18th World IMACS / MODSIM Congress, Cairns, Australia 13-17 July 2009
http://mssanz.org.au/modsim09

Frost Prediction using a Combinational Model of Supervised and Unsupervised Neural Networks for Crop Management in Vineyards

Sallis, P.[1], **M. Jarur**[2], **M. Trujillo**[2] **and A Ghobakhlou**[1]

[1]*Geoinformatics Research Centre, School of Computing and Mathematical Sciences, Auckland University of Technology, New Zealand.* (*www.geo-informatics.org)*
[2]*Escuela de Computación e Informática, Facultad de Ingeniería, Universidad Católica del Maule, Chile*
E-mail: philip.sallis@aut.ac.nz

Frost prediction models could contribute significantly towards the successful growth and production of quality crop yield in horticulture, especially in precision viticulture where the benefits are significant because frost damage is well-known for its potential leading to total harvest failure, with a follow-on regional or national economic impact outcome. This reality has increased interest among scientists and growers to advance their knowledge in relation to the inter-dependencies and possible correlations between meso-micro climate variables and associated plant and soil condition values. Included in this array of variables there are also site specific environmental factors such as pesticide saturation and carbon density. Recent interest in building computational models is focused on predicting frost events using both regional (metrological) and vineyard weather monitoring data gathered via remote low-cost sensors, in addition to vineyard-specific environmental attribute data. This paper outlines some earlier research and then describes the climate and atmospheric data analysis, together with vineyard elevation and wind data in order to determine the inter-dependencies of variable values that inform enhanced frost protection measures. Developing a model with supervised and unsupervised neural networks as a means for characterising the data being analysed, is the focus of the investigation described here, which is part of a wider project called Eno-Humanas. This wider project incorporates the modelling of interrelations between grape wine sensory properties, such as aroma, colour and taste, and climate in addition to environmental factors for crop management and prediction purposes. The Eno-Humanas prediction models are intended to contribute to the body of knowledge aimed at finding grape varieties that best suit different potential future climate scenarios.

Keywords: *Frost Prediction, Neural Networks and Vineyards*

1. INTRODUCTION

Previous work by the authors (Sallis et al, 2008) described the characteristics of frost occurrences and impact potential for grape crop management. It also described a classification method using computational neural networks for predicting frost with a limited variable set. The method employed self-organising map (SOM) techniques to classify weather data in order to develop a computational multilayer perceptron (MLP) model for frost prediction in vineyards. It is proposed that similar algorithms could also be developed for irrigation management.

In the 2008 paper, a wider project referred to as Eno-Humanas (www.geo-informatics.org/Eno-Humanas.aspx) was outlined with its goals being the measurement of micro-climate variations over time in nine geo-spatial locations across five countries in both the Northern and Southern Hemispheres, together with a longevity study of these conditions and their inter-dependence and possible correlation with other atmospheric and environmental factors using data gathered from plants and soil. The analysis of grape variety, growing conditions and expert taster opinions of wine quality for each of these geo-spatial locations also forms part of the study. Published results from this work can also be found at the website cited above.

Not only did the 2008 paper describe a computational model for frost prediction, it also illustrated some alternatives for data dependency depiction and result visualisation. These included coloured weight dependent clusters to illustrate variable data distribution and the relative inter-dependence of values across the data set. Conventional line graphs and histograms were also used but three-dimensional (splat) diagrams to show the tree structures that exist in the relationship matrix were given and proposed as an alternative and meaningful visualisation of the results.

The data used in the 2008 study was gleaned from published Metrological Service sources in Central Chile where the wind direction and thus, a dominant factor in frost fall is from the South-East and is influenced by the mountain range named Los Andes. This study focussed on the relationship of the variables in the hours of the greatest cold (dew point just prior the frost fall) but concluded that a filtering of the data for individual instances over a 24 hour period prior to frost fall would provide a more robust forecasting basis for establishing weather-related dependencies alongside other environmental and growing condition factors such as cloud cover, leaf wetness and soil temperature. Other research in this problem domain conducted in Chile is described for instance in Ovando et al (2005). Related frost prediction research using neural networks can be found in Temeyer et al (2003). Forecasting models for frost fall and the parameters around prediction are also numerous as discussed in (Prahba et al, 2008: Snyder et al, 2005).

