult to achieve because of economic and geographical limitations such as expensive installation and maintenance costs and difficulties selecting installation sites.

Many studies have been conducted to overcome these limitations of AWSs. Straka et al. [6] carried out meteorological observations in mesonet units using meteorological instruments mounted on a vehicle; Cassano [7] observed local temperature variations using a portable weather station (Kestrel 4000) mounted on a bicycle handlebar. However, these methods are geographically limited with respect to installing meteorological observation instruments throughout a large geographic area. Spurred by the recent development of

sensor and network technology, methods for meteorological data collection using sensors mounted on a vehicle have been proposed [8, 9] and meteorological data collection using sensors embedded in smartphones has been conducted [10, 11]. Such methods have the advantage of collecting a large amount of data at a low cost; however, they require the voluntary participation of many people and the data acquisition in sparsely populated regions or locations that are hard to access is limited. Moreover, such sensors are not as reliable as the instruments specialized for meteorological observations.

A Mini-AWS is a miniature weather station developed to overcome the drawbacks of the AWS. The Mini-AWS is approximately seven times less expensive than the AWS (1,300 USD versus 9,000 USD) and requires low maintenance and repair costs. The installation site selection is hardly limited because it requires a very small space; hence, it can be installed in any place deemed suitable, ensuring the stable and steady collection of data. It can also be installed on a mobile object such as a vehicle. However, it is exposed to the external environment depending on the installation site and the collected observation data should be corrected to make them amenable for application in weather forecasting.

We propose a novel correction method based on machine learning to make atmospheric pressure data collected by Mini-AWSs amenable for use as auxiliary meteorological data. If the errors of the corrected Mini-AWS data fall within the range of the maximum permissible error ( $\pm 0.1$  hPa) recommended by the World Meteorological Organization (WMO), they are usable as auxiliary meteorological data [12, 13]. Studies on error correction methods based on machine learning techniques have been conducted in the fields of sensors and meteorology. Smith et al. [14] conducted a study on correcting data collected by sensors installed in a building using machine learning techniques and obtained good results for meteorological data. Lee et al. [15] corrected abnormal meteorological data using machine learning and obtained better results in comparison with those obtained using traditional interpolation methods. Earlier studies on the correction of atmospheric smartphone data conducted by the present research team [16, 17] also yielded results within the standard error. In this study, atmospheric pressure data were collected with Mini-AWSs, preprocessed, and error-corrected based on machine learning using the atmospheric pressure measured at the nearest AWS as the reference value.

The rest of this paper is organized as follows. In Section 2, the specifications of the Mini-AWS are presented and the data collection method is described. The theoretical background of the machine learning approaches used for the experiments is explained in Section 3. The correction method and experimental results are provided in Sections 4 and 5, respectively. Conclusions are drawn in Section 6.

## 2. Mini-AWS

**2.1. Specifications.** A Mini-AWS is a miniature weather station (dimensions: 157 mm  $\times$  167 mm  $\times$  34 mm) capable of measuring and recording air temperature, relative humidity, and atmospheric pressure (Figure 1). Its advantages over the

TABLE 1: Specifications of the Mini-AWS.

Category	Contents
Measurement element	Temperature, humidity, pressure
Accuracy	
Temperature	$\pm 0.3^\circ\text{C}$
Humidity	$\pm 2.0\%$ (typical) RH
Pressure	$\pm 1.5$ hPa
Range	
Temperature	-40 to $125^\circ\text{C}$
Humidity	0–100% RH
Pressure	10–1200 hPa
Resolution	
Temperature	0.01 $^\circ\text{C}$
Humidity	0.04% RH
Pressure	0.012 hPa
Measurement interval	Optional (5 s to hours)
Ambient temperature	-40 $^\circ\text{C}$ to 85 $^\circ\text{C}$
Duration (without power supply)	Maximum 18 hours
Data transmission	Wireless (3G mobile network) communication
Power	USB power 5 V



FIGURE 1: Mini-AWS.

AWS include low-cost installation and maintenance/repair and ease of installation in areas where the AWS installation is constrained for geographic and economic reasons. It can also be mounted on a vehicle and used as a mobile weather station because its precise position can be tracked using GPS (Global Positioning System) and GLONASS. Additionally, the power supply for its sensor and communication units can be automatically switched on and off based on the necessity to save energy. Table 1 presents the detailed specifications of the Mini-AWS.

