

## Research Article

# Estimation of Maximum Daily Fresh Snow Accumulation Using an Artificial Neural Network Model

Gun Lee,<sup>1</sup> Dongkyun Kim ,<sup>2</sup> Hyun-Han Kwon ,<sup>3</sup> and Eunsoo Choi<sup>4</sup>

<sup>1</sup>Graduate Research Assistant, Department of Civil Engineering, Hongik University, Seoul, Republic of Korea

<sup>2</sup>Associate Professor, Department of Civil Engineering, Hongik University, Seoul, Republic of Korea

<sup>3</sup>Professor, Department of Civil Engineering, Chonbuk National University, Jeonju-si, Jeollabuk-do, Republic of Korea

<sup>4</sup>Professor, Department of Civil Engineering, Hongik University, Seoul, Republic of Korea

Correspondence should be addressed to Dongkyun Kim; [kim.dongkyun@hongik.ac.kr](mailto:kim.dongkyun@hongik.ac.kr)

Received 18 September 2018; Revised 7 November 2018; Accepted 12 December 2018; Published 8 January 2019

Academic Editor: Federico Porcù

Copyright © 2019 Gun Lee 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.

For estimation of maximum daily fresh snow accumulation (MDFSA), a novel model based on an artificial neural network (ANN) was proposed. Daily precipitation, mean temperature, and minimum temperature were used as the input data for the ANN model. The ANN model was regularized and trained using a set of 19,923 data points, observed daily in South Korea between 1960 and 2016. Leave-one-out cross validation was performed to validate the model. When the input data were known at the gauged locations, the correlation coefficient between the observed MDFSA and the estimated one by the ANN model was 0.90. When the input data were spatially interpolated at ungauged locations using the ordinary kriging (OK) method, the correlation coefficient was 0.40. The difference in correlation coefficients between the two methods implies that, while the ANN model itself has good performance, a significant portion of the uncertainty of the estimated MDFSA at ungauged locations comes from high spatial variability of the input variables that cannot be captured by the network of in situ gauges. However, these correlation coefficients were significantly greater than the correlation coefficient obtained by spatially interpolating the MDFSA values with the OK method ( $R = 0.20$ ). These findings suggest that our ANN model significantly reduces the uncertainty of the estimated MDFSA caused by its high spatial variability.

## 1. Introduction

On February 17, 2014, the roof of the Mauna Ocean Resort Gym in Gyeongju, Korea (Figure 1), collapsed due to the heavy weight of accumulated snow, resulting in 113 casualties. Although the primary cause of the accident was defective construction [1], the rooftop snow accumulation (Figure 1(b)), which exceeded 50 cm after the accident at the site, played a significant role for the collapse. The operational snow warning system of Korea (<http://www.kma.go.kr/weather/warning/status.jsp>) was unable to monitor this heavy snow accumulation [2]. Only 16 cm/day of snow accumulation was observed at the gauge located 14 km from the accident site (Figure 1). This postsurvey of the accident suggests that the snow accumulation cannot be accurately estimated through simple spatial interpolation, primarily because the factors influencing snow accumulation (such as

winter precipitation and temperature) show significant spatiotemporal variability depending on meteorological [3] and geographical [4, 5] conditions.

In practice, snow accumulation is often approximated using the ten-to-one rule, but the rule can be easily broken due to the high variability of snow density and the shape of ice crystals [6]. The actual ratio between the snow accumulation and the precipitation varies between 2 : 1 and 45 : 1 [7], depending on how the shapes of the snowflakes and ice crystals change as the temperature increases [8]. In addition, the wind can move the accumulated snow from one location to another [9–11]. The ideal approach to estimate the snow accumulation involves analyzing the physical mechanisms of snowmelt and interaction with the wind [8, 12, 13]. However, this approach requires information on variables that are difficult to observe, such as temporal variation in temperature and the internal structure of the

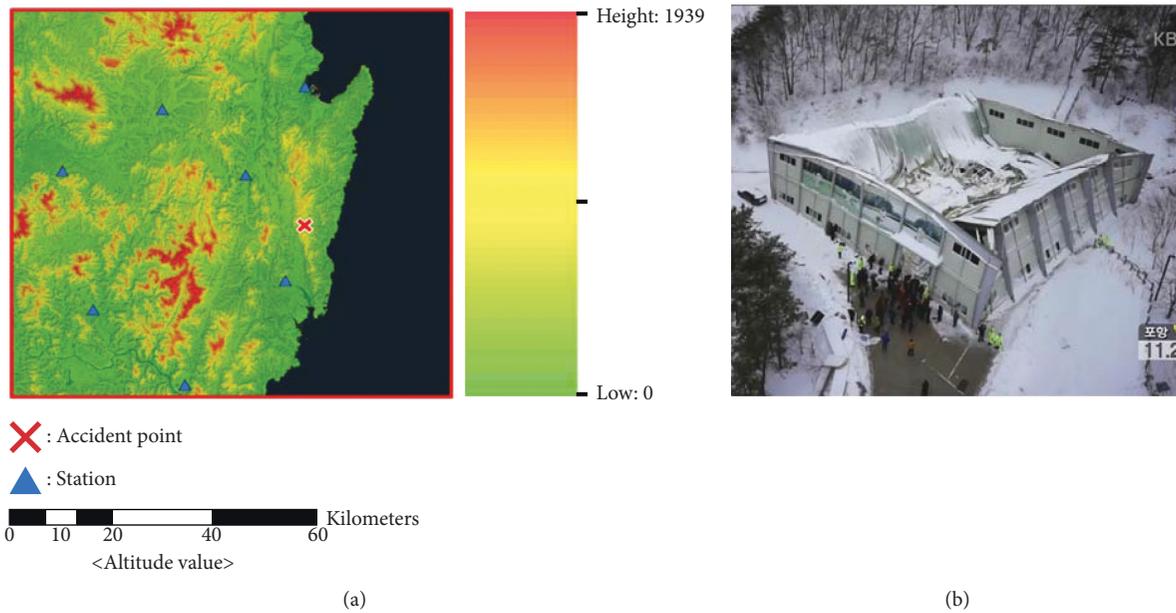


FIGURE 1: (a) Locations of in situ snow accumulation measurement stations and the snow collapse disaster location (Mauna Ocean Resort Gym in Gyeongju). (b) A picture of the disaster site (source: <http://news.kbs.co.kr/news/view.do?ncd=2834382>).

accumulated snow [4], hindering the operational use of this approach [14]. Alternately, methods based on the statistical relationships between various factors affecting the snow accumulation have been studied [15–17]. However, the relationships between the snow accumulation and factors such as precipitation, temperature, and vapor pressure cannot be fully explained with simple statistical models. Furthermore, these factors are observed at point locations on the ground and show significant spatiotemporal variability, which adds to the uncertainty of the snow accumulation estimations at ungauged locations [18, 19].