2. MICRO CLIMATES

It is generally held that numerous microclimates exist in a single vineyard to the extent that growing conditions can vary for one variety or another with some dramatic variances in grape quality (Trought et al, 1999). This is especially true if frost falls only in one part of the vineyard and not another, which can occur. The term ***Terroir*** means that special combination of soil composition and climate which provides distinction to a given viticultural region**.** In our work, we are interested in the variations within a region and indeed, within a vineyard to observe the dynamics of terrior for any given grape variety in any given geo-spatial location. This heterogeneous approach means that we have to deal with varying data types and values that are scaled and calibrated differently one from another. Comparing micro climate data with macro climate data is a significant measure and so is the so-called 'mesa climate' where particular climate and soil conditions can be seen as experiencing extremes of heat, cold and dry because of their exposed geospatial location Ghielmi et al, (2006) The data we have used for the results reported in this paper have been generated from sensor array collection points in what we regard as micro-climate locations.

3. COMPUTATIONAL MODELS

The MLP computational neural network method is perhaps the most used approach to data analysis that employs these modelling techniques. This feed forward network model has three layers (input, hidden and output) expressed as Input (x1...xp) ; Hidden $y_{1..n}$=F_{hidd}{wji*x1…p}; and Output $z_{1..m}$=F_{out}{wkj* $y_{1..n}$}. Individual Inputs populate the Hidden Layer Nodes that are multiplied by a weight and the resulting Hidden Layer Values populate the Output Layer that are then multiplied by a weight and the resulting values are added to produce the values for the final transfer (z) functions of the network (Figure 1), where F_{hidd}{} and F_{out}{} are the transfer function of hidden and out layer respectively.

This method enables translation of the dependent variables into a form that can be used to model, in our case, the combination of factors evident in the influence of frost condition prediction. Neural Network solutions for this problem domain have been long sought and often written about (Paras, et al. 2007; Robinson and Mort, 1996). Probabilistic (PNN), regression (GRNN) and cascade neural network models are also based on algorithms that may provide results of the precision needed for this kind of forecasting. Cascade Neural Networks require only an input and output layer to be defined, which obviates the needs for continuous sensitivity or 'network tuning' because the Hidden Layer is populated 'automatically' by the algorithm. They are self-organising and require little or no supervision. Probabilistic networks utilise the statistical concept of categorical variables, so after clustering a new variable type is generated, which might be useful for frost prediction where humidity for example, is dependent on both temperature and atmospheric moisture thus creating a new class of variable following the analysis of the two data stream variables. In the regression network model, a continuous target variable is evident so the analysis is more linear. Again, either of these methods requires less supervision than the MLP algorithm.

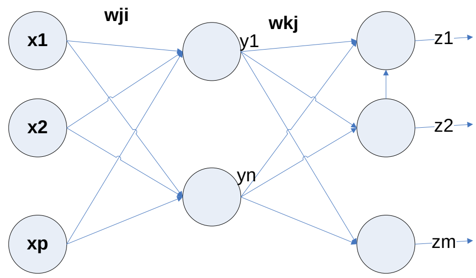


Figure 1. MLP, where 'p' is number of inputs, 'n' number of neurons in hidden layer and 'm' is the number of output neurons. Wji is weight of the connection between input and hidden nodes j and I respectively. Similarly wkj is weight of the connection between hidden and output nodes k and node i respectively

For the research reported in this paper, which extends the earlier 2008 work, we have developed a combinational model, so called because it adopts the concept of Combinatorial Mathematics where multiplication implies the blending of two or more values to create a new and distinctly different value from its input values and yet retains all of their characteristics. It is considered useful here to view the concept this way because of the intrinsic differences in the nature of supervised and unsupervised networks. Fixed weight networks are inappropriate for the data types and kinds of values evident in the frost prediction problem because of the variability of the data over time. Rather than detail the differences in the methods here, which can readily be found in numerous published sources (Kasabov and Kozma 1999) for example, the method we have employed is described below with the results from recent experiments.

4. THE FROST PREDICTION MODEL

4.1. Data Analysis

The database contains values for the following five variables: Temperature (ºC), Humidity (%), Dew Point (ºC), Wind Velocity (Km/hr.) and Wind Direction (º). These variables contain values for occurrences recorded between the months of May and October 2007 for the O'Higgins Region in Chile. This six month period represents the Southern Hemisphere Winter with Autumn and Spring influences. The variable set is generally held to be the principal elements for determination of frost fall. Configuring the neural network is an iterative process that requires an initial set of variable values being used to train it such that it learns the shape of the curve for say temperature compared with humidity as a baseline that can then be used for prediction as new values are ingested over time. The data set consists of a total of 147 records, which are divided into two subsets so that the combined training set has 115 entries (82 No Frost and 33 Frost datums) and the validation set has 32 records (19 without Frost and 13 with Frost data).