**2.2. Data Collection.** For the data collection, we installed eight Mini-AWSs in the Pyeongchang area from January 22 to February 12, 2016. Additionally, we mounted a Mini-AWS on a vehicle and gathered data for three days while driving from Seoul to the Pyeongchang area and returning during the same period. Figure 2 shows the Mini-AWSs installed in a cross-country stadium and on a vehicle. The Mini-AWSs were configured to transmit information on the observation time (year, month, day, hour, minute, and second), location



FIGURE 2: Installation of the Mini-AWS in cross-country stadium (a) and on a mobile vehicle (b).

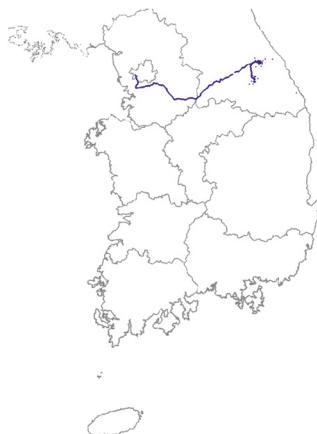


FIGURE 3: Collection sites of Mini-AWS data.

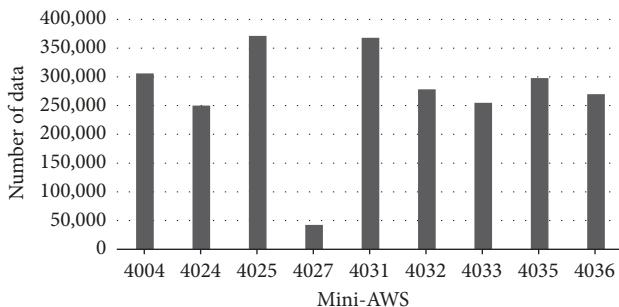


FIGURE 4: Number of collected Mini-AWS data.

(latitude, longitude, and altitude), and weather (air temperature, relative humidity, and atmospheric pressure) every five seconds. A total of 2,438,934 data points were obtained during the collection period. Figures 3 and 4 show the data collection sites and the number of data collected with each of the nine Mini-AWSs, respectively.

### **3. Machine Learning Approaches**

**3.1. Linear Regression.** Linear regression, which is one of the most widely used modeling techniques, is a method to linearly model the relationship between two variables (dependent and independent). It can be classified into simple, multiple, and multivariate linear regression (LR) depending on the number of dependent and independent variables.

Linear regression was performed to obtain the equation of the best fit that minimizes the sum of the squared errors (SSE) of the observed data. The least squares method is generally used as an approach to minimize the SSE.

**3.2. Artificial Neural Networks.** Artificial neural networks (ANNs) are a computational approach in machine learning that simulate the human brain. They are used to analyze the problem through a learning process by connecting neurons in a multilayer structure and controlling the connection strength between individual neurons. There are several types of ANNs depending on the neuron modeling and connection methods. In this paper, the most common type, that is, the multilayer perceptron (MLP) [18], was utilized. The MLP is composed of three layers (input layer, output layer, and hidden layer between them); each layer contains interconnected nodes. In an MLP, nonlinear problems can be analyzed using the hidden layer and a nonlinear activation function. Learning is implemented using a back-propagation algorithm and the gradient descent method in which an update rule is applied to minimize the difference between the target and output values.

**3.3. Support Vector Regression.** Support vector machines (SVMs) [19] are one of the machine learning techniques influenced by statistical learning theory; the maximal margin between two categories is sought. Using the structural risk minimization principle, SVMs have advantages over conventional statistical learning methods that rely on empirical risk minimization in generalization problems [20]. They can solve both classification and regression problems; regression SVMs called support vector regression (SVR) are used for regression problems. They can also solve nonlinear problems using the kernel trick with kernel-dependent performance variations.