Methods based on satellite imageries can reduce the uncertainties derived from spatial variability of the factors influencing the snow accumulation [20–25]. Because satellite imagery does not directly measure the snow accumulation, these approaches extract the required information, such as brightness temperature, to estimate the snow accumulation from the satellite imageries and use this information as the input for the models estimating the snow accumulation. Recently, artificial intelligence techniques have been applied to establish the relationships between the snow accumulation and various factors. Artificial intelligence methods have been especially useful in this regard because they do not require a complete understanding of the complex physical processes of phenomena [26–34]. Davis et al. [35] used an artificial neural network (ANN) algorithm to establish the relationships of five different satellite image brightness temperatures to average snowflake diameter, density, temperature, and depth of the accumulated snow. Gan et al. [36] used an ANN algorithm to estimate the snow water equivalent in the Red River basin of North Dakota and Minnesota in the United States, obtaining a correlation coefficient of 0.71. Gharaei-Manesh et al. [37] estimated the snow accumulation in semiarid regions in Iran using the M5 Decision Tree algorithm, a type of an artificial

intelligence algorithm. Due to the difficulty in obtaining the observational data directly related to the snow accumulation, these authors used geographical characteristics, such as channel network and stream power, as the inputs for their ANN algorithms. Tedesco et al. [38] used an ANN algorithm to establish the relationship between special sensor microwave imager (SSM/I) data and the snow water equivalent. Sun et al. [39] used an ANN to establish the relationships between the parameters of the snow accumulation process and the various brightness temperature values derived from the SSM/I. Dobrev and Klein [40] obtained the spatial distribution of the snow accumulation of a complex mountainous terrain from the moderate resolution imaging spectrometer (MODIS) imagery by analyzing these data with an ANN algorithm. Liang et al. [41] determined the satellite-based microwave brightness temperature and visible/infrared reflectance values using a support vector machine (SVM) algorithm. Czyzowska-Wisniewski et al. [42] used an ANN algorithm to analyze the IKONOS satellite images to estimate the snow accumulation in the Alps region. Park et al. [43] developed an ANN algorithm to determine the occurrence of snow events based on the precipitation and the temperature in Korea. Kim et al. [44] developed an ANN algorithm to estimate the snow accumulation based on the precipitation and the temperature; this ANN algorithm was subsequently used to predict the snow accumulation for a future period under climate change.

This study aims to develop a novel method of estimating the depth of snow accumulation at ungauged locations to help prevent future disasters. While approaches based on the satellite-ANN combination have been used to estimate the snow accumulation, the revisit period of satellites is several days at the shortest. Thus, the application of satellite-based methods is limited to estimating the gradual variation in snow accumulation over a long period in cold regions.

In addition, no study has addressed the applicability of the real-time snow accumulation detection in ungauged locations based on the ANN in the Korean peninsula, where the factors influencing snow accumulation have great spatial variability. We developed an ANN-based model that estimates the snow accumulation based on the in situ temperature and precipitation. Then, we used the developed model to estimate the snow accumulation at ungauged locations based on the spatially interpolated temperature and ground precipitation data. We compared our method to a method that spatially interpolates the in situ snow accumulation with the ordinary kriging (OK) technique. The results of the comparison indicate that our ANN-based approach outperforms the aforementioned approach based on the OK method.

## 2. Methodology

**2.1. Study Area.** South Korea, comprising southern half of the Korean Peninsula and Jeju Island, was chosen as the study area (Figure 2). Snowfall in the study area occurs mostly between late November and early March. During this period, the average precipitation is approximately 88.5 mm, about 6.77 percent of the annual precipitation (1307.7 mm). The factors such as precipitation, temperature, and wind may influence the snow accumulation. These factors show high spatial variability in the study area, primarily due to drastic changes in elevation (0 m–1950 m) over a small area (100,210 km<sup>2</sup>).

**2.2. Data Description.** The Korean Meteorological Administration (KMA) operates 94 in situ gauges in the study area as of 2017, while 123 in situ gauges have been historically operated since April 1904. The KMA records these meteorological variables following a strict standard specified by the government regulation [45]. Precipitations are measured using tipping bucket gauges of which precision is 0.5 mm~1 mm. Temperatures are measured using metal-type gauges of which precision is 0.1°C. At each gauge, the maximum accumulated depth of fresh snow occurring between 0 o'clock and 24 o'clock has been measured on a daily basis, and this measured value is defined as the maximum daily fresh snow accumulation (MDFSA). The MDFSA values are read daily from rulers installed on the ground, and it is these MDFSAs that our study aims to estimate. At each gauge location, the maximum ( $T_{\max}$ ), average ( $T_{\text{mean}}$ ), and minimum daily temperature ( $T_{\min}$ ) and daily precipitation ( $P$ ) values were measured along with MDFSA.

This study used 19,923 datasets of MDFSA,  $P$ ,  $T_{\max}$ ,  $T_{\text{mean}}$ , and  $T_{\min}$  values observed at 90 gauges between 1960 and 2016 to train and validate the ANN model. Figure 2(b) shows the number of the gauges of which data this study used. It showed significant annual variation because the snowfall did not occur at all gauge locations. While the precipitation and the temperature were recorded all possible gauges, we excluded the dataset that does not have any MDFSA recording.

**2.3. Artificial Neural Network.** Figure 3 shows the structure of the artificial neural network used in this study. The network is composed of the input layer, hidden layer, and output layer. This structure mimics the information-transferring process in the human brain, where neurons accept stimulations from the dendrites and then transmit them through axons to other neurons. Just as neurons are interconnected by synapses, the nodes of the artificial neural networks are connected by weights. Each node in a layer receives values from multiple nodes in previous layers. Then, the node calculates the weighted average of the received values. Lastly, the node transforms this value using a given activation function, where the transformed value is subsequently transferred to each of the nodes in the next layer. Here, the activation function works as a threshold of the human brain neurons, based on which the transfer of the stimulation is determined.

The ANN can be also mathematically characterized as follows. Let  $X$  represent the matrix composed of  $n$  number of input variables of the ANN, or  $x_1, x_2, x_3, \dots, x_n$ , including the bias term, and let  $W_i^k$  represent the matrix composed of the strength of the connection (or simply the weight factor) between the  $j$ th input variables and the  $i$ th node value in the  $k$ th hidden layer, or  $w_{ij}^k$ . They can be represented as follows:

$$X = [x_1, x_2, x_3, \dots, x_n, 1], \quad (1)$$

$$W_i^k = [w_{i1}^k, w_{i2}^k, w_{i3}^k, \dots, w_{in}^k, b_i^k]. \quad (2)$$

The net value assigned to the  $i$ th node in the first hidden layer or  $y_i^1$  is calculated as the linear combination of the input variables as follows:

$$y_i^1 = x_1 \cdot w_{i1}^1 + x_2 \cdot w_{i2}^1 + x_3 \cdot w_{i3}^1 + \dots + x_n \cdot w_{in}^1 = XW_i^{1T}. \quad (3)$$

Then, this net value is converted into the input value of the nodes in the next hidden layer based on a given activation function. This study adopted the sigmoid function. This process can be mathematically described as follows:

$$z_i = \frac{1}{1 + e^{-y_i}}, \quad (4)$$

where  $z_i$  represents the converted value at the  $i$ th node. This converted value is to be used as the input value of the next hidden layer;  $z_i$  becomes the new  $x_i$ , and the procedure described in equations (1) through (4) is repeated over the hidden layers until the final output variable is obtained. The process of determining the weight factors, or  $w_{ij}^k$ s, of the ANN is called the learning process. In this learning process, weight factors are determined such that the difference between the observed variable and the estimation from the ANN model is minimized. The Levenberg–Marquardt backpropagation optimization algorithm [46] was employed to train the ANN of this study. The epoch represents the minimum number of repetitions of the training required for the ANN model to reach a given performance standard. Lower the epoch, more stable is the performance of the ANN model. Mitchell [47] provides more detail explanation on the ANN.