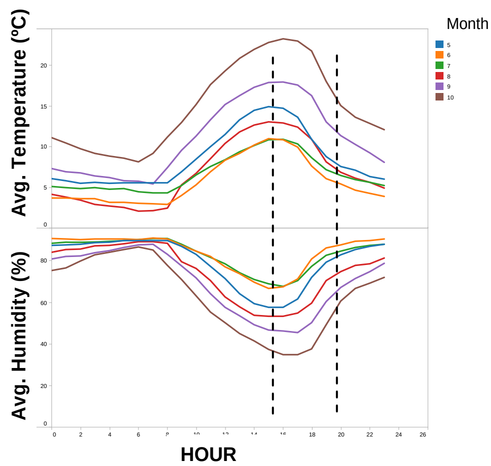


Figure 2. Temperature and Humidity average for every hour in each month.

A two stage process was used to give a priority to the variables and then through visualization and depiction of the network's behaviour, to time step points during the hours of each day.

The Figure 2 depicts a time series of events recorded for average temperature and humidity for every hour over each month in the sample. The horizontal axis presents the hours of the day (0 – 24 hrs.), and vertical axis are the average temperature (Celsius grade) and average humidity (%) respectively.

In Figure 2 it is possible to appreciate the principal declining tendency from 16:00 hrs to 20:00 hrs, when at this time for crop management, it is necessary to begin remedial frost protection activity. We see here the tendency for temperature decline occurring over two days. Using data sampled over a numbers of days prior to a recorded frost fall event, it is possible to predict an upcoming event using the values gathered for each of the variables in the set. Such a depiction can be seen in Figure 3, where horizontal axis represents the dates (4 July at 14 July) and vertical axes are average temperature (ºC), average humidity (%) and dew point (ºC) respectively.

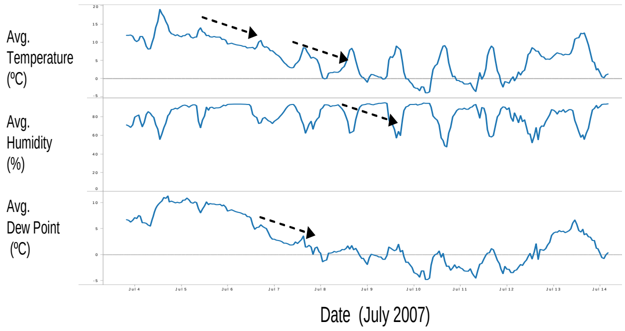


Figure 3. Temperature and Humidity average for every hour in each month.

The initial network training for September data consisted of two cases for each variable. That is, for each instance (i) we had Variables Today (Vt), Variables Yesterday (V) at 16:00 and 20:00 hrs, producing a set for (ii) Variable Today (Vt) at 16:00 and 20:00 hrs. The results were not satisfactory, the models produced had with errors exceeding 46%.

The next step was to decrease the sampling interval for all the variables to 1 hour and increase the range to be considered, so that included variable values from the 16 hr and 21 hr time period. We also took into account the presence of tendency characteristics for this time series through a different sample frequency. Based on this analysis we considered both the supervised and unsupervised training set. We did this by developing a sequence for the Variable Today at 1 hr intervals for the period 16:00 hrs to 21:00 hrs given as V_16, V_17, ..., V_21 hrs.

This means a greater we can observe a greater complexity in the model with a possible total of 30 entries (5 variables by 6 hours). At this level we need to incorporate other analytical tools in order to simplify the network configuration. This is carried out after the initial iterations of the data through the complex model.

An algorithm often used for this next level of analysis to simply (refine) the model is called ID3 (Gillies, 1996). It is used for building decision trees (i.e., classifiers), which then allows for the creation of a very simple classifier. This classifier enables us to implicitly determine the levels of variable value expansion, the effect of which is the creation of a precedence or ranking of the relative importance of each variable. We expect that this mechanism will assist us to determine the critical variable factors needed for reliable prediction criteria. Like any classifier, it requires a set of training and validation data, which have two cases (a) tree creation with all data present (training and validation sets combined) and (b) tree creation with only the training data and then validate being by the resulting classifier. To do this we considered the following two cases: Case 1: Dew Point at 21 hr, Temperature Wind Direction at 16 hrs at 18 hrs. and Case 2: Dew Point at 20 hr, 21 hrs at Dew Point, Temperature at 16 hrs.