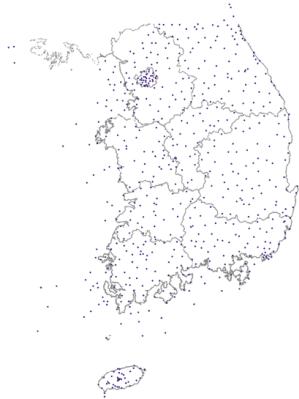


FIGURE 5: Locations of AWSs

The sequential minimal optimization (SMO) algorithm [21] is used for regression analysis in this paper. The SVM-based learning involves the solution of a quadratic programming (QP) optimization problem, whereby SMO accelerates the learning speed by breaking down a large QP problem into a series of smallest possible QP problems.

**3.4. Expectation–Maximization Clustering.** Expectation–maximization (EM) clustering [22] is an unsupervised machine learning method using the EM algorithm to cluster various pattern sets closely aligned within a space by analyzing their patterns. The EM algorithm provides an iterative method to find the parameters with the maximum likelihood in probabilistic models using latent variables in two steps. In the first step (expectation step), the expected value of a latent variable is calculated; in the second step (maximization step), the parameter associated with the value of the latent variable calculated in the expectation step is estimated. The estimated parameter undergoes the next expectation step and increasingly accurate values can be estimated by iteratively running the algorithm. The EM algorithm is widely used, especially for solving statistical estimation and mixture estimation problems [23].

#### 4. Correction Methods

**4.1. Data.** Data collected with the Mini-AWS and corresponding data from 595 AWSs provided by the Korea Meteorological Administration (KMA) were used for the experiments (Figure 5). To match the AWS data points with the Mini-AWS data points, the AWS atmospheric dataset, which consists of data points with 1-minute intervals, was converted into a dataset with data points with 5-second intervals, using the linear interpolation method.

For the correction of a collected Mini-AWS data point, the sea surface pressure measured at the nearest AWS was used as the reference value. If the corresponding value of the nearest AWS was missing, that of the second nearest AWS was used as the reference value. If the value was still missing after repeating the process twice, that data point was

excluded from the analysis in consideration of the efficacy of the experiments.

Atmospheric pressure readings should be reduced to the mean sea level to make the readings of different weather stations comparable by cancelling out altitude-dependent differences. The reduction to the mean sea level was performed on all atmospheric pressure readings based on information about the atmospheric pressure, altitude, and temperature obtained with the Mini-AWS data to provide more information for machine learning. Equation (1) was used for the reduction to the mean sea level.

$$P_0 = P \left( 1 - \frac{0.0065h}{(T + 0.0065h + 273.15)} \right)^{-5.257} \\ = 0.03414 \times \frac{Ph}{(273 + T)}. \quad (1)$$

**4.2. Preprocessing.** The AWS observation data are subject to errors due to various problems such as observation instrument errors or power and communication line disturbances; it is essential to provide accurate data by eliminating erroneous data through quality control [24]. As part of quality control, we performed missing data and location range test, physical limit test, and persistence test and eliminated outliers by applying the  $3\sigma$  rule.

The observed values are stored as missing values in case of missing observation data due to observation instrument errors or communication disturbances. A missing data point is generally coded as -999 or -99; all values coded as -99 or less were deemed missing in this study.

The AWS and Mini-AWS data points with location information deviating from the latitudinal and longitudinal extent of the Korean Peninsula were considered to be errors because the observations took place on the Korean Peninsula. The latitude/longitude position was set as  $33^{\circ}\text{N}$ – $39^{\circ}\text{N}$  and  $124^{\circ}\text{E}$ – $131^{\circ}\text{E}$  as provided by the Korean National Geographic Information Institute (2009).

The WMO standards [12] stipulate outliers as atmospheric pressures lower than 500 hPa and higher than 1,080 hPa, air temperatures lower than  $-60^{\circ}\text{C}$  and higher than  $80^{\circ}\text{C}$ , and relative humidity values lower than 0% and higher than 100%. In this study, all Mini-AWS and AWS data points corresponding to these values were removed.

The persistence test is performed to detect observations that remain unchanged for a certain period of time due to instrument errors or other disturbances. The WMO recommends cases in which changes do not occur for 60 minutes beyond the threshold values of  $0.1^{\circ}\text{C}$  for air temperature, 0.1 hPa for atmospheric pressure, and 1% for relative humidity as “suspect” cases failing the persistence test. We performed the persistence test and removed all “suspect” cases.

