

# Application of artificial neural networks to frost detection in central Chile using the next day minimum air temperature forecast

Marcel Fuentes<sup>1</sup>, Cristóbal Campos<sup>1</sup>, and Sebastián García-Loyola<sup>2\*</sup>

<sup>1</sup>Instituto de Investigaciones Agropecuarias, INIA Quilamapu, Av. Vicente Méndez 515, Chillán, Chile.

<sup>2</sup>Universidad de Concepción, Facultad de Ciencias Naturales y Oceanográficas, Barrio Universitario s/n, Concepción, Chile.

\*Corresponding author (sgarcial@udec.cl).

Received: 23 December 2017; Accepted: 3 June 2018; doi:10.4067/S0718-58392018000300327

## ABSTRACT

Predicting future climatic events is one of the key issues in many fields, whether in scientific or industrial areas. An artificial neural network (ANN) model, based on a backpropagation type, was developed in this study to predict the minimum air temperature of the following day from meteorological data using air temperature, relative humidity, radiation, precipitation, and wind direction and speed to detect the occurrence of radiative frost events. The configuration of the next day ANN prediction system allows operating with low-power computing machines; it is able to generate early warnings that can lead to the development of effective strategies to reduce crop damage, lower quality, and losses in agricultural production. This paper presents a procedural approach to an ANN, which was trained, validated, and tested in 10 meteorological stations in central Chile for approximately 8 yr (2010-2017). The overall mean results were classified by a confusion matrix and showed good performance in predicting minimum temperature with a mean square error (MSE) of 2.99 °C for the network, 1.71 °C for training, 1.77 °C for validation, and 1.74 °C for the testing processes. Frost detection results had an appropriate 98% overall mean accuracy (ACC), 86% sensitivity (TPR), and 2% error rate (ER). Differences and errors in frost detection can be attributed to several factors that are mainly associated with the accuracy of the sensors meteorological stations, local climatic and geographic conditions, and the number of parameters that enter in the ANN training processes.

**Key words:** Artificial neural network (ANN), frost detection, next day forecasting.

## INTRODUCTION

Frost is a climatic event that generates great concern in agricultural activities due to the potential socioeconomic losses it can produce (ODEPA, 2013). Frost is defined in the Glossary of Meteorology (AMS, 2000) as the condition that exists when the temperature of the earth's surface and earthbound objects falls below freezing ( $\leq 0$  °C); frost technically refers to the formation of ice crystals by either frozen dew or a phase change from water vapor to ice (Cunha, 1952). These frost events are usually classified as radiative or advective frosts or a combination of both (Lungu et al., 2010). Radiative frost mainly occurs at night and is caused by intense longwave radiation leaking to the Earth's surface, which moves and is lost to the atmosphere; it cools the area near the surface, specifically between 1.25 m and 2.0 m above the ground, and causes the water near the surface of plant leaves to freeze. This type of frost is generated by certain conditions such as calm wind or low wind speed, thermal inversion, low humidity, dew point and water vapor, and absence of clouds. Radiative frost is the most common and is characteristically sudden because it can happen in hours without warning and gives farmers a short time to prepare (FAO, 2010). On the other hand, advective frost can develop during the day or night as a result of a large-scale incursion of cold and dry air masses mainly coming from polar fronts; it is the most dangerous

for crops because it can remain active for several days and has negative consequences on agricultural yields. This type of frost can be predicted in advance because it is produced by large masses of cold air whose path can be followed via remote sensing and the direction and intensity of the air temperature can be observed over time. A known phenomenon of this type occurred in Chile in September 2013 when fruit and crop growers suffered one of the major advective frost events in the last decades; it affected much of central Chile and low temperatures continued for approximately one week. According to the Office of Agricultural Studies and Policies (ODEPA, 2013), the frost effects were devastating for agriculture in general, and the most affected regions were the Libertador General Bernardo O'Higgins Region with 50% and Metropolitan Region with 45% crop losses. The Maule and Biobío Regions had 40% losses and the Atacama, Coquimbo, and Valparaíso Regions 15%.

Therefore, climate characteristics affect and influence crop growth as well as its geographic distribution, and air temperature variability is one of the most important factors (Woods and Gondal, 2007). It is therefore important to maintain the control and monitoring of this parameter to react on time to sudden changes. Reducing damage mainly caused by radiative frosts could be achieved by predicting the next day minimum air temperature based on key weather indicators (e.g., solar radiation, dew point, wind speed and direction, precipitation, and relative humidity), thus allowing farmers to make decisions and prepare for possible frost events (Retamal-Salgado et al., 2017). It is known that climate is a multidimensional, dynamic, and chaotic process, and these properties make forecasting it a great challenge due to the nonlinear nature of meteorological data (Ashour et al., 2016). This is why the study of modern weather forecasting involves a combination of complex computer models, in situ observations, and knowledge of weather trends and patterns by an advanced contemporary methodology it has attracted the attention of many researchers and scientists in various fields and disciplines associated with Earth science (Malik et al., 2014; Ashour et al., 2016).

One of the most recent methodologies used to work with climate prediction models is the artificial neural network (ANN) (Venkatesan and Ramachandriah, 2018). The ANNs are computer systems with powerful data modeling tools that use an appropriate mathematical configuration inspired by the biological neurological network or animal brain to approximate linear and nonlinear functions as techniques to predict climatic conditions based on its great flexibility and efficiency compared with conventional methods or models known in this area (González-Camacho et al., 2008). Jacob and Murugan (2016) define the ANN as a group of simple processing units (i.e., neurons) that communicate among themselves by sending analog signals. These signals travel across weighted connections with other neurons and each one accumulates the information it receives, learns from this, and generates an output according to an internal activation function (e.g., Gaussian, sinusoidal, linear, and tangent); this output can provide an input for other neurons or can be a part of the network output. Although there are numerous types of ANNs (e.g., mono-layer, multi-layer, backpropagation, and Hopfield), the most common type of ANN is the multi-layer perceptron. This is a feed-forward, fully-connected hierarchical network typically comprised of three types of neuron layers that each include one or several neurons. The first, or the lowest layer, is an input layer in which external information or historical data are received. The last, or the highest layer, is an output layer in which the solution or prediction is obtained. The input layer and output layer are separated by one or more intermediate layers called hidden layers in which the information is processed (Castañeda-Miranda and Castaño, 2017).