5. THE MODELS AND RESULTS

Initially the network setup was based on an MLP configuration, according to the entries mentioned in the section analysis, which can be seen in Figure 4 for case 1 and Figure 5 for case 2.

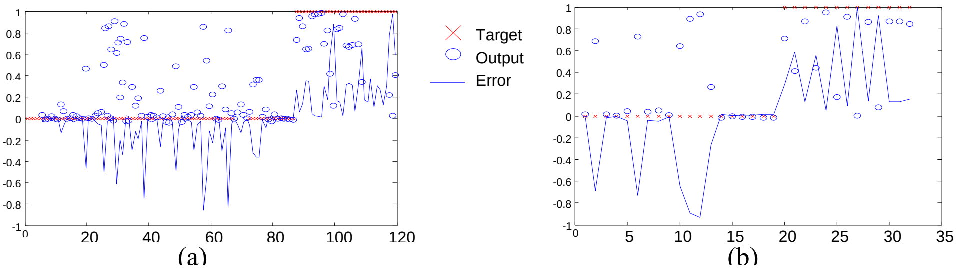


Figure 4. MLP with Dew Point at 21 hr, Temperature at 16 hrs Wind Direction at 18 hrs. (a) Training process and (b) validation process.

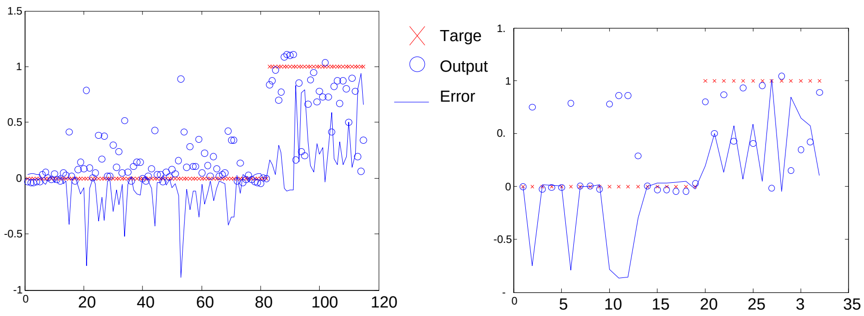


Figure 5. MLP with Dew Point at 20 hr, Dew Point at 21 hrs, Temperature at 16 hrs.

Due to the high level of problem complexity and considering analysing the previous results with an unsupervised (MLP like) neural network, we decided to develop a model using a Radial Basis Function (RBF) (Haykin 2009), which enables a hybrid net to be developed with a mix of supervised and unsupervised training. In this way, a hidden layer that works in a similar sense to how SOMs process data is present but the output layer has a linear function that serves as a combinatorial smoothing device, which is suitable for the blend of variable values being view as belonging to a single data set for every frost fall event in this problem domain.. The literature indicates that as with MLP methods, the RBF networks are multi-layer that can be used for function approximation and classification, but they work in a significantly different way. In RBF networks, the activation of hidden units is based on the distance between the input vector and a prototype vector. This difference can be seen illustrated in Figure 6.

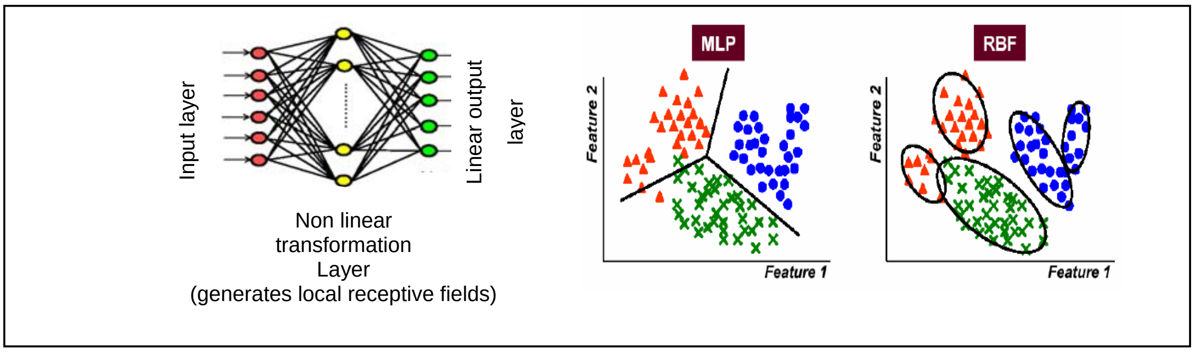


Figure 6. Radial Basis Functions nets, (a) net structure, (b,c) MLP domain recognition v/s RBF domain recognition.