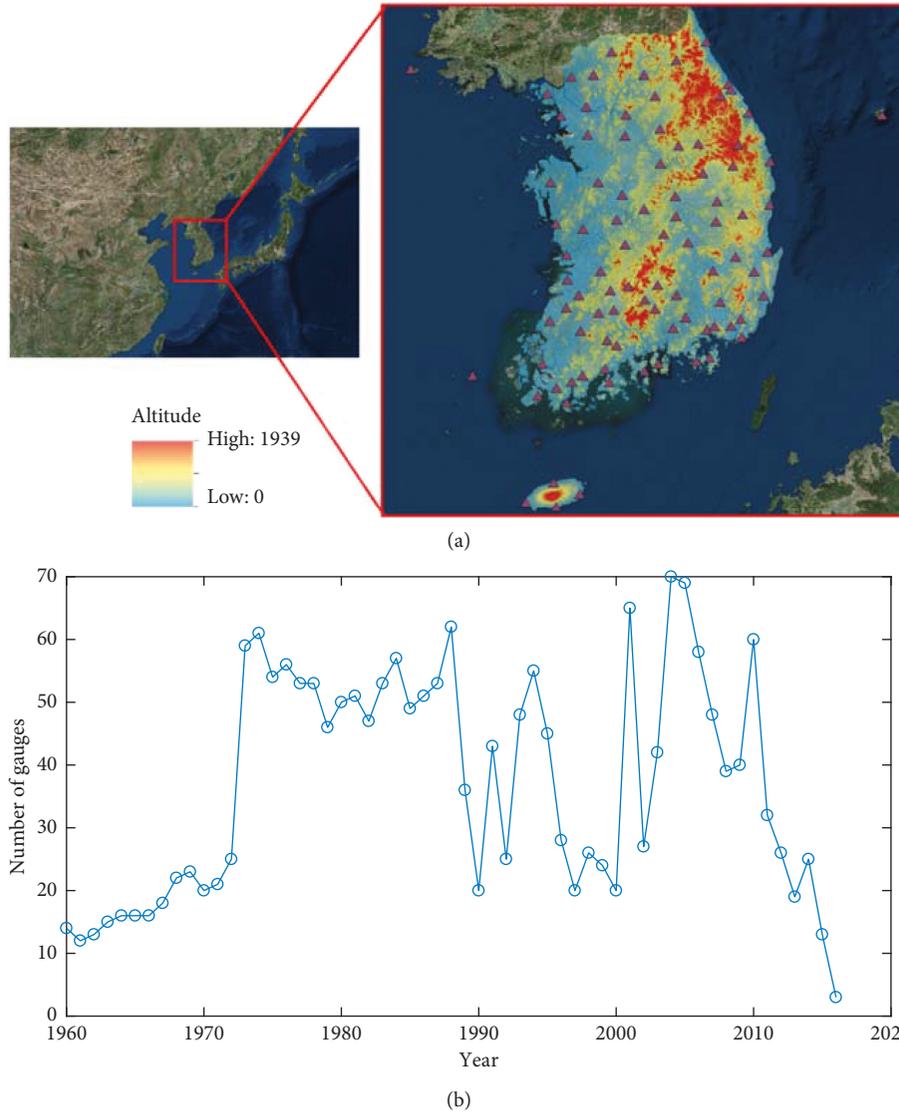


FIGURE 2: (a) Study area (South Korea) and the locations of the 94 ground gauges; (b) number of gauges used for the analysis varying with year.

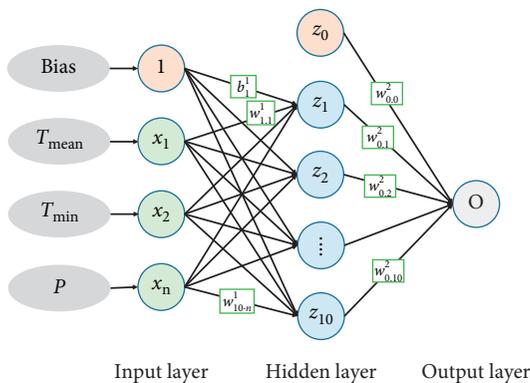


FIGURE 3: Structure of the optimal artificial neural network model determined by this study.

2.3.1. *Regularization of the Artificial Neural Network.* The process of determining the overall structure of the ANN is called regularization. These structure characteristics

include the number and type of input variables, number of hidden layers, number of nodes in each layer, and type of activation function. A complex ANN structure forces the model behavior to be strongly dependent on the observed data [48]. Conversely, a simple ANN structure forces the model to miss the overall relationships between the input variables and output. There is no analytical or formal approach that can be generally applied to the regularization. Instead, the methods based on simple trial and error are used.

2.3.2. *Input Data and Number of Hidden Layers of the ANN Model.* For operational use of the methodology, the input data should be easy to obtain and show high correlations with the MDFSA. After a brief correlation analysis and literature review [4, 49], we chose the precipitation ( $P$ ),  $T_{\text{max}}$ ,  $T_{\text{min}}$ , and  $T_{\text{mean}}$  as the candidate input variables of the ANN model.  $P$  has a very high correlation with MDFSA ( $R = 0.65$ ), so it was considered as an input variable by default, leaving

8 ( $=2^3$ ) additional choices of whether or not to consider the remaining 3 variables. We used the following approach to determine the optimal combination of input variables:

- (1) Develop an ANN composed of one hidden layer with five nodes.
- (2) Train the ANN with a given dataset (out of 8 dataset combinations, which are  $[P]$ ,  $[P, T_{\min}]$ ,  $[P, T_{\text{mean}}]$ ,  $[P, T_{\max}]$ ,  $[P, T_{\text{mean}}, T_{\min}]$ ,  $[P, T_{\max}, T_{\min}]$ ,  $[P, T_{\max}, T_{\text{mean}}]$ , and  $[P, T_{\max}, T_{\text{mean}}, T_{\min}]$ ), excluding the data of one gauge location.
- (3) Estimate the MDFSA based on the ANN using the input variables observed at the excluded gauge location developed in Step 2.
- (4) Compare the estimated MDFSA in Step 3 with the observed MDFSA at the same gauge location in Step 2.
- (5) Repeat the process between Steps 2 and 4 for all gauge locations to acquire the relationship between the observed and the estimated MDFSA. Calculate the correlation coefficient between the two variables.
- (6) If the correlation coefficient calculated in Step 5 is lower than 0.9, add one more node in the hidden layer and repeat Steps 2 through 5.
- (7) If the correlation coefficient estimated in Step 5 does not exceed 0.9 after adding 15 additional nodes (total 20 nodes), reduce the original correlation coefficient threshold value by 0.01 and repeat Steps 2 through 6.
- (8) Repeat Steps 1 through 7 for 20 times and record the average correlation coefficient value for each of the 8 input variable combinations.

Figure 4 shows the boxplot of the mean of the correlation coefficients obtained from repeating Steps 2 through 8 for 50 times with each of the 8 combinations of input variables. Note that the 50 times of repetition here was performed using the ANN model with the optimal structure that was decided from the processes described by Step 1 through Step 8. Therefore, the 50 times of repetition is different from the 20 times of repetition mentioned in Step 7. The case in which  $P$ ,  $T_{\text{mean}}$ , and  $T_{\min}$  were used as the input variables showed the highest correlation coefficients (0.87-0.88). It is notable that as the mean correlation coefficient increased, its variability decreased. This suggests that the choice of proper input variables increased not only the accuracy but also the precision of the ANN model. A high correlation coefficient (0.87-0.88) was obtained for the tested ANN model with one hidden layer. Furthermore, the test using the model structure with multiple numbers of hidden layers did not yield any improved performance. Therefore, this study used one hidden layer for the final ANN model.

*2.3.3. Optimal Number of Nodes in the Hidden Layer.* Blue and red lines in Figure 5 show variations in the correlation coefficients and the epochs in relation to the number of nodes in the hidden layer of the ANN model. These values were extracted using the process described in the previous

section 2.3.1. Higher the correlation coefficient, better was the performance of the ANN model. The correlation coefficient generally increased with an increase in the number of nodes, up to 10 nodes, and then decreased gradually. This showed that the ANN model started to experience overfitting as the number of nodes exceeded the threshold of 10. The epoch monotonically decreased but showed little difference after 10 nodes. On the basis of these results, we used the ANN model with 10 nodes in the hidden layer to maintain the accuracy and the precision.

