Ecological Informatics 36 (2016) 172–180

Contents lists available at ScienceDirect

Ecological Informatics

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

Application of an artificial neural network (ANN) model for predicting mosquito abundances in urban areas

Keun Young Lee [a,*], Namil Chung [b], Suntae Hwang [c]

[a] UBITA (Institute for Ubiquitous Information Technology and Applications), Konkuk University, Seoul, Korea
[b] Department of Ecological Monitoring and Assessment, National Institute of Ecology, Seochon, Choongnam 325-813, Republic of Korea
[c] School of Computer Science, Kookmin University 77, Jeongneung-ro, Seongbuk-gu, Seoul, 136-702, Korea

ARTICLE INFO

Article history:
Received 22 February 2015
Received in revised form 19 July 2015
Accepted 20 August 2015
Available online 10 September 2015

Keywords:
Mosquito abundances
Artificial neural network (ANN)
Multiple regression model (MLR)
Empirical prediction model

ABSTRACT

The mosquito species is one of most important insect vectors of several diseases, namely, malaria, filariasis, Japanese encephalitis, dengue, and so on. In particular, in recent years, as the number of people who enjoy outdoor activities in urban areas continues to increase, information about mosquito activity is in demand. Furthermore, mosquito activity prediction is crucial for managing the safety and the health of humans. However, the estimation of mosquito abundances frequently involves uncertainty because of high spatial and temporal variations, which hinders the accuracy of general mechanistic models of mosquito abundances. For this reason, it is necessary to develop a simpler and lighter mosquito abundance prediction model. In this study, we tested the efficacy of the artificial neural network (ANN), which is a popular empirical model, for mosquito abundance prediction. For comparison, we also developed a multiple linear regression (MLR) model. Both the ANN and the MLR models were applied to estimate mosquito abundances in 2-year observations in Yeongdeungpo-gu, Seoul, conducted using the Digital Mosquito Monitoring System (DMS). As input variables, we used meteorological data, including temperature, wind speed, humidity, and precipitation. The results showed that performances of the ANN model and the MLR model are almost same in terms of R and root mean square error (RMSE). The ANN model was able to predict the high variability as compared to MLR. A sensitivity analysis of the ANN model showed that the relationships between input variables and mosquito abundances were well explained. In conclusion, ANNs have the potential to predict fluctuations in mosquito numbers (especially the extreme values), and can do so better than traditional statistical techniques. But, much more work needs to be conducted to assess meaningful time delays in environmental variables and mosquito numbers.

1. Introduction

There have been increasing concerns about the effects of global warming on ecosystems (Mohseni et al., 2003; Botkin et al., 2007), including the impact on the growth and activity of insects vectors of diseases to humans (Patz et al., 1998; Revich et al., 2012). Among these, mosquitoes are known for causing more human suffering than any other organisms. Great numbers of people die from mosquito-borne disease worldwide each year (WHO, 1996). Increasing temperature due to global warming could elevate the growth rates of larval mosquitoes (Bayoh and Lindsay, 2003; Bayoh and Lindsay, 2004), leading to a larger number of adult mosquitoes and thus more incidences of mosquito-borne diseases (Reiter, 2001).

To predict mosquito abundances, mechanistic models have been developed using various methods. Some mosquito abundance model based on complex biological processes use computer simulations (Focks et al., 1993; Fougue and Baumgartner, 1996), while other researchers have used chemical properties such as temperature and rainfall to develop a mosquito population dynamics model (Ahumada et al., 2004; Shone et al., 2006). Since these dynamic simulations for mosquito abundances include numerous parameters and require complicated domain knowledge, simpler studies have recently been conducted on mosquito abundance prediction. A dynamic hydrology model and a Poisson regression with a genetic algorithm have been used for mosquito abundance prediction (Shaman et al., 2002; Lebl et al., 2013). Although mosquito abundance prediction models have become simpler, it remains difficult to control many parameters and understand the domain knowledge. Therefore, it is necessary to develop a lighter prediction model, such as an empirical approach, for predicting mosquito abundances in particular.

Among the empirical approaches, the artificial neural networks (ANNs), in particular, multilayer perceptrons (MLP), were widely applied in the last decades in the fields of bioinformatics (Dopazo et al., 1997), ecology (Lek et al., 1996; Lek and Guegan, 1999), and environmental engineering (Sahoo et al., 2009; Singh et al., 2009; Hill and

* Corresponding author at: UBITA, Konkuk University 120 Neungdong-ro, Gwangjin-gu, Seoul 05029, Korea.