The hidden layer has a 'radial basis' (Gaussian) function, which is expressed as,

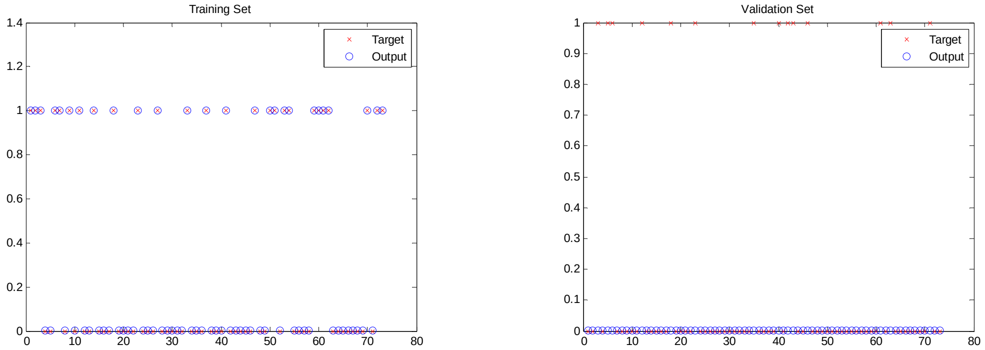
$$\varphi\left(net_j = \left\|x - u_j\right\|\right) = e^{-\left(\frac{\left\|x-u_j\right\|}{\sigma}\right)^2}$$

Where, ‖ … ‖ is the Euclidean norm. The σ is "spread constant", and affect directly the exact or approximated interpolation. The exact interpolation is rarely used due to the possibility that noise may be present in the data and it is necessary to remove it in order to refine the general applicability of the net.

Considering these parameters, we conducted experiments with the data gathered for our variable set over the time period gathered, in order to test the robustness of this blended approach using supervised and unsupervised techniques, The results of these experiments produced results with a low spread and a wide spread as described below.

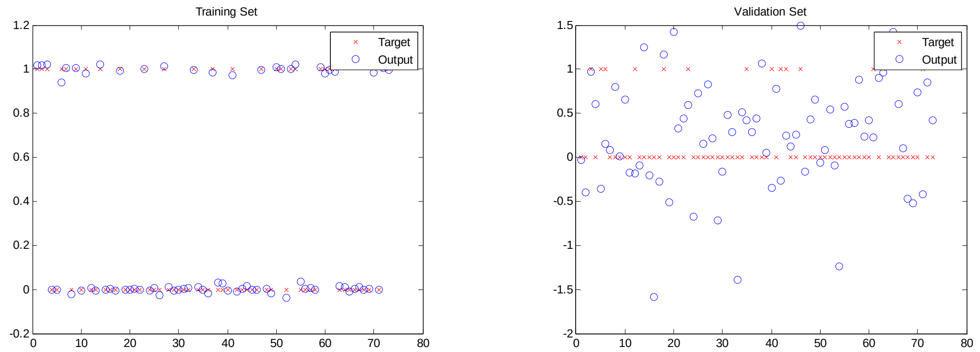
Prior to testing the data set with similar settings as used with the MLP algorithm, we carried out some preliminary tests.

The values used were temperature, humidity and dew point for 16 hrs and 20 hrs in the prediction day and the previous day of prediction and one output denoted by 1 for a 'frost event' and 0 for a 'no frost' event. The results illustrated in Figure 7 show the case when we use a small 'spread' and we obtain a set error for training that is 0, which means that over-fitting has occurred and we have a problem with the validation test and therefore, where the frost prediction (target with value of 1) is wrong. The figure 8 presents a same processed data set with a large spread. Here we can observe a radically clearer dispersion of instance occurrences resulting from the algorithm implementation of that net.