*2.3.4. Number of Training Data Points.* The performance of the ANN generally improved as the size of the dataset used for training increased. However, the rate of performance improvement significantly decreased over a given threshold of training data size. This is because the ANN model became overfitted with an increase in the amount of training data [50]. Figure 6 shows how the size of the training dataset was related to the performance of the ANN model. The performance of the ANN model drastically increased when the size of the training dataset exceeded 20, and it stabilized when the training dataset exceeded 1,000. For this reason, we trained the ANN model using 1,000 datasets randomly chosen from the entire database composed of 19,923 in situ daily observations of  $P$ ,  $T_{\min}$ , and  $T_{\text{mean}}$ . However, note that there is no absolute guideline to determine the optimal size of the training dataset. It should consider not only the relative proportion of the training dataset but also the occurrences of the instability points (points displaying a sudden drop of correlation coefficient even with the increase of the training dataset, circled in Figure 6). Figure 6 suggests that the 1,000 datasets represent a good compromise between the number of training data and the occurrence of instable points.

### 3. Results

*3.1. Performance of the ANN Model Alone.* After the ANN model was regularized, we validated its performance using the leave-one-out cross validation approach. This approach involves training the ANN model on the dataset but excluding the data from just one gauge. Then, the ANN estimates the MDFSA based on the input data observed at the excluded gauge. Finally, the estimated MDFSA for the excluded gauge is compared with the observed one. This approach allows measurement of the pure performance of the ANN without any disturbance due to spatial variability of the input variables.

Figure 7(a) compares the observed MDFSA ( $x$ ) and the MDFSA estimated by the ANN ( $y$ ). Figure 7(b) compares the observed MDFSA ( $x$ ) and the MDFSA estimated by the OK method. The correlation coefficient was 0.90 for the ANN and 0.20 for the OK. Thus, if the precipitation along with minimum and average temperatures of the day is known, the proposed ANN model generally gives more accurate estimate of the MDFSA than that from the spatial interpolation. The high spatial variability of the geographical and meteorological factors influencing the MDFSA is known to cause

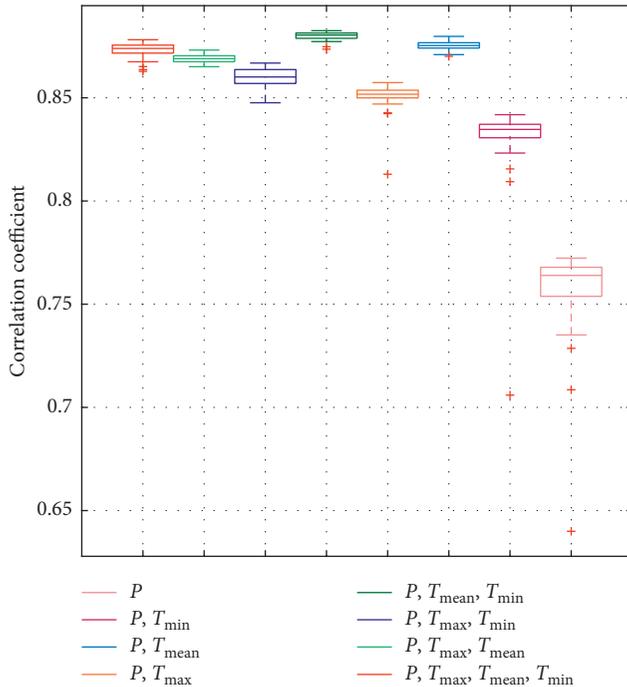


FIGURE 4: Boxplot showing the relationship between input data and correlation coefficient.

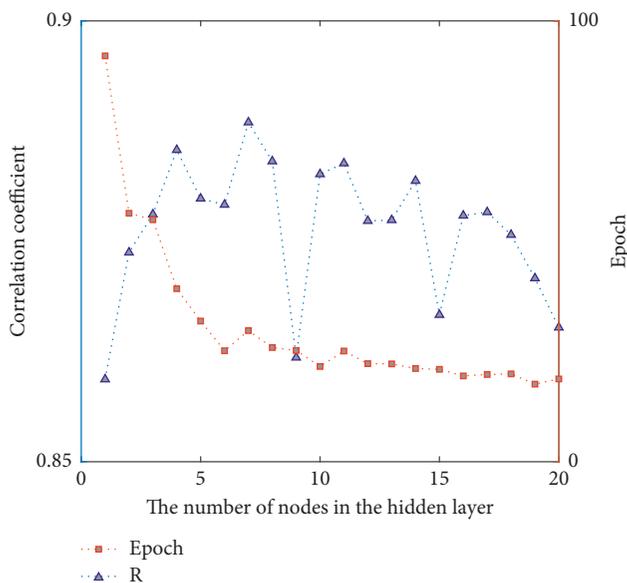


FIGURE 5: Relationship between the number of nodes in the hidden layer and the threshold epoch over which the snow depth residual converged.

high spatial variability of the MDFSA estimation [51], which is the primary reason why spatial interpolation using the OK method did not perform as well as the ANN model.

It is important for the ANN model to estimate the MDFSA in the range that can lead to a disaster. Figure 8(a) shows the performance of the ANN at different intervals of the MDFSA. The performance of the ANN was best between 0 cm and 10 cm, with a correlation coefficient of

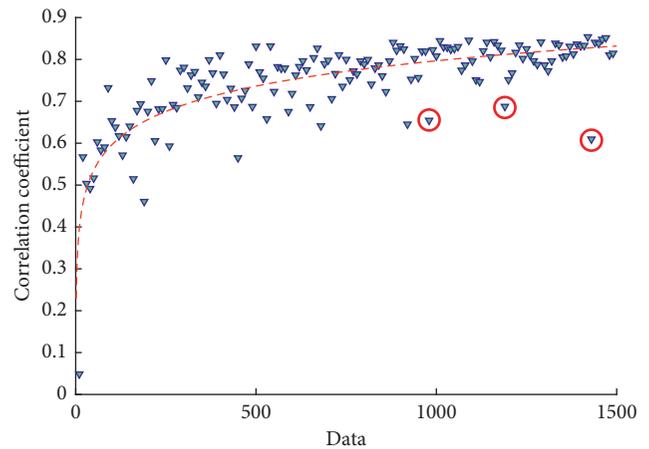


FIGURE 6: Relationship between number of data points for training and correlation coefficient.

0.73. The correlation coefficient varied between 0.26 and 0.44 at the remaining four MDFSA intervals. This abrupt decrease in correlation coefficient was partly attributed to the number of data points available for training of the ANN model. While  $\sim 93\%$  of the dataset was concentrated at the first depth interval (0 cm–10 cm), the proportion of the dataset that was available for the training at the remaining ranges was 5.5, 1.2, 0.46, and 0.19 percent for the second (10 cm–20 cm) through the last (40 cm–50 cm) depth intervals, respectively. The correlation coefficient for the greatest depth interval (40 cm–50 cm) was 0.44 for the ANN and 0.17 for the OK (See Figure 8(b)). This indicates that the ANN method can provide a reasonably accurate estimate of disastrous snow accumulation, which the OK method would fail to predict in most cases.

### 3.2. Performance of the ANN Model in Ungauged Locations.

The ultimate goal of this study was to estimate the MDFSA at ungauged locations. To quantify the performance of the model for this purpose, we estimated the MDFSA using the ANN model based on spatially interpolated input variables. Here, all input variables were interpolated using the OK technique. Figures 9(a) through 9(c) show the observed ( $x$ ) versus spatially interpolated estimates of the variables used as the input of the ANN model. Leave-one-out cross validation was performed to estimate the variables on the  $y$ -axis. The plot corresponding to the precipitation (Figure 9(c)) was shown in the log-log axis because most precipitation is concentrated near the value of zero, but the precipitation that may cause disaster is far greater than zero.