Furthermore, neural networks can be viewed as highly nonlinear functions with the basic form  $F(x,w) = y$  where  $x$  is the input vector presented to the network,  $w$  are the network weights, and  $y$  is the corresponding output vector approximated or predicted by the network. The weight vector  $w$  is commonly ordered first by layer, then by neurons, and finally by the weights of each neuron plus its bias. The ANNs need training to be able to learn different processes that occur in climate; this is why a typical training method is to first partition the data series into three disjoint sets, each one with different ratios (Valipour, 2016). These sets are the training set in which the model learns about the dynamics and behavior of the time series, the validation set in which the results are compared with existing information, and the test set in which the model is tested with new data.

In addition to the abovementioned, it is quite common to combine different types of ANNs to generate better results such as in the case of the multi-layer perceptron with backpropagation, which in a simplified form consists in learning a predefined set of input-output pairs, for example, using a two-phase propagation-adaptation cycle. The first pattern is applied as a stimulus for the input layer of the network, and it is propagated through the following layers to generate the output, which provides the value of the error when compared with the one it wants to obtain. Second, these errors are transmitted backward, starting from the output layer to all the neurons of the intermediate hidden layer that contribute directly to the output, assigning a percentage of the approximate error to the participation of each neuron in the original output. This process is always repeated backward,

layer by layer, until all the neurons in the network have received an error that describes their contribution as related to the total error. Based on this received information, all the connection weights are readjusted, so that the next time the same pattern is present, the difference between the calculated and the desired output decreases (Ovando et al., 2005).

These types of models can be simple or complex depending on their purpose and configuration. They can handle large amounts of data that are able to tolerate incomplete information with noise or inaccurate data, and through their learning can generate a fast and accurate prediction system. It is important to note that data required for ANN input should not necessarily respond to a specific statistical distribution (Bocco et al., 2010).

The ANN models have become increasingly important in the processing and analysis of time series and calculations of future predictions; they are able to self-learn and adapt to any model because they function like a human brain, acquire knowledge by detecting the patterns and relationships in the data, and store them in artificial interneuronal connections known as synaptic weights (Chevalier, 2008; Malik et al., 2014; Ashour et al., 2016). From a practical and economical point of view, a simple ANN model requires neither high computational processing capacity nor a sophisticated and costly device to model meteorological variables, which makes them easily accessible (González-Camacho et al., 2008).

The objective of this study was to develop different artificial neural network models (ANNs) that are able to learn from historical information collected at meteorological stations, such as air temperature, relative humidity, wind, precipitation, and solar radiation, to predict next day minimum temperature to find the best configuration model to detect with the minimal error and anticipate by several hours possible radiative frost events.

## MATERIALS AND METHODS

### Computational characteristics

Ideally, it was thought that this model would be run on ordinary computers; therefore, to process the historical meteorological database and configure the ANN model, we used the free software R-project with the R programming language running the 'neuralnet' package v1.33, which is a toolbox containing the functions and scripts to develop an ANN model (Fritsch et al., 2016). The machine was an Asus computer model X556U series, CPU Intel Core i7-6500U up to 3.1 GHz with OS Ubuntu 16.04.

### Data collection and preprocessing

Time series of approximately 8 yr (2010-2017) were collected at the National Agrometeorological Network station (AGROMET; <http://agromet.cl>) belonging to the Chilean Ministry of Agriculture. We used 10 automatic meteorological stations located in the interior valleys of the Maule and Biobío Regions (Figure 1, Table 1); all sensors are Campbell Scientific (Logan, Utah, USA) and details about each model and its accuracy are included in Table 2. The variables were temperature (°C), relative humidity (%), radiation ( $W m^{-2}$ ), wind magnitude ( $m s^{-1}$ ), wind direction (°), and precipitation (mm). In addition, we calculated dew point temperature (°C) as the temperature at which the saturation vapor pressure and actual vapor pressure are equal; it is estimated according to Campbell and Norman (1998) with the following equations:

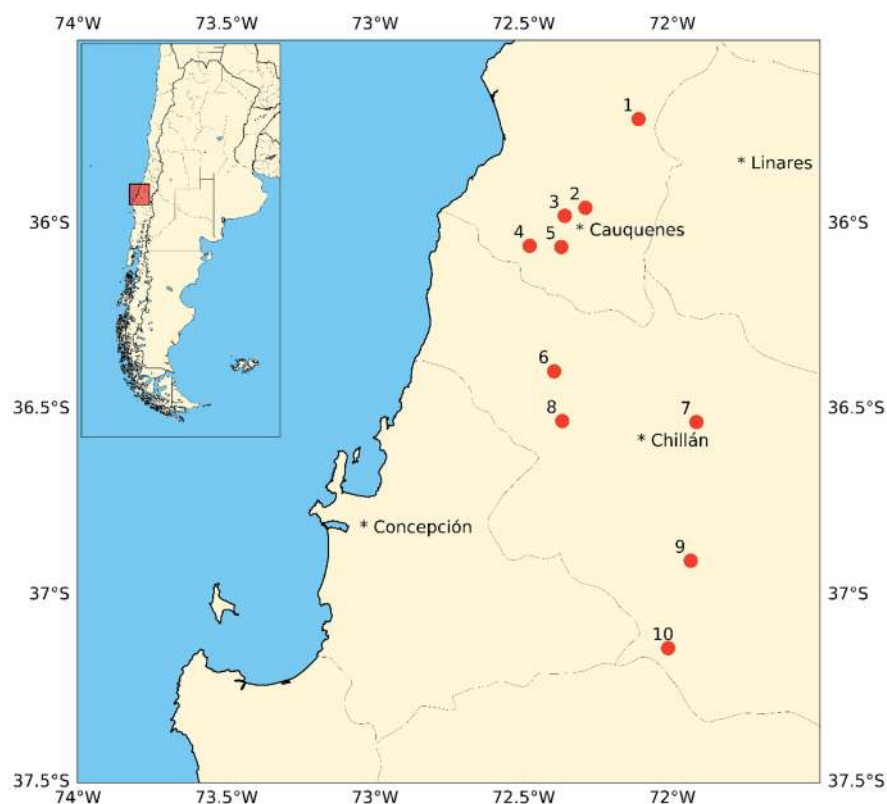
$$e_s = a \exp\left(\frac{bT}{T + c}\right)$$

$$e_a = \frac{RH}{100} e_s$$

$$Td = \frac{c \ln\left(\frac{e_a}{a}\right)}{b - \ln\left(\frac{e_a}{a}\right)}$$

where  $e_s$  is the saturation vapor pressure from temperature and the constants  $a$ ,  $b$ , and  $c$  can be chosen to optimize the fit of the equation for various ranges of data. Our objective was based on environmental biophysics applications for which the constants are  $a = 0.611$  kPa,  $b = 17.502$ , and  $c = 240.97$  (Buck, 1981; Campbell and Norman, 1998) and  $e_a$  is the ambient vapor pressure or actual vapor pressure. For this calculation, we used air temperature ( $T$ ) and relative humidity ( $RH$ ) corresponding to the value at 23:00 h because it is the closest time to the next day that allows better prediction results.

