Using Artificial Neural Networks to Predict Local Disease Risk Indicators with Multi-Scale Weather, Land and Crop Data

A. Pechenick

Donna M. Rizzo

Follow this and additional works at: https://scholarsarchive.byu.edu/iemssconference

Pechenick, A. and Rizzo, Donna M., "Using Artificial Neural Networks to Predict Local Disease Risk Indicators with Multi-Scale Weather, Land and Crop Data" (2008). International Congress on Environmental Modelling and Software. 50.
https://scholarsarchive.byu.edu/iemssconference/2008/all/50

This Event is brought to you for free and open access by the Civil and Environmental Engineering at BYU Scholars Archive. It has been accepted for inclusion in International Congress on Environmental Modelling and Software by an authorized administrator of BYU Scholars Archive. For more information, please contact scholarsarchive@byu.edu, ellen amatangelo@byu.edu.
Using Artificial Neural Networks to Predict Local Disease Risk Indicators with Multi-Scale Weather, Land and Crop Data

A. Pechenick and D.M. Rizzo

College of Engineering & Mathematical Sciences, University of Vermont, Burlington, VT 05405 USA ({apecheni,drizzo}@cems.uvm.edu)

Abstract: The risk of fungal and bacterial crop disease can be predicted using risk models with specific environmental parameters such as temperature, relative humidity, solar radiation, wind speed, and leaf wetness duration (LWD). LWD has long been recognized as key in the management of crop disease. Air temperature and wetness influence the majority of fungal plant diseases. Wetness also impacts insect populations, as well as pollution deposits. Many parameters are well understood, readily defined, and easily measured. Unfortunately, LWD is a complex phenomenon, due to its spatial and temporal variability within a crop canopy. The inconvenience and uncertainty associated with monitoring LWD at the local leaf scale and the complexity of upscaling to the crop level prevent existing disease risk models from being used with reliability. In spite of their imprecision, LWD projections are already included in a number of online weather products. One non-parametric statistical approach receiving scant attention for the modeling of LWD is that of artificial neural networks (ANNs). In this work, two previously untired ANNs estimate this key environmental variable at local crop scales, using local and regional weather station data and site-specific sensing data. The first ANN combines two statistical methods to accomplish this spatial mapping (a K-nearest means classifier and a Bayesian classifier), while the recurrent nature of the second ANN provides a means of leveraging the temporal property of the data. The ultimate goal is to embed the ANN into a highly-portable tool, designed to predict leaf wetness duration as an SOC (system on a chip) in conjunction with local weather stations, and as input to real-time decision support systems.

Keywords: Artificial Neural Networks; Leaf and Surface Wetness; Multi-scale Data; Crop Disease Risk; Decision Support Systems.

1. INTRODUCTION

Disease risk modeling for crop management has been shown to reduce disease incidence and severity [Campbell and Madden, 1990, Funt et al., 1990, Gleason et al., 1994]. Pathogens, pests and air pollution deposits are influenced by several environmental variables including temperature, relative humidity, net solar radiation, wind speed, and surface, or more commonly leaf, wetness duration (LWD) [Huber and Gillespie, 1992; Sharma 1976; Schuepp 1989; Getz 1991]. Among the most critical [Huber and Gillespie, 1992], as well as the most difficult [Sentelhas et al., 2007] of these variables to quantify and forecast are 1) canopy surface wetness (W_{obs}), defined as the observed fraction of plant parts or organs that are wet in a canopy, and 2) canopy surface wetness duration (SWD), defined as the sum of the continuous hours where W_{obs} is greater than 0.1 [Magarey et. al., 2006a, 2006b]. Despite these constraints, decision support systems incorporate LWD
estimates into their products, using both historical, and increasingly, forecast weather data as parameters for their LWD prediction models [Kim et al., 2006]. Model designs include physics-based, empirical and statistical, as well as hybrids of the above; performance trade-offs include complexity, accuracy, convenience and portability [Sentelhas et al., 2008].

2. MODELING SURFACE WETNESS DURATION

Converting \( W_{\text{obs}} \) into SWD requires classifying the data over a given time period (e.g., one hour) into binary categories of wet (1) and dry (0). According to Magarey and Seem [2001], \( W_{\text{obs}} \) is classified as 0 if its value is less than 0.2, or 1 if its value is at least 0.2. Consider the small dataset for a single moisture event, illustrated by Figure 1. Predicted canopy surface wetness (obtained from an artificial neural network described below) and leaf wetness sensor data (obtained from Campbell Scientific leaf wetness sensors [Model 237, Campbell Scientific, Inc., Logan, UT 84321]) are tabulated at left and plotted versus time at right. The table also shows the percentage and binary SW classification for each hour.

<table>
<thead>
<tr>
<th>Hr</th>
<th>LW Sensor</th>
<th>ANN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( W_{\text{obs}} )</td>
<td>( W_{\text{obs}} )</td>
</tr>
<tr>
<td>284</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>285</td>
<td>0.2</td>
<td>0.4</td>
</tr>
<tr>
<td>286</td>
<td>0.5</td>
<td>0.3</td>
</tr>
<tr>
<td>287</td>
<td>0.9</td>
<td>0.8</td>
</tr>
<tr>
<td>288</td>
<td>0.9</td>
<td>0.8</td>
</tr>
<tr>
<td>289</td>
<td>0.8</td>
<td>0.7</td>
</tr>
<tr>
<td>290</td>
<td>0.77</td>
<td>0.7</td>
</tr>
<tr>
<td>291</td>
<td>0.77</td>
<td>0.9</td>
</tr>
<tr>
<td>292</td>
<td>0.9</td>
<td>0.9</td>
</tr>
<tr>
<td>293</td>
<td>0.8</td>
<td>0.9</td>
</tr>
<tr>
<td>294</td>
<td>0.6</td>
<td>0.6</td>
</tr>
<tr>
<td>295</td>
<td>0.2</td>
<td>0.9</td>
</tr>
<tr>
<td>296</td>
<td>0.07</td>
<td>0</td>
</tr>
<tr>
<td>297</td>
<td>0.0</td>
<td>0</td>
</tr>
</tbody>
</table>

\[ \Sigma = 12 \]
\[ \Sigma = 12 \]

Duration (SWD) = 12 hrs for both the sensor and ANN (RMS=0.0).