http://dx.doi.org/10.1016/j.ecoinf.2015.08.011

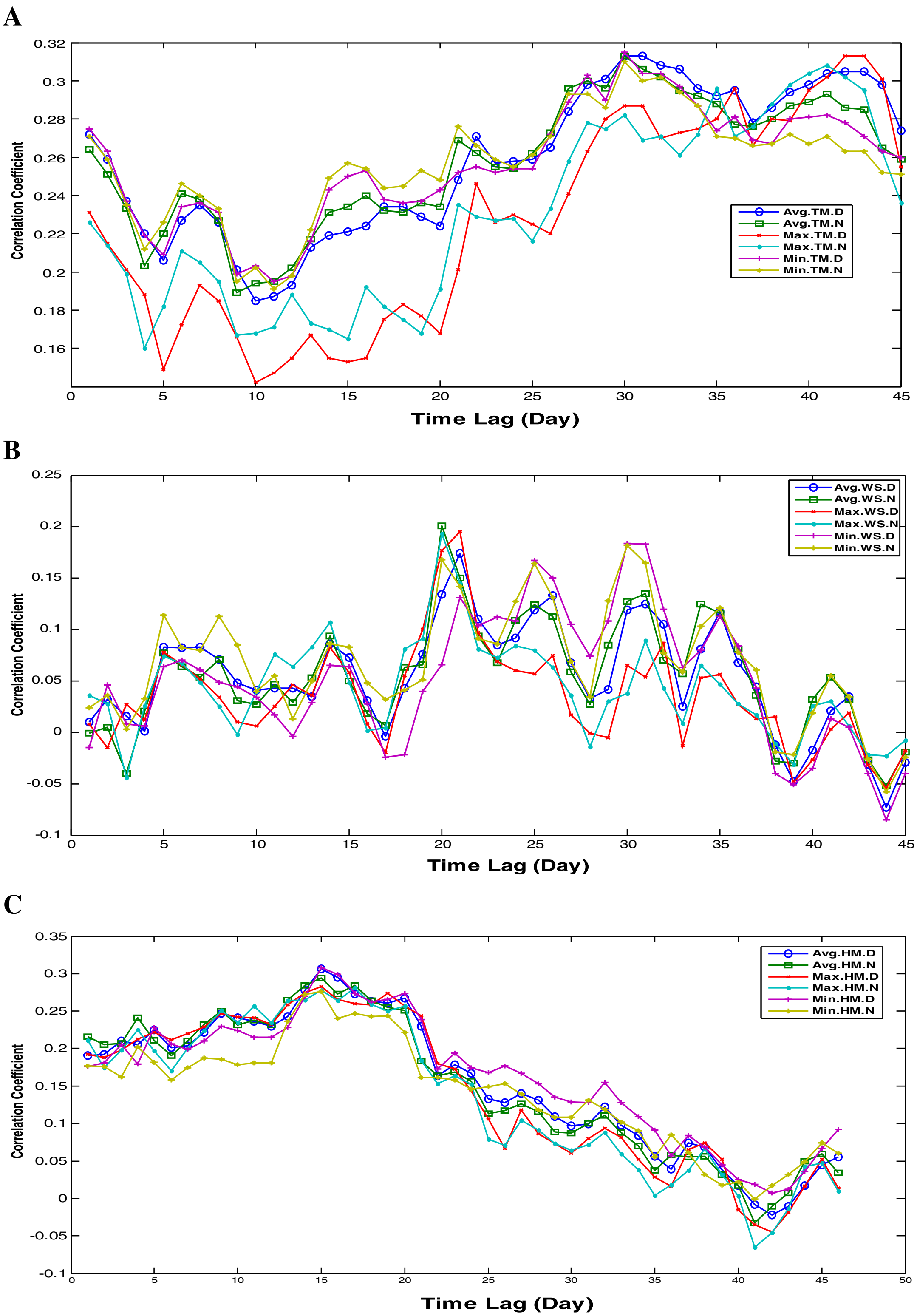

Fig. 1. Correlations and time lags between mosquitoes and meteorological data from temperature (A), wind speed (B), humidity (C), and precipitation (D), where D means daily (24 h) and N means nighttime.

Minsker, 2010). In particular, the good performance of ANNs in various ecological models was verified (Brosse et al., 1999; Park et al., 2003; Song et al., 2013). Moreover, in recent years, ANNs have been successfully applied for classifying and identifying mosquito species (Banerjee et al., 2008).

As a sensitivity analysis, a "weight" analysis can be employed to explain the relations between the input variables and the output variable in ANNs (Garson, 1991). The "weight" analysis computes the strength of the connections between the input factors and the output factors quantitatively. We anticipated that by exploiting this virtue the mosquito activity could be predicted successfully by an ANN model, resulting in a greater degree of estimation accuracy than that provided by other models (Song et al., 2013).

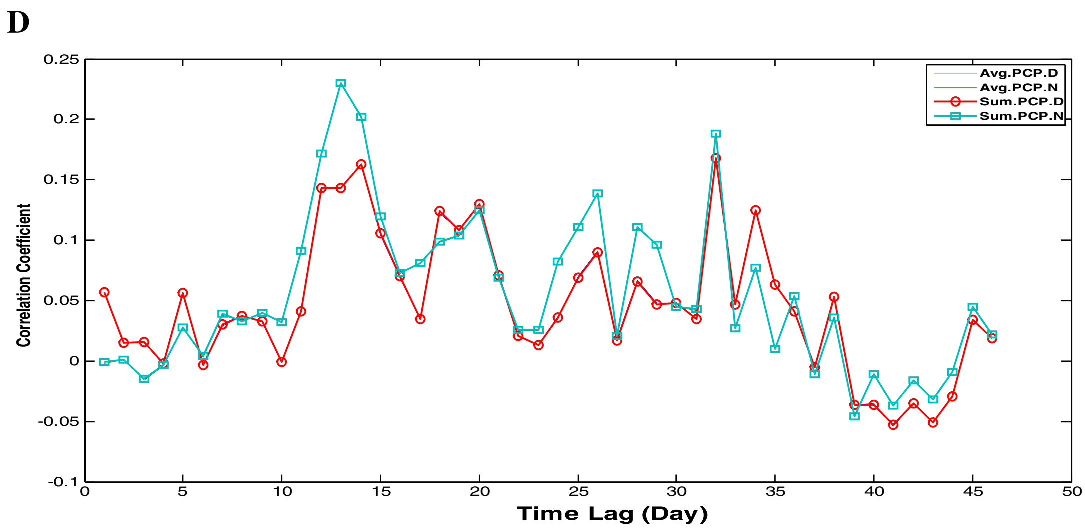

Fig. 1 (*continued*).

In this study, our objective was to estimate the validity of ANNs for forecasting mosquito abundances. To accomplish this, we developed both an ANN model and a multiple linear regression (MLR) for the purpose of comparison. Since ANNs have not yet been applied to predict mosquito abundances in urban areas, our research may facilitate empirical model studies for predicting mosquito abundances. The cross-correlation method can select the key variable to determine the structures of the ANN and MLR. The proposed models can predict mosquito abundances through real sensing. In addition, we also attempted to determine the importance of each meteorological variable using sensitivity analysis.

2. Materials and methods

2.1. Sampling

2.1.1. Mosquito data

Daily mosquito abundance was monitored at a site in Yeongdeungpo-gu, Seoul, Korea, using the Digital Mosquito Monitoring System (DMS; E-TND, Korea) during May–October 2011 and 2012. The DMS attracts mosquitoes by spraying CO2 gas at a rate of 300 ml/min, and traps the lured mosquitoes in a mesh bag by using a fan to create an air current flowing into the trap. The continuous air current inhibits the captured mosquitoes from escaping the mesh trap. The DMS automatically counts the number of mosquitoes trapped by the air current using an infrared sensor. Since the CO_2 gas spray was started at 5:00 p.m. and ended at 7:00 a.m., daily mosquito abundance was considered to be the number of mosquitoes captured from 8 am on 1 day to 7 am on the following day.