**Figure 1. Map of the study area in central Chile. The red dots are the different stations belonging to the National Agrometeorological Network (AGROMET) used in the artificial neural network.**



**Table 1. Stations used to configure and test the artificial neural network, their corresponding number, geographic position, elevation, and total number of days removed in the filter stage.**

Station number	Station name	Latitude (S)	Longitude (W)	Elevation (m a.s.l.)	Total days (filtered days)
1	Sauzal	-35.7149	-72.1113	162	2623 (115)
2	Cauquenes	-35.9560	-72.2899	162	2213 (525)
3	Santa Sofia	-35.9776	-72.3599	151	2600 (138)
4	Coronel del Maule	-36.0591	-72.4777	172	2570 (168)
5	Los Despachos	-36.0621	-72.3714	167	2594 (144)
6	Ninhue	-36.3981	-72.3952	91	2589 (149)
7	Santa Rosa	-36.5352	-71.9164	194	2287 (451)
8	Portezuelo	-36.5320	-72.3683	109	2431 (307)
9	Navidad	-36.9073	-71.9356	314	2215 (523)
10	Yungay	-37.1414	-72.0118	265	2450 (288)

**Table 2. Meteorological variables used as input on the artificial neural network and sensor accuracy.**

Variables	Abbreviation	Temporary aggregation	Model	Accuracy
Temperature, °C	T	Hourly mean	HMP60 - 1000 ohm P.R. Thermometer	± 0.6 °C
Relative humidity, %	RH	Only 23:00 h	HMP60 - Vaisala's INTERCAP	± 3% (0-90%); ± 5% (90-100%)
Radiation, W m <sup>-2</sup>	Rad	Only 14:00 h	LI200X Silicon Pyranometer	± 5% maximum (± 3% typical)
Wind speed, m s <sup>-1</sup>	Mag	Daily mean	03002-L - Anemometer	± 0.5 m s <sup>-1</sup> (1.1 mph)
Wind direction, °	Dir	Daily mean	03002-L - Anemometer	± 5°
Precipitation, mm	PP	Accumulated daily	TE525MM-L - Rain Gage	± 5°
Dew point, °C	*Td	Only 23:00 h	-	-

\*Td is calculated with T and RH at 23:00 h.

**Table 3. Input matrix for artificial neural network called principal matrix. The three points (...) represent the numerical succession.**

Columns →		1	2	3	4	5	...	20	21	22	23	24	25	26	27	28	29	30		
← Rows	Year	Month	Day	T00	T01	T02	T03	T04	...	T19	T20	T21	T22	T23	Rad	Dir	Mag	PP	Td	Tmin
	2010	5	4	11.2	11.0	10.5	10.3	10.3	...	14.5	13.7	12.3	12.5	13.2	148.6	326.7	2.1	0.0	10.2	7.5
2010	5	5	13.2	13.2	12.8	12.2	6.6	...	11.8	10.9	9.3	8.8	7.5	141.5	348.7	4.3	26.5	6.8	2.9	
2010	5	6	6.9	6.4	5.8	5.2	9.3	...	12.2	11.2	10.2	9.5	8.9	539.9	208.7	3.0	0.1	3.7	6.6	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
2017	7	29	6.8	6.3	6.1	6.0	6.0	...	6.3	5.2	3.9	2.1	2.9	263.6	185.4	0.9	0.0	0.9	-2.4	
2017	7	30	1.3	0.0	-0.6	-0.8	-1.3	...	5.2	3.7	2.9	2.3	1.7	493.6	98.5	0.8	0.0	1.3	-1.9	

We generated an array of 30 columns as input for ANN using the previously mentioned time series of meteorological variables (Table 3). The first 24 columns were for hourly temperatures from 00:00 to 23:00 h (T00:00, T01:00, ..., T23:00) and we then added six columns in the following order: radiation at 14:00 h (Rad), wind direction (Dir) daily mean calculated with the hourly value of the wind vectorial components (zonal and meridional components), wind magnitude (Mag) daily mean, accumulated daily precipitation (PP), dew point temperature at 23:00 h (Td), and next day minimum air temperature (Tmin). On the other hand, rows of each array correspond to the temporal daily evolution from 2010 to 2017 (Table 3). From this point on, this matrix was called “principal matrix”.

Additionally, simple filters were applied to some variables to detect and remove outliers/wrong data as well as sensor errors and extreme values that may occur in the study area that we then replaced with NaN values (i.e., Not-a-Number in programming language), and we applied a linear interpolation to complete the row that contains NaN information from the principal matrix. The filters were mainly based on acceptance thresholds such as a range from -15 to 45 °C and running variance for temperature, 0% to 100% for relative humidity, and 0 to 1300 W m<sup>-2</sup> for radiation. We did not apply filters to the other variables.

### Configuration of backpropagation artificial neural network

Backpropagation is a systematic method of training multilayer ANNs based on a gradient descent method. Error difference is calculated with respect to weights for a given input by propagating the error backwards from the output layer to the hidden layer and further to the input layer; the purpose is to adjust weights according to the error function and minimize the error through a series of iterations (Gill et al., 2010; Singh et al., 2011; Wang et al., 2015). This method is accompanied by a Levenberg-Marquardt (LM) training function, which is the fastest and most stable algorithm developed by Kenneth Levenberg and Donald Marquardt (Marquardt, 1963), who designed it to minimize the sum of squared error function and provide a numerical solution to the problem of minimizing nonlinear functions (Golsorkhi and Tehrani, 2014). The LM algorithm is a very simple but robust method for approximating a function. Basically, it consists in solving the equation:

$$(J^T J + \lambda i) \delta = J^T E$$

where  $J$  is the Jacobian matrix for the system,  $\lambda$  is the Levenberg damping factor that ensures the positiveness of the Hessian,  $i$  is the identity matrix,  $\delta$  is the weight update vector we want to find, and  $E$  is the error vector containing the output errors for each input vector used in training the network. The  $\delta$  tells us by how much we should change our network weights to achieve a possibly better solution. The  $J^T J$  matrix is also known as the approximate Hessian. The  $\lambda$  damping factor is adjusted at each iteration and guides the optimization process. If  $E$  is rapidly reduced, a lower value can be used, thus bringing the algorithm closer to the Gauss-Newton algorithm. However, if an iteration insufficiently reduces the residual,  $\lambda$  can be increased and make it a step closer to the gradient descent direction (Kanzow et al., 2004).