(a) (b)

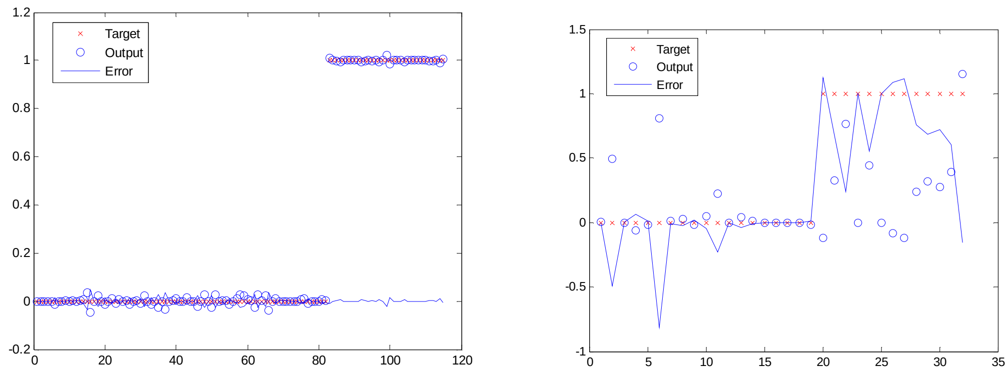
Figure 7. Results of RBF with small spread**.** (a) Validation set and (b) Validation set.



(a) (b)

Figure 8. Results of RBF with large spread. (a) Validation set and (b) Validation set.

Finally, we tested case 2 (Dew Point at 20 hr, Dew Point at 21 hrs, Temperature at 16 hrs) with RBF and this produced better results (Figure 9).



(a) (b)

Figure 9. RBF with Dew Point at 20 hr, Dew Point at 21 hrs, Temperature at 16 hrs. (a) Validation set and (b) Validation set.

6. CONCLUSIONS

While the results differ and can be seen to have improved in precision using the RBF in a supervised mix with an unsupervised network, it is also obvious that the sample frequency is fundamental for the improvement of such a model and its prediction potential. Similarly, it is necessary to obtain a large amount of time-dependent data for each of the variables in the set so that greater accuracy of prediction can be obtained. We propose however, that the methodology employed in this research is appropriate for the frost fall prediction problem domain and can potentially produce robust models using these techniques that could assist greatly the crop management scenarios needed by growers to avoid damage caused by frost. Our work continues as we obtain real-time data from nine climate monitoring stations in situ for five countries and continue to experiment with these prediction methods.

The techniques although simple in terms of generating a classifier system, such as, a decision tree, do translate into a tool that can be used to establish a robust preliminary analysis of the variables. The tasks ahead, while seeking to refine the methodology will be to increase the volume of data submitted to the network for analysis and to have more frequent sampling that will incorporate at least 48 previous data sets to determine the optimal frequency for the prediction robustness of the network type.

REFERENCES

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, pp. 101-114 Elseiver (2006).

Gillies, D. , (1996) Artificial intelligence and scientific method, Ed. Oxford University Press, 1996

Haykin, Simon S. (2009). Neural networks and learning machines. New York: Prentice Hall (2009).

Kasabov, N and Kozma, R. (1999) *Neuro-Fuzzy Techniques for Intelligent Information Systems*. Physica-Verlag, Heidelberg, 1999. ISBN 3-7908-1187-4.

Ovando, G., Bocco, M., Sayago, S. Neural Networks for Modeling Frost Prediction, Agric. Téc., vol.65, no.1, p.65-73. Chile, (2005)

Paras, et al. (2007) A Feature Based Neural Network Model for Weather Forecasting, Proceedings of World Academy of Science, Engineering and Technology, Vol. 23, (2007)

Prabha, T., Hoogenboom, G. Evaluation of Weather Research and Forecasting model for two Frost Events, Computers and Electronics in Agriculture 64, pp. 234-247, Elseiver (2008).

Robinson, C., and Mort, N. A (1996) Neural Network Solution to the Problem of Frost Prediction, UKACC International Conference on Control. Control '96, pp 136 -139, UK (1996)

Sallis, P., Jarur, M., and Trujillo, M., (2008) Frost Prediction Characteristics and Classification using Computational Neural Networks in Australian Journal of Intelligent Information Processing Systems (AJIIPS) volume 10.1, 2008 (ISSN 1321-2133) pp50-58.

Snyder, R.L., de Melo-Abreu, J.P., Matulich, S. (2005) Frost Protection: Fundamentals, Practice, and Economics. Vol.1/2. FAO Environmental and Natural Resources Series, 10, Rome, Italy.

Temeyer, B.R. et al. (2003) Using Artificial Neural Network to Predict Parameters for Frost Deposition on Iowa Bridgeways. Proceedings of the 2003 Mid-Continent Transportation Research Symposium, Iowa

Trought, M.C.T, Howell, G.S., Cherry, N. (1999) Practical Considerations for Reducing Frost Damage in Vineyards. Report to New Zealand Winegrowers 1999. www.nzwine.com/assets/frost_review.pdf