Since the number of mosquitoes according to the automatic monitoring using the DMS was abnormally overestimated, mainly because of the back-and-forth movement of spiders through the infrared sensor of the DMS, the raw data were processed to filter out noise data. According to the advice of the mosquito control experts who had monitored the same sites for several years, the daily maximum number of mosquitoes captured was set at 200 individuals for a site located in a residential areas based on land use type. The raw data on any given date exceeding the daily maximum values were replaced with the average values of the data for the previous and the following day. If the noise data were recorded more than 2 days consecutively, the raw values were deleted and left as blank. The same rules were applied to the missing values. The processed data were used for further analysis.

2.1.2. Weather data

Hourly estimates of air temperature (TM), wind speed (WS), humidity (HM), and precipitation (PCP) were obtained from the Korea Meteorological Administration, Korea (www.kma.go.kr), during 2011–2012. The whole area of Yeongdeungpo-gu was completely covered by three lattices having a resolution of 5 km × 5 km. The distance between our site and the center of each of these three lattices was calculated using a geographical information system to pair the sets. The hourly values were averaged to calculate the daily means of TM, WS, HM, and PCP. Hourly precipitation estimates were summed to give the daily precipitation as well.

2.2. Model development

2.2.1. Data preprocessing

It is well known that meteorological variable such as temperature, humidity, and rainfall influence mosquito population (Wang et al., 2011; Chuang et al., 2012; Lebl et al., 2013). Thus, to predict mosquito abundances, it is reasonable to consider meteorological variables as input. In this study, we therefore used meteorological data including TM, WS, HM, and PCP as input variables. More precisely, we used daily (24 h) and nighttime (5:00 p.m.–7:00 a.m.) average, maximum and minimum values for TM, WS, and HM, respectively, and daily and nighttime average and summation values for PCP. In addition, the number of mosquitoes on the preceding was used as an input variable. The measured numbers of mosquitoes were considered as an output variable in all models. The data set for the entire 2 years consisting of 317 records for all variables was used in this study. The first 220 observations were used to develop the prediction models. The remaining observations were used as test data to verify the prediction models. The MLR and ANN models were developed to predict the mosquito abundances in urban areas.

To extract the key input variables, we analyzed the correlation coefficients with an imposed time lag. It was assumed that the daily abundance of mosquitoes in day t (Y_t) was related to the values of the weather variables in day t-k (X_{t-k}). The association between the values of Y_t and X_{t-k} was estimated by comparing cross-correlation between Y_t and X_{t-k} when k varies from 0 to 45. It did not make any sense to compare the association between the mosquito abundance in day t and the weather variables in day $t + k$ ($k > 0$). The cross-correlation function quantifies the association between the two variables with a time lag of k days. The cross-correlation function is based on the Pearson correlation function, except that the X variable is shifted in time with a lag of

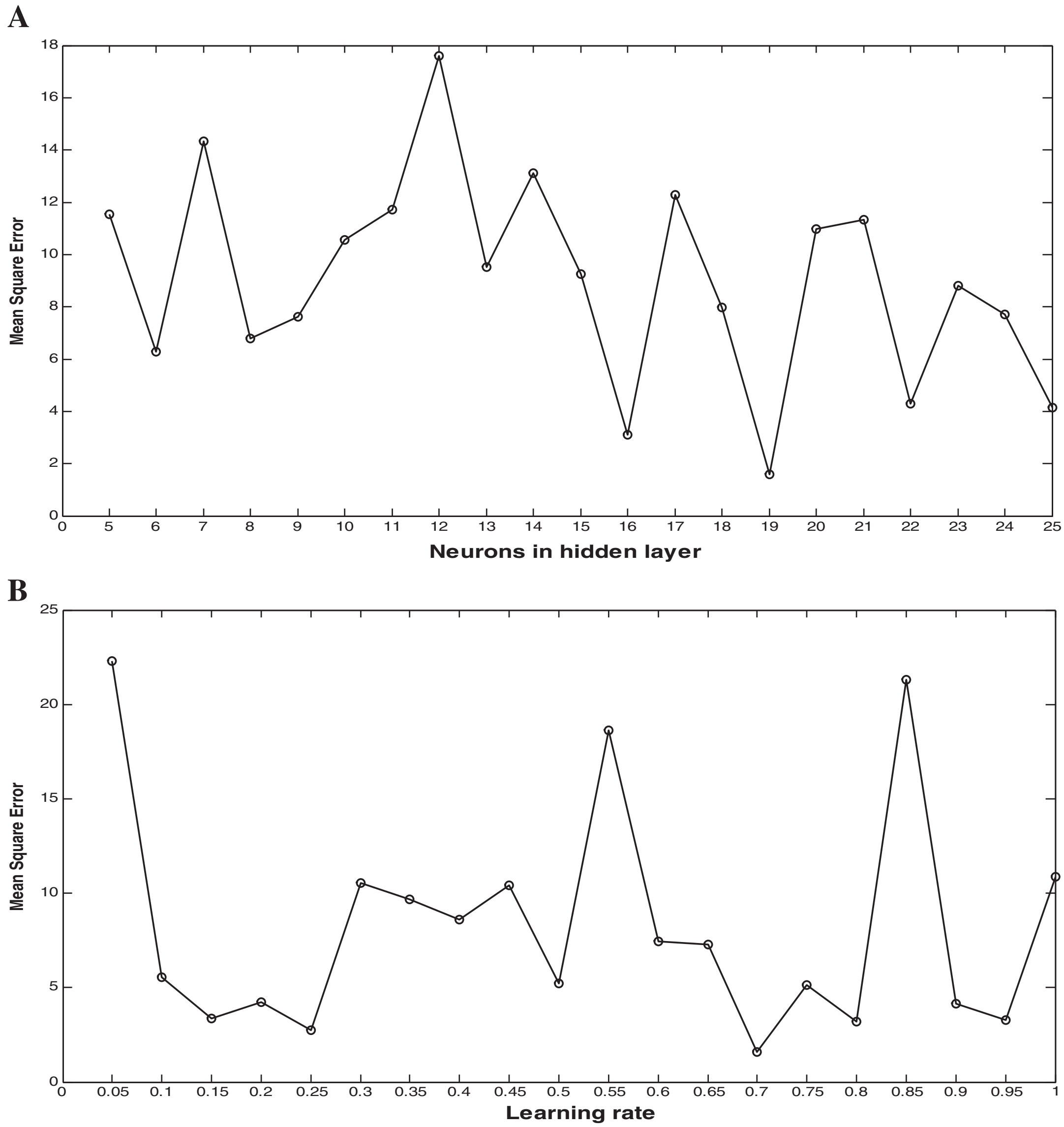

Fig. 2. Mean square error between measured data and model output from variation with the number of neurons in the hidden layer (A) and variation with the learning rate (B).