The training process automatically stops when generalization no longer improves, as indicated by an increase in the mean square error (MSE) of the validation process. The MSE is calculated with the mean square difference between the ANN output and the observed value associated with that input (target). The partial derivatives of that error are used in this method to adjust the weights using a gradient descent algorithm (Han et al., 2015) and the lower MSE values are better, while zero means no error (Sapna et al., 2012). A Spearman correlation analysis was also performed to measure the correspondence between outputs and targets, and the  $Rho$  values closer to 1 show a good correlation with a statistical significance at p-value (p-val) < 0.05.

The ANN was configured as a multi-layer ANN arranged in three layers: input, hidden, and output layers. The input layer receives data one case at a time, hidden layers connect the input and output layers, and the output layer is interpreted as the prediction (Shank et al., 2008). Figure 2 shows the general architecture of the three-layer ANN using backpropagation training. The input layer contains the columns described in Table 3, namely 29 input layers with the first 29 columns of the principal matrix. A hidden layer was then used for different experiments or model configurations (Figure 3) containing 1, 5, 10, 15, 20, 25, and 30 neurons. Figure 2 displays an ANN architectural schematic diagram where  $IW\{1,1\}$  is the weight matrix of the order of  $S \times R$  and  $b\{1\}$  the bias vector of the order of  $S \times 1$ . Finally, there is one output layer from the last column of the principal matrix, which is interpreted as the variable to be predicted, where  $LW\{2,1\}$  is the weight array of the order of  $S \times 1$  and  $b\{2\}$  is the output bias value;  $S$  is the number of neurons and  $R$  is the input number. The ANN has a sigmoid transfer function  $f(n) = 2/(1 + \exp(-2n)) - 1$  for the hidden layer, which is faster and easier to derive than hyperbolic functions frequently used in this context, and this function has a tangent activation threshold corresponding to 0. A linear transfer function was used for the output layer. An ANN with that configuration can represent any functional relationship between input and outputs (Hagan et al., 2014).

Our supervised ANN was trained by the backpropagation method, directly on the training set with a data period from 2010 to 2015; its ability to generalize is monitored on the validation set at the same time, and its ability to forecast is measured on the test set for 2016 and 2017. We used several different rates for the training, validation, and test processes (not shown), but we concluded that 80% in training, 10% in validation, and 10% in testing were the best configuration to use because if we over train the ANN (e.g., 90% training, 5% validation, and 5% testing), it is very possible to generate good predictions on the training dataset but not in testing; the selection datasets can be a problem for ANN if the percentages are not balanced (Shank et al., 2008). When these three steps were performed for all the data series of the principal matrix, one iteration occurred. Our network was then configured to make 1000 iterations to find the best performance.

Many statistical indices were used to assess frost predictions, including sensitivity, specificity, precision, accuracy, error rate, and F1-score (Powers, 2011; Sammut and Webb, 2011; Zhu et al., 2014). All these indices need information from the classification confusion matrix (Kohavi and Provost, 1998) to be calculated, which mainly consist of four elements adapted to our model approach: i) true positive (tp) or forecasted frost, ii) false positive (fp) or false frost, iii) false negative (fn) or lost frost, and iv) true negative (tn) or no-frost. According to Liu et al. (2005), Zhu et al. (2014), and Powers (2011), sensitivity (recall), also called true positive rate (TPR), is the proportion of the real positive cases of frost that was correctly predicted and TPR is estimated as  $tp/(tp + fn)$ . Specificity or true negative rate (TNR) is the proportion of the real negative cases of no-frost that was correctly estimated and TNR is estimated as  $tn/(tn + fp)$ . Precision or positive predictive value (PPV) is the proportion of the predicted positive frost that was correctly real positive, and PPV is estimated as  $tp/(tp + fp)$ . Accuracy (ACC) is the proportion of the total number that was correctly estimated and ACC is estimated as  $(tp + tn)/(tp + tn + fp + fn)$ . The misclassification rate or error rate (ER) is the proportion of the false frost and ER is estimated as  $(fp + fn)/(tp + tn + fp + fn)$ . Finally, the F1-score (F1) is the harmonic average of TRP and PPV and F1 is estimated as  $2tp/(2tp + fp + fn)$ . The indices vary from 0 to 1 and when it is close to 0, it indicates poor model results; however, when it is close to 1, it indicates good model results. The abovementioned applies to TPR, TNR, PPV, ACC, and F1, while it is the opposite for ER.

**Figure 2. General architecture of the backpropagation artificial neural network for the proposed model.  $IW\{1,1\}$  is weight matrix of the order of  $S \times R$  and  $b\{1\}$  the bias vector of the order of  $S \times 1$  in the hidden layer with 25 neurons.  $LW\{2,1\}$  is the weight matrix of the order of  $S \times 1$  and  $b\{2\}$  is the output bias value in the output layer with one neuron;  $S$  is the number of neurons and  $R$  is the input number.**