The correlation coefficient was 0.69 for  $T_{\min}$ , 0.70 for  $T_{\text{mean}}$ , and 0.59 for  $P$ . The spatial correlation of precipitation is low because most heavy snow events in the study area occurs in the form of orographic precipitation that is caused by orographic lift of the moist air that the cold Siberian High takes from the warm sea beneath it [52]. The high spatial variability of elevation in Korea causes high spatial variability in orographic lift and winter precipitation, which the current in situ precipitation gauge network could not capture well. Comparison of the correlation coefficients for

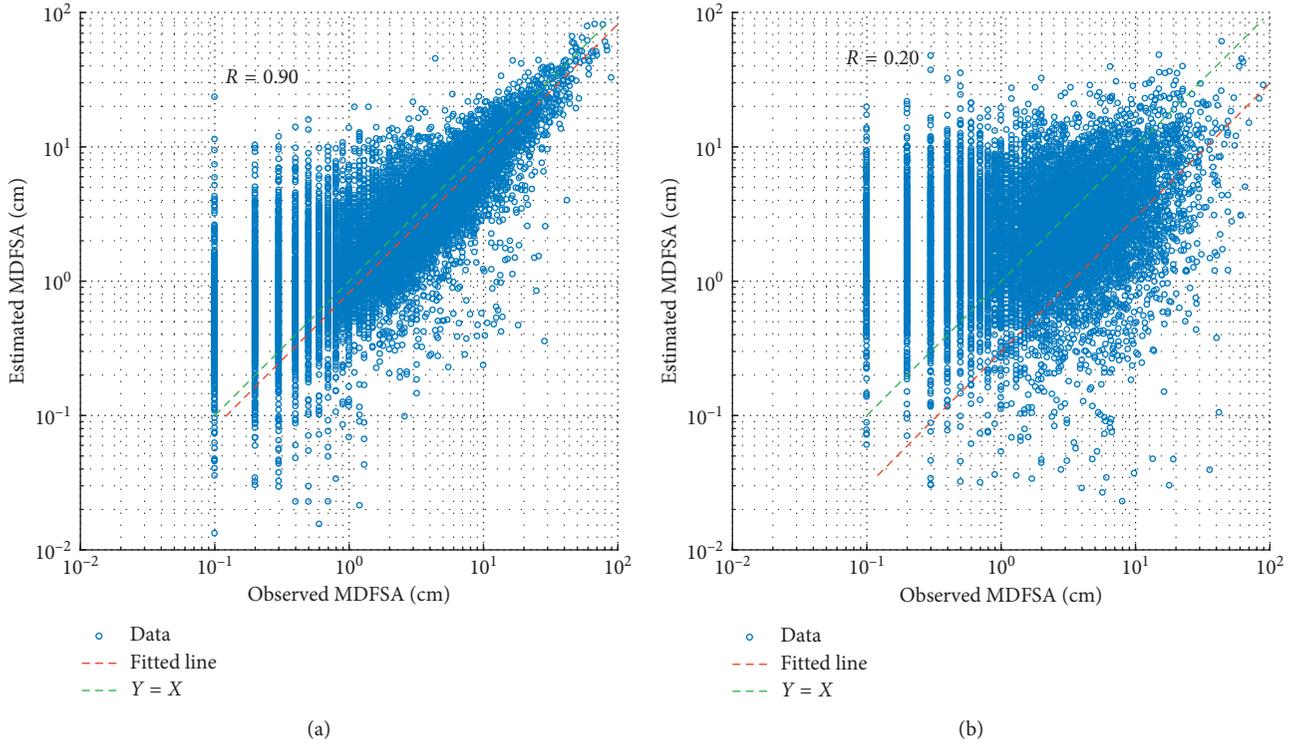


FIGURE 7: (a) Observed MDFSA ( $x$ ) versus the MDFSA estimated by the ANN model ( $y$ ). Here, the input variables of the ANN are the observed ones, not spatially interpolated; (b) observed MDFSA ( $x$ ) versus the MDFSA estimated by spatially interpolating the MDFSA values observed at the nearby gauges with the ordinary kriging method.

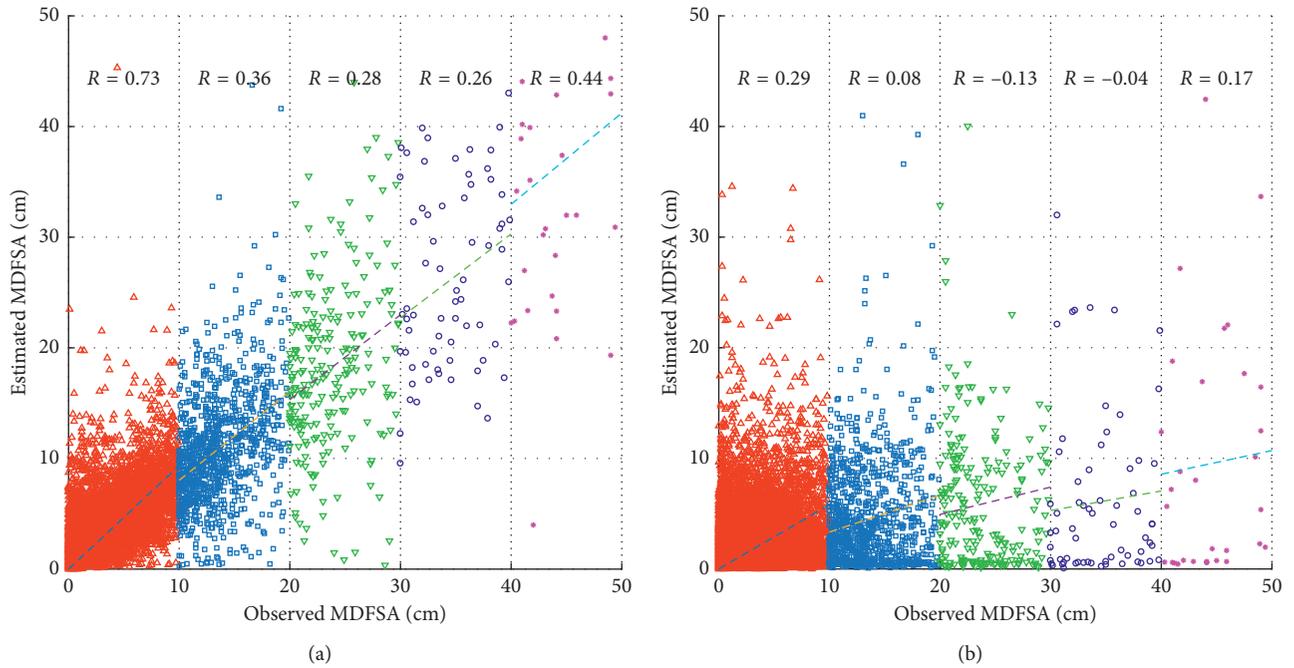


FIGURE 8: Cross validation of the ANN for different snow depth intervals. The results corresponding to the ANN and the ordinary kriging (OK) are shown. (a) ANN. (b) Ordinary kriging.

the three variables implies that uncertainty in the estimated MDFSA at ungauged locations is primarily caused by high spatial variability of winter precipitation in Korea.