In consideration of the cases in which the observations are influenced by surrounding conditions other than meteorological conditions, we removed all Mini-AWS and AWS observation data deviating from the  $3\sigma$  limits.

**4.3. Verification.** Various statistical analysis techniques can be used for performance verification; the most important

TABLE 2: Experimental results of the data divided into Mini-AWSs.

Mini-AWS	Number of data	MAE		RMSE	
		LR	MLP	LR	MLP
4004	297,825	0.518	0.252	0.628	0.325
4024	227,665	0.491	0.363	0.640	0.459
4025	321,815	0.390	0.310	0.523	0.399
4027	8,783	0.256	0.101	0.345	0.130
4031	208,423	0.385	0.264	0.496	0.331
4032	128,439	0.329	0.180	0.438	0.230
4033	252,109	0.38	0.258	0.496	0.339
4035	242,173	0.985	0.485	1.199	0.623
4036	253,600	0.395	0.239	0.524	0.311

quantitative criterion is accuracy. In this study, performance verification was conducted by comparing the mean absolute error (MAE) and root-mean-square error (RMSE). The RMSE is used as the standard statistical metric when testing model performances in meteorological, atmospheric, and climatological studies; the MAE is also a widely used model evaluation parameter [25]. The parameter  $\epsilon$  represents the error between the values predicted by the model and observed for the  $i$ th sample among a total of  $n$  samples. The RMSE and MAE can be calculated using (2) and (3), respectively.

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n \epsilon_i^2} \quad (2)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |\epsilon_i|. \quad (3)$$

The MAE is generally smaller than the RMSE for the same result because the same weight is applied to all errors with respect to the MAE, whereas heavier penalties are given to larger errors in RMSE.

All experiments were cross-validated. Cross-validation is an experimental method for the evaluation of the performance of a supervised learning model. In an  $N$ -fold cross-validation, the dataset is partitioned into  $N$  folds, whereby  $N-1$  folds make up the training set and the remaining one makes up the test set, repeating the experiment  $N$  times. Each of the  $N$  folds is used once for testing. In this study, 10-fold cross-validation was conducted. Cross-validation prevents overfitting to obtain generalized results.

## 5. Experimental Results

**5.1. Benchmarks.** The LR and MLP were carried out using the experimental data and Weka software [26], accepting the default parameter values of the Weka software as the options for each method. A total of 1,940,903 experimental data points, which were extracted from 2,438,934 raw data points through preprocessing, were trained separately for each Mini-AWS using 10-fold cross-validation. The experimental results of the tested model of the data of each Mini-AWS are

TABLE 3: Weighted mean values of the experimental results for the data divided into Mini-AWSs.

ML methods	Weighted mean MAE	Weighted mean RMSE	Average time taken to build the model
LR	0.489	0.737	0.62 s
MLP	0.300	0.389	180.44 s

TABLE 4: Experimental results for 1,500 samples.

ML methods	Weighted mean MAE	Weighted mean RMSE	Average time taken to build the model
LR	0.062	0.079	0.06 s
MLP	0.041	0.052	1.41 s
SMOreg	0.057	0.077	1.51 s

given in Table 2. Table 3 presents the weighted mean values of the experimental results of all Mini-AWSs.

**5.2. Results.** The SMOreg (regression by SMO) is known to have superior performance with respect to generalization problems. However, the training time increases in proportion to the number of data points. Figure 6 shows the training time for the SMOreg implementation of the experimental data; the number of sampled data increases from 1,000 to 9,000 by 1,000. As the number of data used for training increased, the time taken to build the model increased on a logarithmic scale.

In an attempt to reduce the SMOreg training time, experiments were conducted using the last 1,500 and 5,000 data points extracted from each Mini-AWS. The experimental results are listed in Tables 4 and 5, respectively.

The results in Tables 4 and 5 show that the learning was over in less than one minute and MAE weighted average value satisfied the range of the WMO stipulated maximum permissible error (0.1 hPa). The experiment with 1,500 samples yields better results than that with 5,000 samples, so it can be seen that the shorter the learning cycle, the smaller the error. If more data are collected and validated in the future, the results may be further improved.

TABLE 5: Experimental results for 5,000 samples.