k days (Zuur et al., 2007). The cross-correlation was calculated by (Diggle, 1990; Chartfield, 2003)

$$\hat{\rho}_{YX}(k) = \begin{cases} \dfrac{1}{N}\dfrac{\sum_{t=1}^{N-k}(Y_t-\overline{Y})(X_{t+k}-\overline{X})}{s_Y s_X} & if\ k \geq 0 \\ \dfrac{1}{N}\dfrac{\sum_{t=1-k}^{N}(Y_t-\overline{Y})(X_{t+k}-\overline{X})}{s_Y s_X} & if\ k < 0 \end{cases}$$

where s_Y and s_X are sample standard deviation of the time series Y_t (mosquito abundances) and X_t (weather variables), respectively. The results of the cross-correlation can be plotted in a graph in which various time lags are plotted along the horizontal axis and the correlations along the vertical axis.

In this study, the cross correlations were calculated for the daily mosquito abundances at the monitoring site and the weather condition. The weather variables included TM, WS, HM, and PCP. The values for temperature, wind speed, and humidity comprised of the daily (24 h) and the nighttime (5:00 p.m.–7:00 a.m.) average, maximum and minimum values of each variable. The values for precipitation were summed for daily and nighttime hours.

Since the correlation coefficients can be considered a measure of the extent to which a certain influential variable corresponds to the data, we were able to determine the important variables of the prediction model. Fig. 1 illustrates the correlation of mosquito abundance data with other time lagged input data. Each input variable appears to provide some information that could facilitate mosquito abundance prediction. Here, the minimum value of TM for the preceding 29 days (TM_{t-29}), the average value of WS for the preceding 19 days (WS_{t-19}), the average value of HM for the preceding 14 days (HM_{t-14}), and the average value of PCP for the preceding 12 days (PCP_{t-12}) were selected as sensitive variables since they influence the prediction more than any other variables. To evaluate the prediction models, we used the correlation coefficients (R) and the index of agreement (IA). The IA is a relative measure and, therefore, allows different models to be compared using different data sets. The coefficient of determination (R^2), although a measure that is easy to understand, is not suitable for comparison purposes (Willmott, 1982). The relative importance of input variables was studied through a sensitivity analysis in the case of the ANN model.

2.2.2. Multiple linear regression (MLR)

The MLR model constitutes a technique for forecasting process design, optimization, and process control (Kim et al., 2010). The general

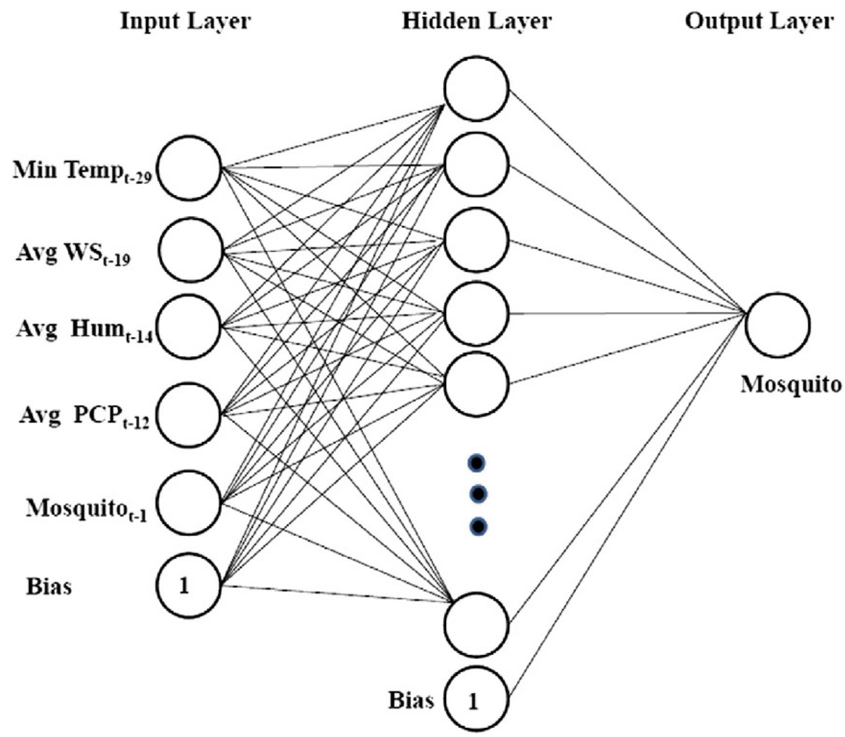

Fig. 3. A presentation of feed-forward artificial neural network.

MLR is represented by

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n$$

where X_i denotes the input variables, while Y the output variable, and n is the number of input variables included in the model, and β_i the regression coefficients. The goal of MLR is to find an approximation function for the prediction future response of the system output. We estimated mosquito abundances using the MLR model in MATLAB ver. 2010b.

2.2.3. Artificial neural network

A multilayer ANN is used to create models of a system state using nonlinear combinations of the input variables (Bishop, 1995; Duda et al., 2001; Hastie et al., 2001). The ANN employed in this study is a feed-forward network with sigmoid activation functions in the hidden layers and a linear activation function in the output node in MATLAB ver. 2014a. Since according to Bishop's (1995) study more than one hidden layer is often not necessary, our architectures have only one hidden layer. The ANN is trained using a back-propagation algorithm with gradient descent and momentum terms.

The ANN requires that the learning rate, number of nodes in a single hidden layer, and maximum number of training epochs are specified (Hill and Minsker, 2010). In this study, we used the optimal number error approach. The number of nodes in the hidden layer was varied between 5 and 23, and the learning rate was varied from 0.01 to 1.0 in increments of 0.05. For each configuration, the mean square error (MSE) between the model output and the measured data was calculated. Fig 2 illustrates the optimal number of neurons in the hidden layer and the optimal learning rate having the maximum model performance as indicated by MSE. The number of neurons in the hidden layer and the optimal learning rate were selected using a trial-and-error method. The final ANN structure had five input variables with one node accounting for bias, 19 hidden neurons with one node accounting for bias, a 0.7 learning rate, and one output variable of the output layer (Figs. 2 and 3).