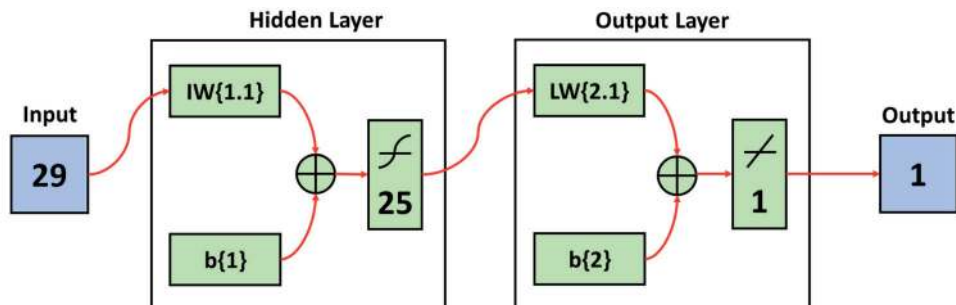
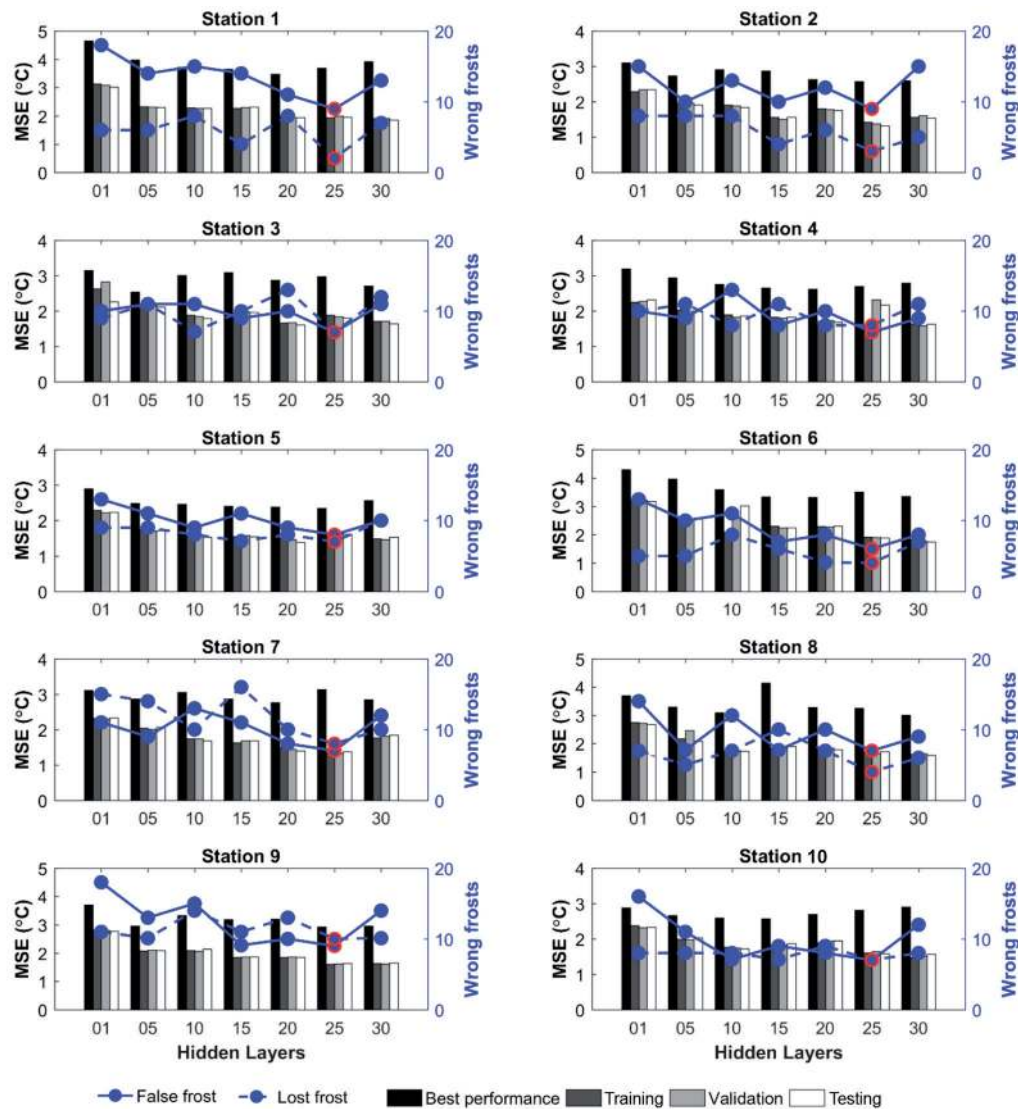


Figure 3. Experiments with variation in the number of neurons in the hidden layer (1, 5, 10, 15, 20, 25, and 30). Bars show mean square error (MSE) resulting from best performance, training, validation, and testing, respectively.



\*Solid blue line is predicted false frost; dotted line is predicted lost frost; red circle is experiment with an overall minimal error in the prediction.

## RESULTS AND DISCUSION

To study the performance of the proposed backpropagation ANN, a set of meteorological time series was obtained from the AGROMET database and it was trained, validated, and tested to predict the following minimum air temperature to detect the possible next day radiative frost events and be able to warn growers to prepare for them. Based on Castañeda-Miranda and Castaño (2017), we developed a multi-layer perceptron neural network using a backpropagation algorithm. The experiments were performed using the principal matrix with air temperature, relative humidity, radiation, wind direction/magnitude, precipitation, and dew point variables as ANN input (Table 3). This same input configuration was performed by Sallis et al. (2009), who worked with data processing and the characteristics of frost prediction in central Chile. They mentioned the importance of each variable in the configuration of optimal conditions for the formation of frost; for example, one of their conclusions was that it is important to have wind data for this region associated with the Los Andes mountain range. Low velocities and mainly east directions bring low temperatures in central valleys, which could contribute to generate frost events.

With a different number of neurons in the hidden layer (1, 5, 10, 15, 20, 25, and 30 neurons), we observed that 25 neurons was usually the most optimal network because the lowest numbers of lost and false frost were exhibited. In addition, the lowest MSE for the best performance, training, validation, and testing processes were also for 25 neurons in the hidden layer (Figure 3). When we used a few neurons in the hidden layer as did Smith et al. (2006), the model ran quite fast compared with applying a large number of neurons. However, as a result of MSE, the false and lost frosts were higher and the correct detection of frost was lower. Thus, larger numbers of neurons in the hidden layer give the network more flexibility because the network has more parameters to optimize (Malik et al., 2014). This is why we prefer our model to take more time to work or train (approximately 30 min) but deliver better frost predictions to promote good network results instead of optimizing time. According to Baboo and Shereef (2011), the number of hidden layers and how many neurons in each hidden layer cannot be well defined in advance and could change for each network configuration and type of data. Nevertheless, the addition of a hidden layer could allow the network to learn more complex patterns.

The ANN test results showed a very good adjustment in estimated (predicted) and known (real) minimum temperature for all stations (Figure 4, Table 4); in general, an MSE value of 2.99 °C was observed for the best performance, specifically MSE of 1.71 °C for training, 1.77 °C for validation, and 1.74 °C for testing processes. The overall mean of the difference between predicted and known minimum temperature was  $\pm 1$  °C with 71% of the data and  $\pm 3$  °C with 96% of the data, which means that the error is usually low in the prediction. This is corroborated with a good Spearman correlation of 0.86 ( $p$ -val < 0.05) as a mean of all stations (Figure 4). In addition, if we only observe differences between erroneous frosts (false and lost frosts), we can see that the overall mean was  $\pm 3$  °C in 78% of the cases.

One of the possible causes of the differences and errors can be the monitoring error of the sensors because each sensor can exhibit an error in accuracy associated with the manufacturing process of the equipment (Table 2). This was in fact proven by simple exercises in which we performed two simulations to the network by adding the maximum and minimum accuracy error for each sensor, respectively. Figure 4 shows shaded areas associated with the main continued line, which represent how the minimum temperature prediction could vary by forcing this error. The deviation of this variability was averaged and resulted in a mean standard deviation of  $\pm 1$  °C ( $\pm 1.07$  °C for estimated and  $\pm 0.88$  °C for known minimum temperature). This means that the final differences between predicted and observed temperature could be forced in part by the accuracy of the meteorological sensors. One case is an experiment at Station 2 (Cauquenes) when we simulated the model with the lowest accuracy values, now called Station 2<sub>min</sub> (Table 4), and model precision improved from 0.85 to 0.90. Another example is Station 5 (Los Despachos) where we simulated the model with the highest accuracy values, now called Station 5<sub>max</sub>, and sensitivity and precision metrics improved from 0.79 to 0.82 and from 0.77 to 0.82, respectively (Table 4).