ML methods	Weighted mean MAE	Weighted mean RMSE	Average time taken to build the model
LR	0.087	0.111	0.17 s
MLP	0.056	0.071	4.39 s
SMOreg	0.083	0.110	32.50 s

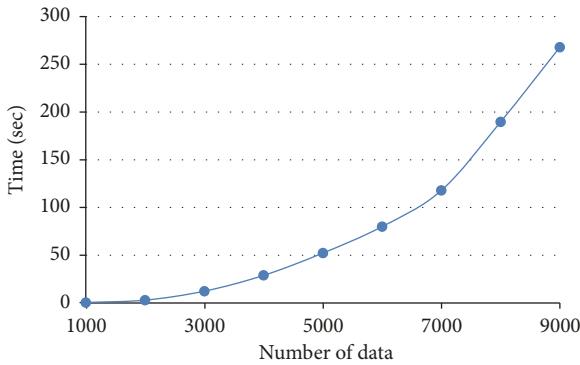


FIGURE 6: Time taken to build the model according to the number of data in the SMOreg.

Considering the different characteristics of the data collected from each Mini-AWS, we created models by Mini-AWSs. Figure 7 shows the experimental results of the individual Mini-AWSs. Overall, the results were better than those in Table 4, but the performance of some models such as 4033 was worse. Of the three machine learning techniques, only MLP could correct data within the range of the WMO stipulated maximum permissible error.

To improve the results mentioned in Figure 7, we devised a method to add categorical information to datasets using a clustering method. Through EM clustering, we classified data samples into groups and added the information to the collected data. We could confirm that, with this additional information, the model performance was significantly improved as given in Table 6, which presents the number of clusters yielded from each Mini-AWS and the MAE and RMSE values.

The Mini-AWSs with the best MAE or RMSE values are highlighted in bold. All Mini-AWSs, except for 4025 and 4036, show superior MLP results. The LR and SMOreg results of 4033 and 4035 are greater than 0.1.

In the experiments mentioned in Figure 7, we used the default parameters of Weka. Since changing parameter values affects the results, we also made experiments to change the parameter values of MLP and SMOreg. The MLP node options were changed as follows: the sum of the number of features and number of classes was halved and the number of hidden layers was reduced to two (option *a,a*). The SMOreg option was changed to the Pearson Universal Kernel (option *puk*). The experimental results for the changed options are displayed in Figure 8.

The experimental results reveal that SMOreg (option *puk*) outperformed MLP (option *a,a*) with respect to the error of all Mini-AWSs, demonstrating MAE and RMSE values lower than 0.05 and 0.06, respectively. The weighted mean MAE and RMSE of SMOreg (option *puk*) are 0.020 and 0.034, respectively, approximately 1.5-fold lower than those of MLP (option *a,a*), which are 0.039 and 0.052.

## 6. Conclusion

We present a study aimed at correcting data collected with Mini-AWSs using three different machine learning approaches and atmospheric pressure readings of the nearest AWSs as reference values. The weighted means of the experimental machine learning data divided into Mini-AWSs did not reach WMO standards. In the case of SMOreg, the time taken to build the model increased on a logarithmic scale with increasing number of training data. However, the correction results of the SMOreg implementation with the last 1,500 observation data points sampled with each Mini-AWS, which was conducted additionally in an attempt to reduce the training time, fall within the range of the standard permissible error set by the WMO when EM clustering and SMOreg (option *puk*) are applied. The error correction performance of machine learning varies slightly depending on the applied approach, ultimately yielding superior results in the order of SMOreg, MLP, and LR.

Experiments of sampling and clustering were conducted for performance improvement. We used a sampling method that extracts the most recently collected data, and using the sampled data led to better results compared to doing the whole ones. The smaller the number of samples, the lower the correction error, so we concluded that short learning cycles help to reduce correction errors. But since the data used for verification were collected during a short period of time, it is necessary to collect more data in order to correctly verify the model. In addition, we need to decide how often we will produce a correction model, in the case that we apply the collection and correction of real-time data. Clustering was used to provide additional categorical information on datasets, and we could confirm that performance was significantly improved through the clustering process. However, it was not easy to know which characteristics the data classified by clustering have and how they contribute to performance improvement.