Figure 10(a) compares the observed MDFSA to that estimated by the ANN model. The correlation coefficient was significantly higher for the ANN model ( $R = 0.4$ ) than the

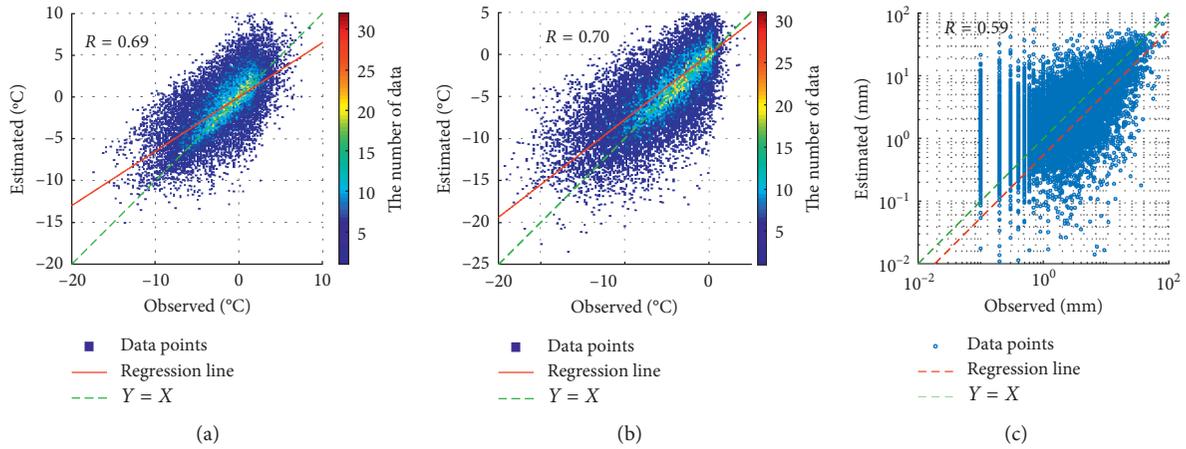


FIGURE 9: Relationship between spatially interpolated input data of the ANN model and the observed data. (a) Minimum temperature. (b) Average temperature. (c) Precipitation.

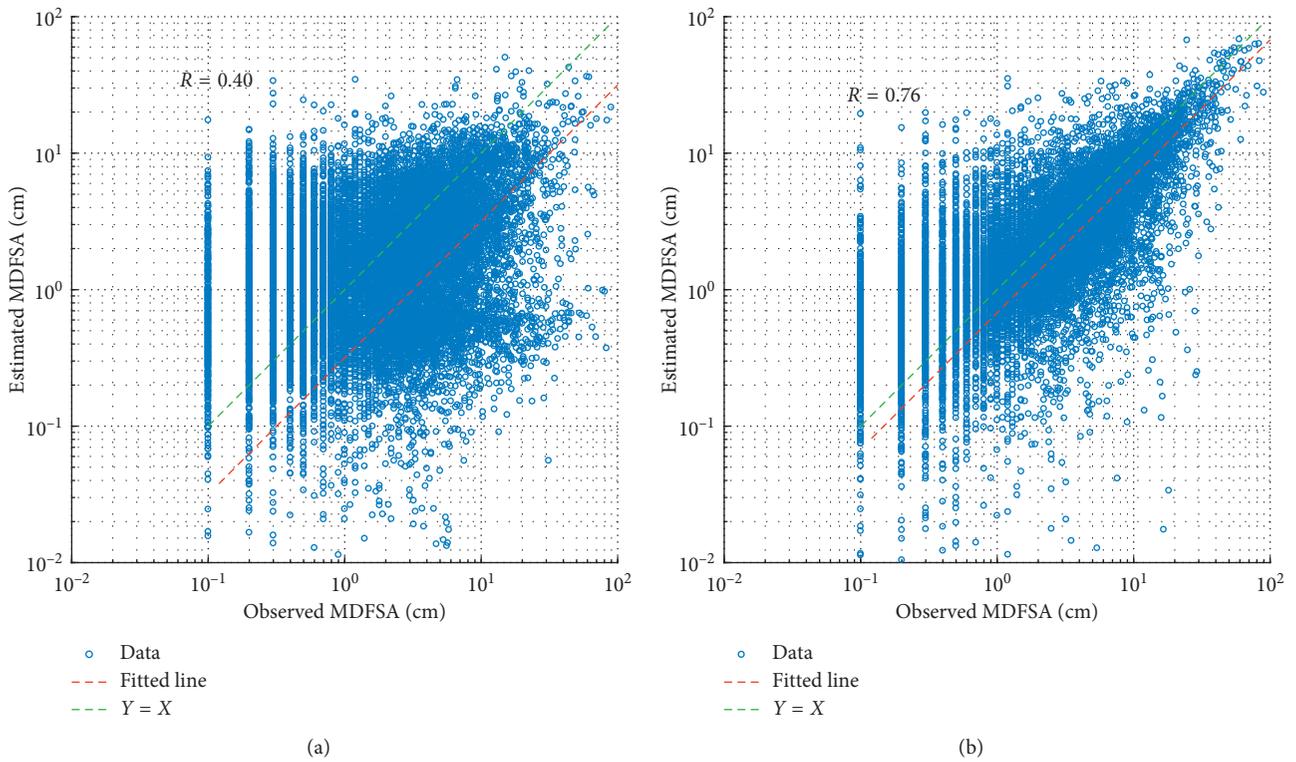


FIGURE 10: Observed versus estimated MDFSA when (a) all input variables were spatially interpolated and (b) P was observed.  $T_{\text{mean}}$  and  $T_{\text{min}}$  were spatially interpolated.

OK method ( $R = 0.2$ , See Figure 7(b)), which implies that the ANN model can estimate MDFSA with greater accuracy than the OK method for ungauged locations. However, the correlation coefficient using the interpolated variables as the input dataset ( $R = 0.4$ ) was lower than the case in which the input variables were observed values at the in situ gauges (Figure 7(a),  $R = 0.90$ ). This implies that a significant amount of the uncertainty in estimating the MDFSA at ungauged locations with the ANN model came from the spatial variability of the input variables that could not be captured by

the in situ gauge network. As shown in Figure 9(c), the spatial variability of precipitation was especially important. Figure 10(b) compares the observed ( $x$ ) versus estimated MDFSA ( $y$ ) using the in situ precipitation values as the input of the ANN model instead of the spatially interpolated values. The remaining input variables ( $T_{\text{mean}}$  and  $T_{\text{min}}$ ) were spatially interpolated, as in the case of Figure 10(a). Therefore, the difference between Figures 10(a) and 10(b) show the isolated impact of spatial variability in precipitation on the performance of the ANN. The correlation

coefficient significantly increased from 0.40 to 0.76, which implies that the uncertainty corresponding to the 47 percent ( $= 0.76 - 0.40/0.76$ ) was added by the spatial variability of precipitation to the correlation coefficient. A similar analysis was performed for the remaining input variables ( $T_{\text{mean}}$  and  $T_{\text{min}}$ ), and the isolated adverse impact of each variable on the correlation coefficient was minimal ( $T_{\text{mean}}$ , 0.01, 1.4 percent;  $T_{\text{min}}$ , 0.04, 5.2 percent).

#### 4. Summary and Conclusion

This study developed an ANN model that estimates the maximum daily fresh snow accumulation (MDFSA) based on daily precipitation, mean temperature, and minimum temperature. Regularization and training of the ANN model were performed through a trial-and-error method based on a set of 19,923 in situ data points observed at 90 gauges in the Korean Peninsula between 1960 and 2016. The final ANN model comprising one hidden layer with 10 nodes was proposed.

Definite relationships between the MDFSA and these three factors were established by the ANN model. The correlation coefficient between the observed and estimated MDFSA was 0.90. The accuracy of the ANN model was greatest in the MDFSA interval between 0 cm and 10 cm, with a correlation coefficient of 0.70. For the remaining MDFSA intervals, the correlation coefficient varied between 0.20 and 0.40. The reduction in the correlation coefficient at higher MDFSA intervals was most likely explained by the difference in amount of data available at those intervals for training of the ANN model.