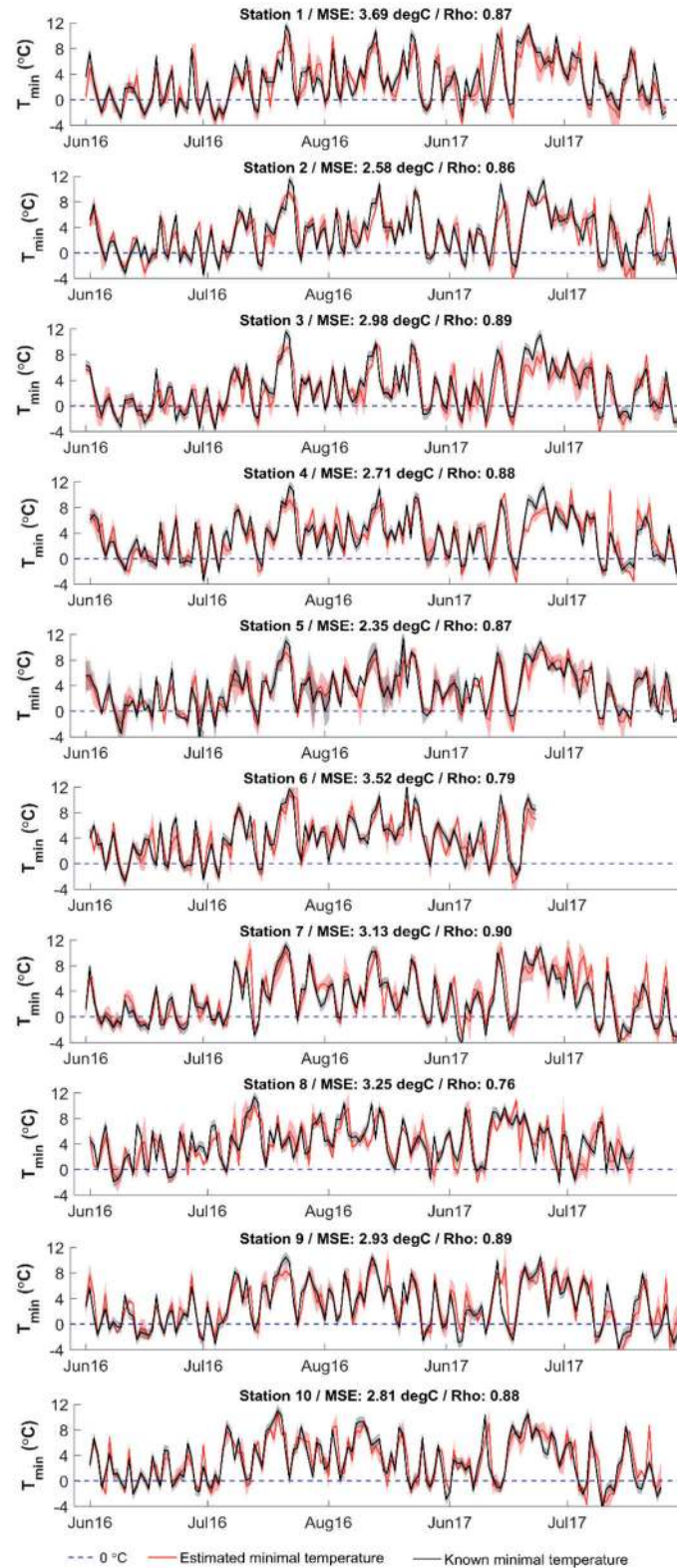
The overall mean for known frost in the study area was 44 from January 2016 to June 2017 where our model successfully predicted 38 frosts as a mean of all stations; however, it presented eight false frost and six lost frost events (Table 4). Each specific value for each station is listed in Table 4.

These results generated good performance in the confusion matrix classification metrics such as 0.86 in the overall mean sensitivity (TRP), 0.82 in overall mean precision (PPV), and 0.84 in overall mean F1-score (F1). Station 1 (Sausal) and Station 2 stand out with higher TPRs of 0.96 and 0.94, respectively; this indicates that out of 53 frost events at Station 1, 51 were predicted and only 2 were lost, and out of 52 frost events at Station 2, 49 were predicted and 3 were lost. On the other hand, the station with the best precision (PPV) was Station 3 (Santa Sofía) with 0.88, 7 false frost and 7 lost frost events. The only stations with an F1-score < 80% precision were Station 5 and Station 8 (Portezuelo), with values of 0.78 and 0.76, respectively; however, these values are not so low.

If we change the inputs in the matrix configuration (principal matrix) in another experiment, some station's statistics metrics should be benefitted. For example, at Station 4 (Coronel del Maule) we used the same matrix structure without wind and precipitation inputs, now called Station 4<sub>b</sub>; results were better in that sensitivity improved from 0.74 to 0.90 and reduced false frosts from 8 to 4 (Table 4). Another example was Station 9 (Navidad) where we used the original matrix but without wind magnitude, now called Station 9<sub>b</sub>; sensitivity and precision improved from 0.80 to 0.86 and from 0.81 to 0.82, respectively, thus reducing false frosts from 10 to 7. The aforementioned raises the question of what would happen if a specific configuration were performed for each station. There would then be a great possibility of improving the resulting forecasts; in fact, Ghielmi and Eccel (2006) observed the same situation in a similar valley area associated with a mountain range where their results varied between stations even though stations were relatively close to one another and associated this with typical conditions of each particular place. This could be because each area associated with the



**Figure 4. Time series of estimated (red line) and known (black line) minimum air temperatures for June-August 2016 and June-July 2017. The shaded areas associated with each line represent the two simulations performed on the network by adding the maximum and minimum accuracy error for each meteorological sensor. MSE is the mean square error and Rho is the Spearman correlation statistic and all values have p-value < 0.05.**



**Table 4. Values of mean square error (MSE) and values and metrics associated with confusion matrix classification per station.**