For the purpose of prediction, the most important property of a model is its competence to generalize. While generalization competence indicates a model's power to perform well on data that were not used to train it, overfitting prevents model generalization in the face of new situations (Schlink et al., 2003). To avoid overfitting, early stopping the most frequently used regularization technique was employed. In order to apply it, the data set was randomly split into two sets, 80% for model training (to compute the gradient and updating of the network parameters, such as weights and biases—the training set) and 20% for model testing (to test the model error validation—the validation set). The model weights were randomly initialized and the training process was stopped when the network began to overfit the data, i.e., the error on the validation set.

In order to investigate the explanatory competence of ANN, we applied the aforementioned ‘weight’ sensitivity analyses to determine the relative contribution and roles of input variables in mosquito activity (Garson, 1991; Song et al., 2013). The weights method was developed by Garson (1991). The percentages of the influence of input

Table 1
Mean and standard deviation values of meteorological characteristics and mosquitoes during the 2-year experimental period.

	TM (°C)	WS (m/s)	HM (%)	PCP (mm)	Mosquitoes
Mean ± SD	18.62 ± 4.79	2.61 ± 1	76.04 ± 12.4	0.4 ± 1.23	6.71 ± 12.11
Ranges	4.21–27.9	0.9–8.38	44.42–99.05	0–10.93	0–78

Table 2
Standardized coefficients of estimated mosquito activity of multiple linear regression models.

Variables	Standardized coefficient
TM$_{t-29}$	0.006
WS$_{t-19}$	0.082
HM$_{t-14}$	0.003
PCP$_{t-12}$	−0.05
Mosquito$_{t-1}$	0.96

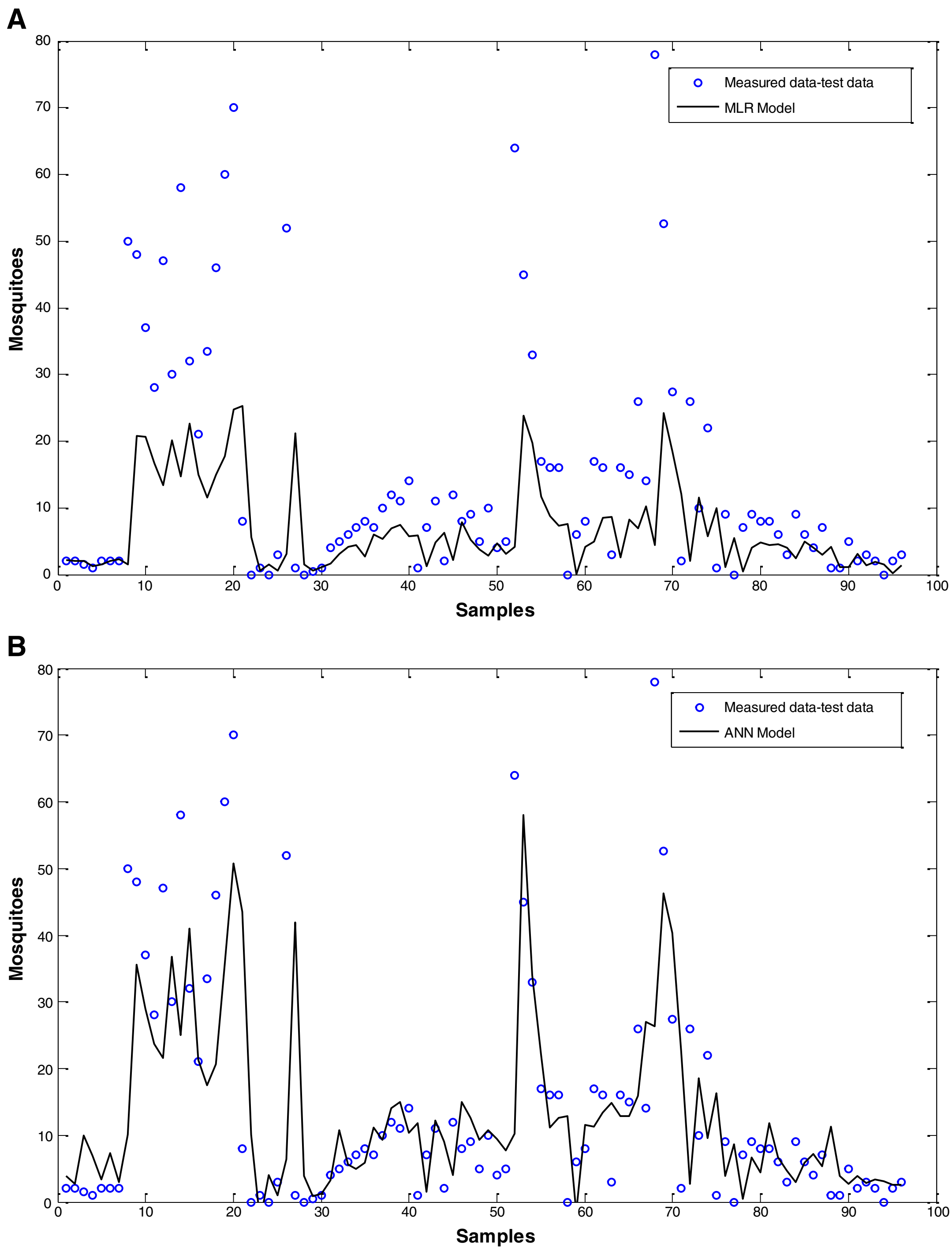

Fig. 4. Measured and predicted mosquito abundances in urban area obtained by MLR (A), ANN (B), MLR_C (C), and ANN_C (D).

variable on the output value, Q_{ik} (%), indicating the importance of input variables were determined by

$$Q_{ik}(\%) = \frac{\sum_{j=1}^{n}\left(\frac{|w_{ij}|}{\sum_{i=1}^{m}|w_{ij}|}|v_{jk}|\right)}{\sum_{i=1}^{m}\left(\sum_{j=1}^{n}\left(\frac{|w_{ij}|}{\sum_{i=1}^{m}|w_{ij}|}|v_{jk}|\right)\right)} \times 100$$

where w_{ij} represents the weights between the input neuron i (=1, 2, …, m) and the hidden neuron j (=1, 2, …, n), and v_{jk} represents the weights between the hidden neuron j and the output neuron k (=1, 2, …, l) (see, Song et al., 2013).