We confirm the feasibility of the error correction method presented in this paper to render Mini-AWS atmospheric pressure data usable as observation data for weather forecasting. However, additional validation is necessary, given the limited data collection period and amount of mobile data. In a follow-up study, additional validation of the Mini-AWS data will be performed taking seasonal and geographical variations into account to test methods and compare errors by applying various preprocessing methods, such as internal consistency tests, in addition to the preprocessing method used in this study. Such studies are expected to continuously enhance the usability of Mini-AWSs and thus contribute to improving the accuracy of numerical weather prediction such as [27, 28].

TABLE 6: Experimental results for clustered Mini-AWS data.

Mini-AWS	Number of clusters	MAE			RMSE		
		LR	MLP	SMOreg	LR	MLP	SMOreg
4004	7	0.066	<b>0.042</b>	0.066	0.082	<b>0.055</b>	0.083
4024	6	0.037	<b>0.030</b>	0.037	0.046	<b>0.038</b>	0.047
4025	12	0.036	0.036	<b>0.035</b>	<b>0.047</b>	0.047	0.049
4027	4	0.032	<b>0.024</b>	0.032	0.041	<b>0.030</b>	0.041
4031	4	0.038	<b>0.029</b>	0.038	0.049	<b>0.038</b>	0.049
4032	7	0.044	<b>0.035</b>	0.044	0.058	<b>0.047</b>	0.059
4033	2	0.103	<b>0.064</b>	0.102	0.127	<b>0.081</b>	0.132
4035	5	0.087	<b>0.031</b>	0.079	0.111	<b>0.043</b>	0.126
4036	13	0.043	0.043	<b>0.042</b>	<b>0.055</b>	0.057	0.055

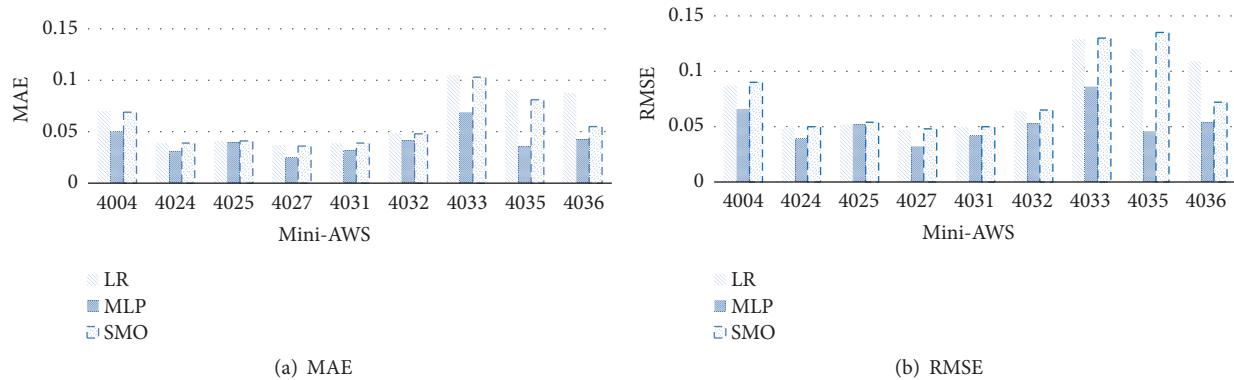


FIGURE 7: Experimental results for the Mini-AWSs and 1,500 samples.