Stations	General mean	1	2	3	4	5	6	7	8	9	10	2 <sub>min</sub>	5 <sub>max</sub>	4 <sub>b</sub>	9 <sub>b</sub>
Mean square error (MSE)															
Best performance	2.99	3.69	2.58	2.98	2.71	2.35	3.52	3.13	3.16	2.93	2.81	2.58	2.29	2.73	2.96
Training	1.71	1.94	1.42	1.89	1.62	1.65	2.26	1.35	1.72	1.60	1.61	1.45	1.48	1.77	1.84
Validation	1.77	2.00	1.38	1.85	2.32	1.62	2.20	1.33	1.71	1.62	1.64	1.43	1.54	1.74	1.89
Testing	1.74	1.96	1.33	1.80	2.18	1.62	2.26	1.37	1.65	1.63	1.59	1.47	1.55	1.80	1.95
Confusion matrix classification															
Total known frost	44	53	52	58	39	34	28	60	21	49	43	63	22	39	49
True positive (tp), forecasted frost	38	51	49	51	31	27	24	53	17	39	36	57	18	35	43
False positive (fp), false frost	8	9	9	7	7	8	6	7	7	9	7	9	4	7	9
False negative (fn), lost frost	6	2	3	7	8	7	4	7	4	10	7	6	4	4	6
True negative (tn), no frost	519	512	516	512	531	535	506	510	528	518	521	505	551	531	518
Confusion matrix classification metrics															
True positive rate (TPR)*	0.86	0.96	0.94	0.88	0.79	0.79	0.86	0.88	0.81	0.80	0.84	0.86	0.82	0.90	0.88
True negative rate (TNR)**	0.99	0.98	0.98	0.99	0.99	0.99	0.99	0.99	0.99	0.98	0.99	0.99	0.99	0.99	0.98
Positive predictive value (PPV)***	0.82	0.85	0.84	0.88	0.82	0.77	0.80	0.88	0.71	0.81	0.84	0.90	0.82	0.83	0.83
Accuracy (ACC)	0.98	0.98	0.98	0.98	0.97	0.97	0.98	0.98	0.98	0.97	0.98	0.97	0.99	0.98	0.97
Error Rate (ER)****	0.02	0.02	0.02	0.02	0.03	0.03	0.02	0.02	0.02	0.03	0.02	0.03	0.01	0.02	0.03
F1-score (F1)	0.84	0.90	0.89	0.88	0.81	0.78	0.83	0.88	0.76	0.80	0.84	0.88	0.82	0.86	0.85

\*Sensitivity or recall, \*\*Specificity, \*\*\*Precision, \*\*\*\*Misclassification rate.

stations has different geographical characteristics (orography and terrain morphology) and local climatic conditions, which might be more important than a generalized condition for the total area in frost formation such as elevation, inclination, woods, or nearby hills.

Another possible explanation for differences can be attributed to a number of input variables and their total time to generate the input matrix. This means that if we add another important variable, which is not accounted for in the present study (e.g., cloud cover or atmospheric pressure), then it could add information to be learned, or if we add more years of data to the training process, it could learn from these other data and improve considerably its performance. This same idea was worked out by Bonilla et al. (2006), who mention that the model's efficiency could increase if a greater number of meteorological variables involved in the ANN training process were used as well as a period close to or > 20 yr of data.

## CONCLUSIONS

This paper proposes a simple technique of minimum temperature forecasting to predict frost by using backpropagation artificial neural networks. The results show good performance in frost detection and reveal a low associated error in the difference between estimates and known minimum temperatures. This network is capable of providing good results and can be considered as an alternative to traditional complex meteorological approaches because it can be an important tool to help generate an early warning system for growers.

The differences and errors in frost detection can be attributed to several factors that are mainly associated with the accuracy of the sensor measurement of the meteorological stations, local climatic and geographic conditions, and the number of parameters that enter in the ANN training processes as well as other factors such as the time series data. Performance can be improved if the networks are made with specific configurations for each location and with the configuration of a more complex network.

## ACKNOWLEDGEMENTS

This work was carried out within the framework of the project "Validación y ampliación de la cobertura del sistema de monitoreo y alerta temprana de heladas como servicio de información de valor agregado de la Red Agroclimatólogica Nacional" and funded by a grant from FIA Chile.