3. Results

The averaged meteorological properties and mosquito abundances at the test site are provided in Table 1. High variability was observed in HM and the number of mosquitoes. The number of mosquitoes was positively correlated with all the meteorological properties (Fig. 1). Mosquito$_{t-1}$ showed the highest coefficients in multiple linear regressions (Table 2). The regression model predicted well for low values of mosquitoes, while it could not predict relatively high values (Fig. 4A).

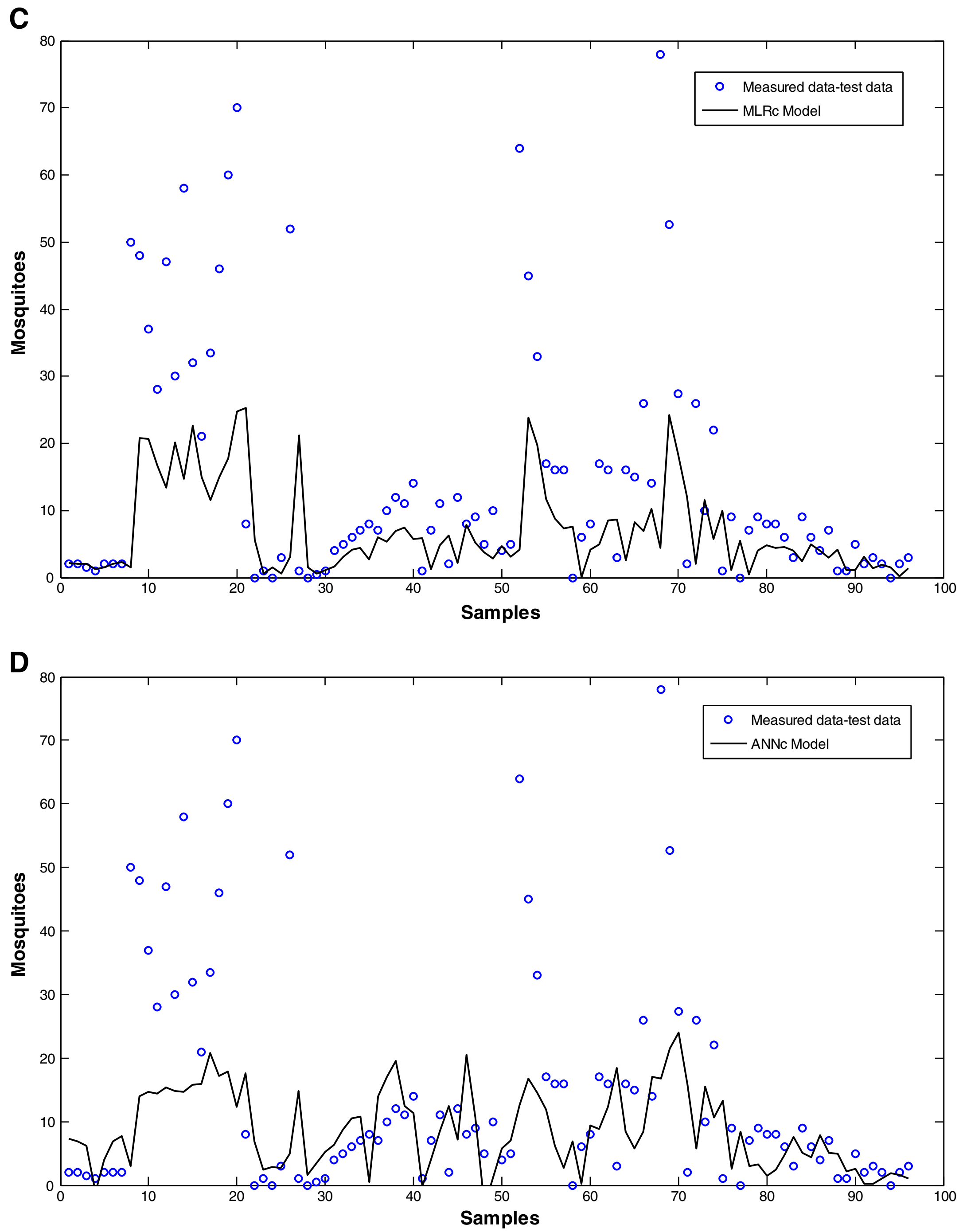

Fig. 4 (*continued*).

The regression model provides both of an average *R* value of 0.56 and a low IA value of 0.53 for the model test with RMSE of 17.53 (Table 3).

Table 3
Mosquito abundance model performance statistics: root mean square error (RMSE), index of agreement (IA), and correlation coefficient (*R*) between measured and estimated values.

	Mean	SD	RMSE	IA	*R*
MLR	6.73	5.99	17.53	0.53	0.56
ANN	13.28	12.52	14.38	0.75	0.61
MLR_C	6.97	6.66	17.23	0.56	0.56
ANN_C	8.67	6.15	17.04	0.49	0.5

The ANN model predicted well for high mosquitoes in comparison to the regression model. However, in general, from performance of model point of view, the ANN model did not appear to be better than the regression model fit to the data because of an average *R* value of 0.61 and an average IA of 0.75 for the model test with an RMSE of 14.38. The reason for this estimates high fluctuation mosquito abundances and relative small correlation coefficients between meteorological variables and mosquito abundances.

In the 'weight' sensitivity analysis, HM accounted for almost 20% of the variability in mosquito activity. Three major input variables including, TM, WS, and HM contributed to 65% of the variation in mosquito activity (Fig. 5).