The developed ANN model was used to estimate the MDFSA at ungauged locations. The correlation coefficient between the observed and estimated MDFSA was 0.40. The spatial variability of the precipitation that could not be captured by the in situ gauge network played a significant role in reducing the correlation coefficient. The isolated adverse impact of each input variable on the correlation coefficient was 47 percent for precipitation, 1.4 percent for  $T_{\text{mean}}$ , and 5.2 percent for  $T_{\text{min}}$ .

A key finding of this study is that the MDFSA at ungauged locations can be estimated with high accuracy with the help of artificial intelligence techniques if precipitation and temperature can be estimated accurately. The MDFSA estimated using the artificial intelligence methods will be more accurate than the one estimated through a more direct manner using spatial interpolation, such as the ordinary kriging method. The accuracy of the model was especially sensitive to the accuracy of the input precipitation data. Therefore, accurate estimation of the precipitation at ungauged locations is crucial to successful utilization of the proposed ANN model in practice. In this context, more in situ rain gauges should be installed in mountainous areas where snow accumulation frequently leads to disasters and precipitation is difficult to predict. This is especially because mountainous areas can have very distinct precipitation characteristics that cannot be easily inferred from the information acquired at nearby locations [53, 54].

#### Data Availability

All meteorological data used in this study were provided by the Korean Meteorological Administration. Most of them can be downloaded for free from the following website: <https://data.kma.go.kr/cmmn/main.do>.

#### Conflicts of Interest

The authors declare that they have no conflicts of interest.

#### Acknowledgments

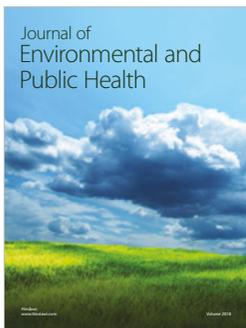
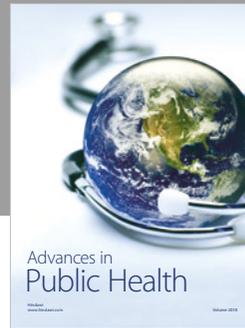
This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (project no. 2015-041523).

#### References

- [1] Y. K. Kim and H. G. Sohn, "Disasters from 1948 to 2015 in Korea and power-law distribution," in *Disaster Risk Management in the Republic of Korea*, pp. 77–97, Springer, Singapore, 2018.
- [2] Y. Lee, C. Lee, and S. Ahn, "Estimation of freshly fallen snow unit weight and maximum probable snow load," *Journal of Korean Society of Hazard Mitigation*, vol. 15, no. 1, pp. 47–55, 2015.
- [3] E. Kuusisto, *Snow Accumulation and Snowmelt in Finland*, Water Research Institute, National Board of Waters, Helsinki, Finland, 1984.
- [4] R. L. Bras, *Hydrology: An Introduction to Hydrologic Science*, Addison Wesley Publishing Company, Boston, MA, USA, 1990.
- [5] M. k. Woo, R. Heron, P. Marsh, and P. Steer, "Comparison of weather station snowfall with winter snow accumulation in High Arctic basins," *Atmosphere-Ocean*, vol. 21, no. 3, pp. 312–325, 2010.
- [6] A. Judson and N. Doesken, "Density of freshly fallen snow in the central Rocky Mountains," *Bulletin of the American Meteorological Society*, vol. 81, no. 7, pp. 1577–1587, 2000.
- [7] Y. Kim, N. Kang, S. Kim, and H. Kim, "Evaluation for snowfall depth forecasting using neural network and multiple regression models," *Journal of Korean Society of Hazard Mitigation*, vol. 13, no. 2, pp. 269–280, 2013.
- [8] S. L. Dingman, *Physical Hydrology*, Prentice-Hall, Vol. 173–178, Prentice-Hall, Upper Saddle River, NJ, USA, 2002.
- [9] B. E. Goodison, "Accuracy of Canadian snow gage measurements," *Journal of Applied Meteorology*, vol. 17, no. 10, pp. 1542–1548, 1978.
- [10] T. Hongo, M. Tsuchiya, T. Tomabechei, and H. Ueda, "Wind effects on snow drifts around two-dimensional fence," in *Proceedings of the Fourth International Conference on Snow Engineering. Snow Engineering: recent Advances and Developments*, Norwegian University of Science and Technology, Trondheim, Norway, Trondheim, Norway, June 2000.
- [11] M. Tsuchiya, T. Hongo, T. Tomabechei, and H. Ueda, *Wind Effects on Snow Accumulation on Two-Level Flat Roofs*, *Snow Engineering*, pp. 323–328, Balkema, Leiden, Netherlands, 2000.
- [12] E. Brun, P. David, M. Sudul, and G. Brunot, "A numerical model to simulate snow-cover stratigraphy for operational