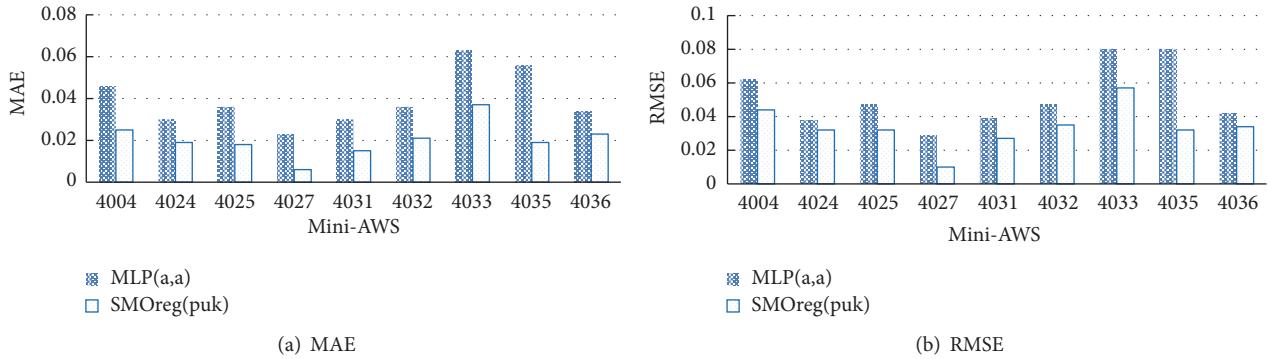


FIGURE 8: Experimental results for MLP (option a,a) and SMOreg (option puk).

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## Acknowledgments

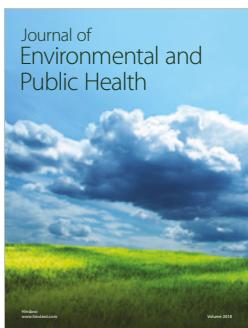
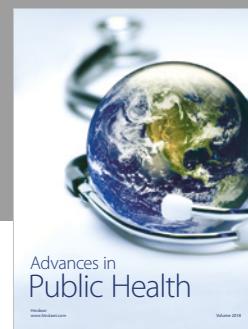
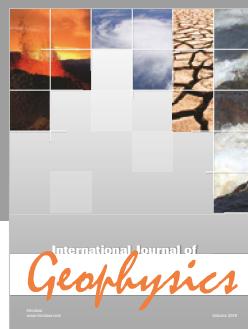
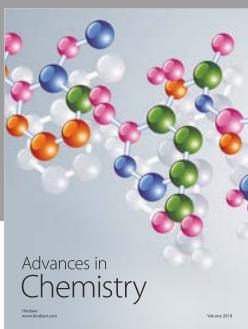
This work was supported by the “Meteorological Industry Support and Utilization Technology Development Program” (no. KMIPA2015-4060), which is one of the Meteorological See-At Technology Development Programs through the Korea Meteorological Industry Promotion Agency (KMIPA) in 2015. This research was also supported by the Information

and Communication Promotion Fund through the Ministry of Science and ICT (MSIP) of Korea (C0510-18-1007) and may differ from the official opinion of the MSIP.

## References

- [1] R. Kimura, “Numerical weather prediction,” *Journal of Wind Engineering & Industrial Aerodynamics*, vol. 90, no. 12-15, pp. 1403–1414, 2002.
- [2] L. F. Richardson, “Weather prediction by numerical methods,” Cambridge: Cambridge University Press, 1922.
- [3] C. D. Allen, A. K. Macalady, H. Chenchouni et al., “A global overview of drought and heat-induced tree mortality reveals