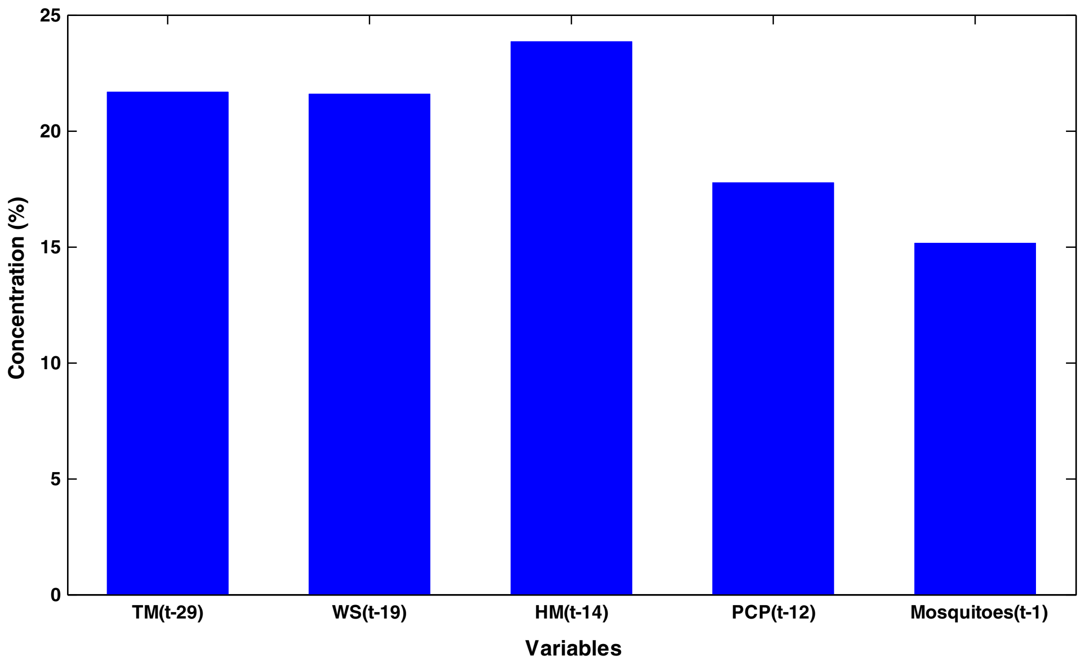

Fig. 5. Contribution of input variable according to 'weight' analysis.

To verify the influence of the time lagged variables, we did additional works. As the time lagged variables, we used TM_{t-29}, WS_{t-19}, HM_{t-14}, PCP_{t-12}, and $Mosquito_{t-1}$ as input variables. As comparable model, we developed prediction models using present meteorological data, that is, TM_t, WS_t, HM_t, PCP_t, and $Mosquito_{t-1}$ as input variables. We call these models MLR_C and ANN_C. We found that MLR_C achieved an average R value of 0.56 and a low IA value of 0.56 but ANN_C achieved a relatively low R value of 0.5 and an IA of 0.49 (Table 3). We found that the MLR_C has almost same performance of the MLR. Then it follows that time lag could not influence the predictability of the MLR. However, the ANN_C lack as much performance compared to the ANN in terms of IA and RMSE. This reflects that time lag influences the performance of the ANN model significantly.

4. Discussion

In this study, two different mosquito abundance models were developed, and their estimation performances of mosquito abundance was evaluated through a comparison with the measured mosquito abundances. In these two models, we found that different environmental parameters were selected as the major factors of mosquito activity. That is, WS_{t-19} and $Mosquito_{t-1}$ were selected as the major factors in the MLR model based on their high coefficient, while HM_{t-14} and TM_{t-29} were selected as the major factors in ANN model based on weighted contribution in the ANN. Moreover, we found that the shape of the contributions of the variables in the MLR and ANN models was differentiated.

However, we found a little difference in the model's performance. MLR predicted normally with average R and IA values. This reflects that MLR could forecast the average variability of mosquito abundances but was not able to predict the high variability. In case of the ANN model, its performance is almost same performance of the MLR in terms of R and IA. The difference of both of two is predictability for the high variability of mosquito abundances. The ANN model was able to predict the high variability as compared to MLR. The reason for this estimates immanence nonlinearity and generalization of ANNs during the learning status.

'Weight' analysis presents an overall contribution of input variables (Gevrey et al., 2003). Using 'weight' analysis, we could investigate the linearity/nonlinearity of the relationships between the variables and mosquito abundances. For example, the coefficient of HM_{t-14} from MLR is 0.03, but the contribution of HM_{t-14} based on 'weight' analysis is 23%. This reflects that the relationship between HM_{t-14} and mosquito abundances is nonlinear. Thus, sensitivity analysis helped to determine the role of the variables in the ANNs. Since the ANN is regarded as a black box model, we could not determine the precise relationships between the input variables and output variables in the ANN itself. However, the application of sensitivity analysis can resolve this problem and shed light on relationships between meteorological factors and mosquito abundances.

To develop a robust ANN, we have to consider the selection of the number of layers, the number of neurons in the hidden layer, the learning rates, and the number of epochs for model training carefully. For example, if we consider an insufficient number of neurons in the hidden layer, then the ANN cannot reflect nonlinearity within the training data. Conversely, if we consider too many neurons, then the ANN has an overfitting problem, and hence, this leads a lack of generalizability. In this study, we applied a trial-and-error method, which is known to be the best method to determine the appropriate number of neurons and learning rate (Shamseldin, 1997; Hill and Minsker, 2010), and an early stopping technique to hinder overfitting.

In summary, both the ANN model and the MLR model are not quite good enough to predict mosquito abundances in this study. However, ANNs have the potential to predict fluctuations in mosquito numbers (especially the extreme values) and can do so better than traditional statistical techniques. Much more work needs to be conducted to assess meaningful time delays in environmental variables and mosquito numbers. In addition, a sensitivity analysis confirmed the contribution of major factors and their relationships with mosquito activity.

Acknowledgments

The authors would like to thank Prof. Karpjoo Jeong and Prof. Sangmin Yoon for valuable comments. This work was funded by the

Weather Information Service Engine Project under Grant KMIPA-2012-0001.

References

Ahumada, J.A., Lapointe, D., Samuel, M.D., 2004. Modeling the population dynamics of *Culex quinquefasciatus* (Diptera: Culcidae), along an elevation gradient in Hawaii. J. Med. Entomol. 41, 1157–1170.

Banerjee, A.K., Kiran, K., Murty, U.S.N., Venkateswarlu, Ch., 2008. Classification and identification of mosquito species using artificial neural networks. Comput. Biol. Chem. 32, 442–447.

Bayoh, M.N., Lindsay, S.W., 2003. Effect of temperature on the development of the aquatic stages of *Anopheles gambiae sensu stricto* (Diptera: Culicidae). Bull. Entomol. Res. 93 (05), 375–381.

Bayoh, M.N., Lindsay, S.W., 2004. Temperature-related duration of aquatic stages of the Afrotropical malaria vector mosquito *Anopheles gambiae* in the laboratory. Med. Vet. Entomol. 18 (2), 174–179.

Bishop, C., 1995. Neural Networks for Pattern Recognition. Oxford University Press, New York.