## REFERENCES

- AMS. 2000. Glossary of meteorology. 2<sup>nd</sup> ed. 850 p. Glickman, T.S. (ed.) American Meteorological Society, Boston, Massachusetts, USA.
- Ashour, M.A., ElZahaby, S.A., and Abdalla, M.I. 2016. Backpropagation neural network approach for mean temperature prediction. *International Journal of Recent Research and Applied Studies* 29:12-18.
- Baboo, S.S., and Shereef, I.K. 2011. An efficient weather forecasting system using artificial neural network. *International Journal of Environmental Science and Development* 2(1):49-54. doi:org/10.7763/IJESD.2011.V2.96.
- Bocco, M., Willington, E., and Arias, M. 2010. Comparison of regression and neural networks models to estimate solar radiation. *Chilean Journal of Agricultural Research* 70:428-435.
- Bonilla, J.E., Ramírez, J., y Ramírez, O. 2006. Metodología para el diseño de un modelo univariado de red neuronal para el pronóstico de la temperatura mínima en la zona de Mosquera (Cundinamarca, Colombia). *Meteorología Colombiana* 10:111-120.
- Buck, A.L. 1981. New equations for computing vapor pressure and enhancement factor. *Journal of Applied Meteorology* 20:1527-1532.
- Castañeda-Miranda, A., and Castaño, V.M. 2017. Smart frost control in greenhouses by neural networks models. *Computers and Electronics in Agriculture* 137:102-114.
- Campbell, G.S., and Norman, J.M. 1998. An introduction to environmental biophysics. 2<sup>nd</sup> ed. Springer Science & Business Media, New York, USA.
- Chevalier, R.F. 2008. Air temperature prediction using support vector regression and GENIE: The Georgia extreme-weather neural network informed expert. 72 p. MSc Thesis. University of Georgia, Athens, Georgia.
- Cunha, J.M. 1952. Contribuição para o estudo do problema das geadas em Portugal. Relatório final do Curso de Engenheiro Agrônomo. Instituto Superior de Agronomia, Lisboa, Portugal.
- FAO. 2010. Protección contra las heladas: fundamentos, práctica y economía. Volumen 1. Snyder, R.L., Melo-Abreu, J.P., and Villar-Mir J.M. (eds.) Serie FAO Sobre el Medioambiente y la Gestión de los Recursos Naturales. Nr 10. 257 p. FAO, Roma, Italia.
- Fritsch, S., Guenther, F., Suling, M., and Mueller, S.M. 2016. Package 'neuralnet', training of neural networks. Available at <https://cran.r-project.org/web/packages/neuralnet/neuralnet.pdf> (accessed November 2017).
- Ghielmi, L., and Eccel, E. 2006. Descriptive models and artificial neural networks for spring frost prediction in an agricultural mountain area. *Computers and Electronics in Agriculture* 54:101-114. doi:10.1016/j.compag.2006.09.001.
- Gill, E.J., Singh, E.B., and Singh, E.S. 2010. Training back propagation neural networks with genetic algorithm for weather forecasting. p. 465-469. IEEE 8<sup>th</sup> International Symposium on Intelligent Systems and Informatics, Subotica, Serbia. 10-11 September. Institute of Electrical and Electronics Engineers (IEEE), Piscataway, New Jersey, USA. doi:10.1109/SISY.2010.5647319.
- Golsorkhi, N.A., and Tehrani, H.A. 2014. Levenberg-Marquardt method for solving the inverse heat transfer problems. *Journal of Mathematics and Computer Science* 13:300-310.
- González-Camacho, J.M., Ojeda-Bustamante, W., y López-Cruz, I. 2008. Predicción de la evapotranspiración de referencia mediante redes neuronales artificiales. *Ingeniería Hidráulica en México* 23(1):127-138.
- Hagan, M.T., Demuth, H.B., Beale, M.H., and De Jesús, O. 2014. Neural network design. 2<sup>nd</sup> ed. Hagan, M. (ed.) PWS, Boston, Massachusetts, USA.
- Han, S., Pool, J., Tran, J., and Dally, W.J. 2015. Learning both weights and connections for efficient neural networks. p. 1135-1143. In Proceedings of the 28<sup>th</sup> International Conference on Neural Information Processing Systems - Volume 1, NIPS'15. MIT Press, Cambridge, Massachusetts, USA.
- Jacob, R., and Murugan, R. 2016. Implementation of neural network based for PID controller. p. 2769-2771. In International Conference on Electrical, Electronics, and Optimization Techniques, Chennai, Tamil Nadu. 3-5 March. DMI College of Engineering, Tamil Nadu, India.
- Kanzow, C., Yamshita, N., and Fukushima, M. 2004. Levenberg-Marquardt methods with strong local convergence properties for solving nonlinear equations with convex constraints. *Journal of Computational and Applied Mathematics* 172:375-397.
- Kohavi, R., and Provost, F. 1998. On applied research in machine learning. In Editorial for the Special Issue on Applications of Machine Learning and the Knowledge Discovery Process. Volume 30. Columbia University, New York, USA.
- Liu, C., Berry, P.M., Dawson, T.P., and Pearson, R.G. 2005. Selecting thresholds of occurrence in the prediction of species distributions. *Ecography* 28:385-393. doi:10.1111/j.0906-7590.2005.03957.x.
- Lungu, M., Panaitescu, L., Albu, A., Cracu, G., and Nita, S. 2010. Frost and thaw - Climatic risk to crops in Southern Dobrudja. *Research Journal of Agricultural Science* 42:673-677.
- Malik, P., Singh, S., and Arora, B. 2014. An effective weather forecasting using neural network. *International Journal of Emerging Engineering Research and Technology* 2:209-212.
- Marquardt, D.W. 1963. An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics* 11:431-441.

- ODEPA. 2013. Efecto heladas de septiembre en frutales y hortalizas entre la Región de Coquimbo y la del Maule. Oficina de Estudios y Políticas Agrarias (ODEPA), Santiago, Chile. Available at [http://www.odepa.cl/wp-content/files\\_mf/1388163908estudioDanoHeladas.pdf](http://www.odepa.cl/wp-content/files_mf/1388163908estudioDanoHeladas.pdf) (accessed May 2017).
- Ovando, G., Bocco, M., y Sayago, S. 2005. Redes neuronales para modelar predicción de heladas. *Agricultura Técnica* 65:65-73.
- Powers, D.M.W. 2011. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. *Journal of Machine Learning Technologies* 2(1):37-63.
- Retamal-Salgado, J., Vásquez, R., Fischer, S., Hirzel, J., and Zapata, N. 2017. Decrease in artificial radiation with netting reduces stress and improves rabbit-eye blueberry (*Vaccinium virgatum* Aiton) Ochlockonee productivity. *Chilean Journal of Agricultural Research* 77:226-233. doi:10.4067/S0718-58392017000300226.
- Sallis, P., Jarur, M., and Trujillo, M. 2009. Frost prediction characteristics and classification using computational neural networks. p. 1211-1220. Springer-Verlag, Berlin, Heidelberg, Germany.
- Sammut, C., and Webb, G.I. 2011. *Encyclopedia of machine learning*. Springer, New York, USA. doi:10.1007/978-0-387-30164-8.
- Sapna, S., Tamilarasi, A., and Kumar, M. 2012. Backpropagation learning algorithm based on Levenberg Marquardt algorithm. *Computer Science and Information Technology*. p. 393-398. doi:10.5121/csit.2012.2438.
- Shank, D.B., Hoogenboom, G., and McClendon, R.W. 2008. Dew point temperature prediction using artificial neural networks. *Journal of Applied Meteorology and Climatology* 47:1757-1769. doi:10.1175/2007JAMC1693.1.
- Singh, S., Bhambri, P., and Gill, J. 2011. Time series based temperature prediction using back propagation with genetic algorithm technique. *International Journal of Computer Science Issues* 8:28-32.
- Smith, B.A., McClendon, R.W., and Hoogenboom, G. 2006. Improving air temperature prediction with artificial neural networks. *International Journal of Computational Intelligence* 3(3):179-186.
- Valipour, M. 2016. Optimization of neural networks for precipitation analysis in a humid region to detect drought and wet year alarms. *Meteorological Applications* 23(1):91-100.
- Venkatesan, K., and Ramachandraiah, U. 2018. Climate responsive control using artificial neural networks. *Journal of Building Engineering* 19:191-204.
- Wang, L., Zeng, Y., and Chen, T. 2015. Back propagation neural network with adaptive differential evolution algorithm for time series forecasting. *Expert Systems with Applications* 42(2):855-863. <http://dx.doi.org/10.1016/j.eswa.2014.08.018>.
- Woods, M., and Gondal, I. 2007. A procedural approach to forecasting nocturnal air temperature for frost prediction. p. 412-417. 6<sup>th</sup> IEEE/ACIS International Conference on Computer and Information Science, Melbourne, Queensland. Institute of Electrical and Electronics Engineers (IEEE), Piscataway, New Jersey/International Association for Computer and Information Science (ACIS), Mount Pleasant, Michigan, USA. doi:10.1109/ICIS.2007.42.
- Zhu, L., Cao, Z.G., Zhuo, W., Yan, R., and Ma, S. 2014. A new dew and frost detection sensor based on computer vision. *Journal of Atmospheric and Oceanic Technology* 31:2692-2712.