- emerging climate change risks for forests," *Forest Ecology and Management*, vol. 259, no. 4, pp. 660–684, 2010.
- [4] P. Mondal and A. J. Tatem, "Uncertainties in measuring populations potentially impacted by sea level rise and coastal flooding," *PLoS ONE*, vol. 7, no. 10, Article ID e48191, 2012.
- [5] S.-H. Chiang and K.-T. Chang, "The potential impact of climate change on typhoon-triggered landslides in Taiwan, 2010–2099," *Geomorphology*, vol. 133, no. 3–4, pp. 143–151, 2011.
- [6] J. M. Straka, E. N. Rasmussen, and S. E. Fredrickson, "A mobile mesonet for finescale meteorological observations," *Journal of Atmospheric and Oceanic Technology*, vol. 13, no. 5, pp. 921–936, 1996.
- [7] J. J. Cassano, "Weather bike: A bicycle-based weather station for observing local temperature variations," *Bulletin of the American Meteorological Society*, vol. 95, no. 2, pp. 205–209, 2014.
- [8] B. Mahoney, S. Drobot, P. Pisano, B. McKeever, and J. O'Sullivan, "Vehicles as mobile weather observation systems," *Bulletin of the American Meteorological Society*, vol. 91, no. 9, pp. 1179–1182, 2010.
- [9] W. P. Mahoney III and J. M. O'Sullivan, "Realizing the potential of vehicle-based observations," *Bulletin of the American Meteorological Society*, vol. 94, no. 7, pp. 1007–1018, 2013.
- [10] A. Overeem, J. C. R. Robinson, H. Leijnse, G. J. Steeneveld, B. K. P. Horn, and R. Uijlenhoet, "Crowdsourcing urban air temperatures from smartphone battery temperatures," *Geophysical Research Letters*, vol. 40, no. 15, pp. 4081–4085, 2013.
- [11] S. Sosko and S. Dalyot, "Crowdsourcing user-generated mobile sensor weather data for densifying static geosensor networks," *ISPRS International Journal of Geo-Information*, vol. 6, no. 3, article no. 61, 2017.
- [12] WMO-No.8, Guide to meteorological instruments and methods of observation, World Meteorological Organization, Geneva, Switzerland, 2014.
- [13] J. Estévez, P. Gavilán, and J. V. Giráldez, "Guidelines on validation procedures for meteorological data from automatic weather stations," *Journal of Hydrology*, vol. 402, no. 1–2, pp. 144–154, 2011.
- [14] M. K. Smith, C. C. Castello, and J. R. New, "Machine learning techniques applied to sensor data correction in building technologies," in *Proceedings of the 2013 12th International Conference on Machine Learning and Applications, ICMLA 2013*, pp. 305–308, USA, December 2013.
- [15] M.-K. Lee, S.-H. Moon, Y.-H. Kim, and B.-R. Moon, "Correcting abnormalities in meteorological data by machine learning," in *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics (SMC '14)*, pp. 888–893, San Diego, Calif., USA, October 2014.
- [16] N.-Y. Kim, Y.-H. Kim, Y. Yoon, H.-H. Im, R. K. Y. Choi, and Y. H. Lee, "Correcting air-pressure data collected by MEMS sensors in smartphones," *Journal of Sensors*, vol. 2015, Article ID 245498, 10 pages, 2015.
- [17] Y.-H. Kim, J.-H. Ha, Y. Yoon et al., "Improved Correction of Atmospheric Pressure Data Obtained by Smartphones through Machine Learning," *Computational Intelligence and Neuroscience*, vol. 2016, Article ID 9467878, 2016.
- [18] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning internal representations by error propagation," *California Univ. San Diego La Jolla Inst. for Cognitive Science*, 1985.
- [19] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.
- [20] V. N. Vapnik, "An overview of statistical learning theory," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 10, no. 5, pp. 988–999, 1999.
- [21] J. Platt, "Sequential minimal optimization: a fast algorithm for training support vector machines," Microsoft Research, 1998.
- [22] T. K. Moon, "The expectation-maximization algorithm," *IEEE Signal Processing Magazine*, vol. 13, no. 6, pp. 47–60, 1996.
- [23] S. Borman, "The expectation maximization algorithm-a short tutorial," *Submitted for publication*, pp. 1–9, 2004. [https://www.cs.utah.edu/~piyush/teaching/EM\\_algorithm.pdf](https://www.cs.utah.edu/~piyush/teaching/EM_algorithm.pdf).
- [24] I. Zahumenský, *Guidelines on quality control procedures for data from automatic weather stations*, World Meteorological Organization, Switzerland, Geneva, 2004.
- [25] T. Chai and R. R. Draxler, "Root mean square error (RMSE) or mean absolute error (MAE)?—arguments against avoiding RMSE in the literature," *Geoscientific Model Development*, vol. 7, no. 3, pp. 1247–1250, 2014.
- [26] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The WEKA data mining software: an update," *ACM SIGKDD Explorations Newsletter*, vol. 11, no. 1, pp. 10–18, 2009.
- [27] J.-H. Seo, Y. H. Lee, and Y.-H. Kim, "Feature selection for very short-term heavy rainfall prediction using evolutionary computation," *Advances in Meteorology*, vol. 2014, Article ID 203545, 15 pages, 2014.
- [28] Y.-H. Kim and Y. Yoon, "Spatiotemporal pattern networks of heavy rain among automatic weather stations and very-short-term heavy-rain prediction," *Advances in Meteorology*, vol. 2016, Article ID 4063632, 2016.



Hindawi  
Submit your manuscripts at  
[www.hindawi.com](http://www.hindawi.com)