Botkin, D.B., Saxe, H., et al., 2007. Forecasting the effects of global warming on biodiversity. Bioscience 57 (3), 227–236.

Brosse, S., Lek, S., Dauba, F., 1999. Predicting fish distribution in a mesotrophic lake by hydroacoustic survey and artificial neural networks. Limnol. Oceanogr. 44, 1293–1303.

Chartfield, C., 2003. The Analysis of Time Series: An Introduction. 6th edition. Chapman and Hall, London.

Chuang, T.W., Lonides, E.L., Knepper, R.G., Stanuszek, W.W., Walker, E.D., Wilson, M.L., 2012. Cross-correlation map analyses show weather variation influences on mosquito abundance patterns in Saginaw County, Michigan, 1989–2005. J. Med. Entomol. 49, 851–858.

Diggle, P.J., 1990. Time Series: A Biostatistical Introduction. Oxford University Press, London.

Dopazo, J., Huaichun, W., Carazo, J.M., 1997. A new type of unsupervised growing neural network for biological sequence classification that adopts the topology of a phylogenetic tree. Lect. Notes Comput. Sci. 1240, 932–941.

Duda, R., Hart, P., Stork, D., 2001. Pattern Classification. Wiley-Interscience, New York.

Focks, D.A., Haile, D.G., Daniels, E., Mount, G.A., 1993. Dynamic life table model for *Aedes-Aegypti* (Diptera, Culcidae)—analysis of the literature and model development. J. Med. Entomol. 30, 1003–1017.

Fouque, F., Baumgartner, J., 1996. Simulating development and survival of *Aede vexans* (Diptera, Culcidae) preimaginal stages under field conditions. J. Med. Entomol. 33, 32–38.

Garson, G.D., 1991. Interpreting neural network connection weights. Artif. Intell. Expert 6, 47–51.

Gevrey, M., Dimopoulos, I., Lek, S., 2003. Review and comparison of methods to study the contribution of variables in artificial neural network models. Ecol. Model. 160, 249–264.

Hastie, T., Tibshirani, R., Friedman, J., 2001. The Elements of Statistical Learning. Springer-Verlag, New York.

Hill, D.J., Minsker, B.S., 2010. Anomaly detection in streaming environmental sensor data: a data-driven modeling approach. Environ. Model Softw. 25, 1014–1022.

Kim, M.H., Kim, Y.S., Lim, J.J., Kim, J.T., Sung, S.W., Yoo, C.K., 2010. Data-driven prediction model of indoor air quality in an underground space. Korean J. Chem. Eng. 27 (6), 1675–1680.

Lebl, K., Brigger, K., Rubel, F., 2013. Predictiong on *Culex pipiens/restuans* population dynamics by interval lagged weather data. Parasites & Vectors 6, 129.

Lek, S., Guégan, J.F., 1999. Artificial neural networks as a tool in ecological modeling, an introduction. Ecol. Model. 120, 65–73.

Lek, S., Belaud, A., Baran, P., Dimopoulos, I., Delacoste, M., 1996. Role of some environmental variables in trout abundance models using neural networks. Aquat. Living Resour. 9, 23–29.

MATLAB and Statistics Toolbox Release, 2014a. The MathWorks, Inc., Natick, Massachusetts, United States.

MATLAB and Statistics Toolbox Release, 2014b. The MathWorks, Inc., Natick, Massachusetts, United States.

Mohseni, O., Stefan, H.G., Eaton, J.G., 2003. Global warming and potential changes in fish habitat in U.S. streams. Clim. Chang. 59 (3), 389–409.

Park, Y.S., Cereghino, R., Compin, A., Lek, S., 2003. Applications of artificial neural networks for patterning and predicting aquatic insect species richness in running waters. Ecol. Model. 160, 265–280.

Patz, J.H., Martens, W.J., Focks, D.A., Jetten, T.H., 1998. Dengue fever epidemic potential as projected by general circulation models of global climate change. Environ. Health Perspect. 106, 147–153.

Reiter, P., 2001. Climate change and mosquito borne disease. Environ. Health Perspect. 109, 141–161.

Revich, B., Tokarevich, N., Parkinson, A.J., 2012. Climate change and zoonotic infections in the Russian Arctic. Int. J. Circumpolar Health 71, 1–8.

Sahoo, G.B., Schladow, S.G., Reuter, J.E., 2009. Forecasting stream water temperature using regression analysis, artificial neural network, and chaotic non-linear dynamic models. J. Hydrol. 378, 325–342.

Shaman, J., Stieglitz, M., Stark, C., Blancq, S.L., Cane, M., 2002. Using a dynamic hydrology model to predict mosquito abundances in flood and swamp water. Emerg. Infect. Dis. 8, 6–13.

Shamseldin, A.Y., 1997. Application of a neural network technique to rainfall-runoff modeling. J. Hydrol. 199, 272–294.

Schlink, U., Doring, S., Pelikan, E., Nunnari, G., Cawley, G., Junninen, H., Greig, A., Foxall, R., Eben, K., Chatterton, T., Vondracek, J., Richter, M., Dostal, M., Bertucco, L., Kolehmainen, M., Doyle, M., 2003. A rigorous inter-comparison of ground-level ozone predictions. Atmos. Environ. 37, 3237–3253.

Shone, S.M., Curriero, F.C., Lesser, C.R., Glass, G.E., 2006. Characterizing population dynamics of *Aedes sollicitans* (Diptera, Culcidae) using meteorological data. J. Med. Entomol. 43, 393–402.

Singh, K.P., Basant, A., Malik, A., Jain, G., 2009. Artificial neural network modeling of the river water quality-case study. Ecol. model. 220, 888–895.

Song, K., Park, Y.S., Zheng, F., Kang, H., 2013. The application of artificial neural network (ANN) model to the simulation of denitrification rates in mesocosm-scale wetlands. Ecol. Inform. 16, 10–16.

Wang, J., Ogden, N.H., Zhu, H., 2011. The impact of weather conditions on *Culex pipiens and Culex restuans* (Diptera, Culcidae) abundance: a case study in peel region. J. Med. Entomol. 48, 468–475.

WHO, 1996. World Health Report, 1996. World Health Organization, Geneva.

Willmott, C.J., 1982. Some comments in the evaluation of model performance. Bull. Am. Meteorol. Soc. 63, 1309–1313.

Zuur, A.F., Ieno, E.N., Smith, G.M., 2007. Analysing Ecological Data. Springer, New York.