- avalanche forecasting,” *Journal of Glaciology*, vol. 38, no. 128, pp. 13–22, 2017.
- [13] J. Kondo and T. Yamazaki, “A prediction model for snowmelt, snow surface temperature and freezing depth using a heat balance method,” *Journal of Applied Meteorology*, vol. 29, no. 5, pp. 375–384, 1990.
- [14] O. Eisen, M. Frezzotti, C. Genthon et al., “Ground-based measurements of spatial and temporal variability of snow accumulation in East Antarctica,” *Reviews of Geophysics*, vol. 46, no. 2, 2008.
- [15] J. J. Lee, Y. H. Jung, and S. W. Lee, “A study on the evaluation of probable snowfall depth in Korea,” *Journal of Korean Society of Hazard Mitigation*, vol. 7, no. 2, pp. 53–63, 2007.
- [16] V. Meløysund, B. Leira, K. V. Høiseth, and K. R. Lisø, “Predicting snow density using meteorological data,” *Meteorological Applications*, vol. 14, no. 4, pp. 413–423, 2007.
- [17] N. Ohara, M. L. Kavvas, and Z. Q. Chen, “Stochastic upscaling for snow accumulation and melt processes with PDF approach,” *Journal of Hydrologic Engineering*, vol. 13, no. 12, pp. 1103–1118, 2008.
- [18] C. Daly, “Guidelines for assessing the suitability of spatial climate data sets,” *International Journal of Climatology*, vol. 26, no. 6, pp. 707–721, 2006.
- [19] M. New, M. Hulme, and P. Jones, “Representing twentieth-century space-time climate variability. Part I: development of a 1961–90 mean monthly terrestrial climatology,” *Journal of Climate*, vol. 12, no. 3, pp. 829–856, 1999.
- [20] Y. Cao, X. Yang, and X. Zhu, “Retrieval snow depth by artificial neural network methodology from integrated AMSR-E and in-situ data—A case study in Qinghai-Tibet Plateau,” *Chinese Geographical Science*, vol. 18, no. 4, pp. 356–360, 2008.
- [21] A. T. C. Chang, J. L. Foster, and D. K. Hall, “Nimbus-7 SMMR derived global snow cover parameters,” *Annals of glaciology*, vol. 9, pp. 39–44, 2016.
- [22] E. G. Josberger and N. M. Mognard, “A passive microwave snow depth algorithm with a proxy for snow metamorphism,” *Hydrological Processes*, vol. 16, no. 8, pp. 1557–1568, 2002.
- [23] K. F. Kunzi, S. Patil, and H. Rott, “Snow-Cover parameters retrieved from nimbus-7 scanning multichannel microwave radiometer (SMMR) data,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. GE-20, no. 4, pp. 452–467, 1982.
- [24] J. T. Pulliainen, J. Grandell, and M. T. Hallikainen, “HUT snow emission model and its applicability to snow water equivalent retrieval,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, no. 3, pp. 1378–1390, 1999.
- [25] V. Roy, K. Goita, A. Royer, A. E. Walker, and B. E. Goodison, “Snow water equivalent retrieval in a Canadian boreal environment from microwave measurements using the HUT snow emission model,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 9, pp. 1850–1859, 2004.
- [26] L. K. Abidoye and D. B. Das, “Artificial neural network modeling of scale-dependent dynamic capillary pressure effects in two-phase flow in porous media,” *Journal of Hydroinformatics*, vol. 17, no. 3, pp. 446–461, 2015.
- [27] E. Arnone, A. Francipane, L. V. Noto, A. Scarbaci, and G. La Loggia, “Strategies investigation in using artificial neural network for landslide susceptibility mapping: application to a Sicilian catchment,” *Journal of Hydroinformatics*, vol. 16, no. 2, pp. 502–515, 2014.
- [28] J. Bahrami, M. R. Kavianpour, M. S. Abdi, A. Telvari, K. Abbaspour, and B. Rouzkhosh, “A comparison between artificial neural network method and nonlinear regression method to estimate the missing hydrometric data,” *Journal of Hydroinformatics*, vol. 13, no. 2, pp. 245–254, 2010.
- [29] N. S. Hanspal, B. A. Allison, L. Deka, and D. B. Das, “Artificial neural network (ANN) modeling of dynamic effects on two-phase flow in homogenous porous media,” *Journal of Hydroinformatics*, vol. 15, no. 2, pp. 540–554, 2013.
- [30] K. S. Kasiviswanathan and K. P. Sudheer, “Quantification of the predictive uncertainty of artificial neural network based river flow forecast models,” *Stochastic environmental research and risk assessment*, vol. 27, no. 1, pp. 137–146, 2012.
- [31] S. E. Kim and I. W. Seo, “Artificial neural network ensemble modeling with exploratory factor analysis for streamflow forecasting,” *Journal of Hydroinformatics*, vol. 17, no. 4, pp. 614–639, 2015.
- [32] P. Singh and B. Borah, “Indian summer monsoon rainfall prediction using artificial neural network,” *Stochastic Environmental Research and Risk Assessment*, vol. 27, no. 7, pp. 1585–1599, 2013.
- [33] J.-d. Wu, N. Li, H.-j. Yang, and C.-h. Li, “Risk evaluation of heavy snow disasters using BP artificial neural network: the case of Xilingol in Inner Mongolia,” *Stochastic Environmental Research and Risk Assessment*, vol. 22, no. 6, pp. 719–725, 2007.
- [34] W. Zhang, G. Liu, and H. Dai, “Simulation of food intake dynamics of holometabolous insect using functional link artificial neural network,” *Stochastic Environmental Research and Risk Assessment*, vol. 22, no. 1, pp. 123–133, 2007.
- [35] D. T. Davis, Z. Chen, L. Tsang, J.-N. Hwang, and A. T. C. Chang, “Retrieval of snow parameters by iterative inversion of a neural network,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 31, no. 4, pp. 842–852, 1993.
- [36] T. Y. Gan, O. Kalinga, and P. Singh, “Comparison of snow water equivalent retrieved from SSM/I passive microwave data using artificial neural network, projection pursuit and nonlinear regressions,” *Remote Sensing of Environment*, vol. 113, no. 5, pp. 919–927, 2009.
- [37] S. Gharaei-Manesh, A. Fathzadeh, and R. Taghizadeh-Mehrjardi, “Comparison of artificial neural network and decision tree models in estimating spatial distribution of snow depth in a semi-arid region of Iran,” *Cold Regions Science and Technology*, vol. 122, pp. 26–35, 2016.
- [38] M. Tedesco, J. Pulliainen, M. Takala, M. Hallikainen, and P. Pampaloni, “Artificial neural network-based techniques for the retrieval of SWE and snow depth from SSM/I data,” *Remote sensing of Environment*, vol. 90, no. 1, pp. 76–85, 2004.
- [39] C. Sun, H. D. Cheng, J. J. McDonnell, and C. M. Neale, “Identification of mountain snow cover using SSM/I and artificial neural network,” in *International Conference on Acoustics, Speech, and Signal Processing, ICASSP-95*, pp. 3451–3454, vol. 5IEEE, Detroit, MI, USA, May 1995.
- [40] I. D. Dobreva and A. G. Klein, “Fractional snow cover mapping through artificial neural network analysis of MODIS surface reflectance,” *Remote Sensing of Environment*, vol. 115, no. 12, pp. 3355–3366, 2011.
- [41] J. Liang, X. Liu, K. Huang et al., “Improved snow depth retrieval by integrating microwave brightness temperature and visible/infrared reflectance,” *Remote Sensing of Environment*, vol. 156, pp. 500–509, 2015.
- [42] E. H. Czyzowska-Wisniewski, W. J. D. van Leeuwen, K. K. Hirschboeck, S. E. Marsh, and W. T. Wisniewski, “Fractional snow cover estimation in complex alpine-forested environments using an artificial neural network,” *Remote Sensing of Environment*, vol. 156, pp. 403–417, 2015.
- [43] H.-S. Park, S. Jeong, and G. Chung, “Frequency analysis of future fresh snow days and maximum fresh snow depth using artificial neural network under climate change scenarios,”

- Journal of Korean Society of Hazard Mitigation*, vol. 14, no. 6, pp. 365–377, 2014.
- [44] Y. Kim, S. Kim, N. Kang, T. Kim, and H. Kim, “Estimation of frequency based snowfall depth considering climate change using neural network,” *Journal of Korean Society of Hazard Mitigation*, vol. 14, no. 1, pp. 93–107, 2014.
- [45] KMA, “Revision of standard of automatic weather observation gages,” 2018, [http://www.kma.go.kr/notify/information/law\\_notice\\_list.jsp?bid=lawnotice&mode=view&num=85&page=1&field=&text=](http://www.kma.go.kr/notify/information/law_notice_list.jsp?bid=lawnotice&mode=view&num=85&page=1&field=&text=).
- [46] K. Levenberg, “A method for the solution of certain non-linear problems in least squares,” *Quarterly of Applied Mathematics*, vol. 2, no. 2, pp. 164–168, 1944.
- [47] T. M. Mitchell, “Artificial neural networks,” *Machine learning*, vol. 45, pp. 81–127, 1997.
- [48] Oh, *Pattern Recognition*, Kyobo, Seoul, South Korea, 2008.
- [49] K. F. Dewey, “Daily maximum and minimum temperature forecasts and the influence of snow cover,” *Monthly Weather Review*, vol. 105, no. 12, pp. 1594–1597, 1977.
- [50] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, *Data Mining: Practical Machine Learning Tools and Techniques*, Morgan Kaufmann, Burlington, MA, USA, 2016.
- [51] P. D. Broxton, N. Dawson, and X. Zeng, “Linking snowfall and snow accumulation to generate spatial maps of SWE and snow depth,” *Earth and Space Science*, vol. 3, no. 6, pp. 246–256, 2016.
- [52] G. Y. Choi and J. S. Kim, “Surface synoptic climatic patterns for heavy snowfall events in the Republic of Korea,” *Journal of the Korean Geographical Society*, vol. 45, 2010.
- [53] W. Cho, J. Lee, J. Park, and D. Kim, “Radar polygon method: an areal rainfall estimation based on radar rainfall imageries,” *Stochastic Environmental Research and Risk Assessment*, vol. 31, no. 1, pp. 275–289, 2016.
- [54] J. Lee, J. Ahn, E. Choi, and D. Kim, “Mesoscale spatial variability of linear trend of precipitation statistics in Korean peninsula,” *Advances in Meteorology*, vol. 2016, Article ID 3809719, 15 pages, 2016.



Hindawi

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