**Figure 1.** Example calculation of processing \( W_{\text{obs}} \) to SWD for a given moisture event.

Two important measures of error may be used to describe these data. One is the difference between canopy surface wetness sensor measurements and a model prediction prior to classification over each hour, and the second is the difference between sensor measurements and model predictions after classification to 0 or 1, over each hour. The root-mean-square (RMS) error is defined as

\[ \text{RMS} = \sqrt{\frac{\sum_{j}(T_j - \hat{W}_{\text{obj}})^2}{n}} \],

where \( T_j \) is the measurement data (either \( W_{\text{obs}} \) visually recorded by an observer or, in the case of Figure 1, estimated by a Campbell Scientific leaf wetness sensor), \( \hat{W}_{\text{obs}} \) is estimated by a model, and \( n \) is total number of data records over the moisture event. It is important to note that for the moisture event of Figure 1, the RMS error is computed as 0.2330 when comparing canopy surface wetness predicted by a model (indicated by squares) to sensor measurements (circles). However, when \( W_{\text{obs}} \) is post-processed into SWD (up-scaled through time), the RMS error for the same moisture event is 0. In an effort to provide conservative estimates of error (and
because the post-processing of \( W_{obs} \) into SWD is straightforward), we will present model forecasts in terms of \( W_{obs} \).

Overall, an accurate visual representation of \( W_{obs} \) is required for successful model validation and sensor calibration. While temperature, relative humidity, wind speed, and rainfall have become standard measurements in a weather station set-up, surface wetness remains without a standard of measurement [Seem and Magarey, personal communication, 2002; Sentellhas et al., 2007]. Although leaf wetness sensors are presently used as surrogates, experts in quantifying canopy surface wetness rely on visual observations on some number of individual leaves to reflect the spatial aspect of the proportion of leaves wet in an entire canopy. Unfortunately, it is impossible to collect sufficient visual measurements at the same snapshots in time over the entire crop space, and the individual leaf wetness measurements must be averaged. If greater than 10% of the leaves in a canopy appear to have moisture for at least 12 minutes in an hour, the hour is considered “wet”. \( W_{obs} \) is then calculated by averaging the binary values of the leaves over the time period of observation.

3. ARTIFICIAL NEURAL NETWORKS (ANNs)

ANNs have been used to predict wheat leaf wetness [Franci and Panigrahi, 1997], to estimate moisture occurrence and duration [Chtioui et. al, 1999], to forecast treatment for \textit{Plasmopara viticola} infection [Dalla Marta et al. 2005] and to distinguish the type of infection and classify the infection period in a wheat field environment [de Wolf and Franci, 1997; 2000]. ANNs are not programmed; they are data-driven, and learn by example. Typically, an ANN is presented with a training set of examples (training patterns, typically represented as vectors) from which the network can learn. The two ANNs selected for this work utilize supervised learning, during which the ANN is also presented with a target output pattern: the \textit{known} answer (classification) for the corresponding input pattern. The ANN methodology involves training the ANN to iteratively determine the hidden weights that will accurately map appropriate environmental parameters to \( W_{obs} \). During training, examples of the mapping are presented to the network, and the weights of the hidden and output layers are adjusted over a series of iterations until the network has satisfactorily mapped the training set inputs to the training set outputs. Once a satisfactory mapping has been obtained, training ends and the weights are fixed. These fixed weights are then used during the interpolation phase to map new inputs (that the network has never seen) to predict \( \hat{W}_{obs} \). In this work, the initial composition of input vector consisted of local weather variables (specifically temperature, relative humidity, wind speed, net canopy radiation, and leaf area index), with output being the canopy wet surface area, \( W_{obs} \), for the given canopy elevation.

4. MODELING WETNESS DURATION USING COUNTERPROPAGATION ANN

4.1 What Distinguishes This Approach

Two ANNs have been developed, trained, and tested for the prediction of temporal trends associated with local crop scale SWD using multiple types of data measured from local scale and regional weather station data and site specific sensing data. The first is a counterpropagation network (CP), based on the work of Hecht-Nielsen [1987; 1988], the second a recurrent backpropagation network (RBP), independently generalized from standard backpropagation by Pineda [1987] and Almeida [1987]. The majority of earlier ANN attempts to model LWD use standard backpropagation, whose weaknesses, unlike CP, include a tendency towards local minima, the required optimization of multiple parameters, and a “black box” result [Frasconi, et al., 1993; Cristianini, 2001]. As shown in Figure 2,
CP is a combination of two classifiers – K-nearest means (Kohonen) and Bayesian (Grossberg) – and is guaranteed to converge. For details of the method see Hecht-Nielsen [1987]; for application of the method and pseudo-code see Rizzo and Dougherty [1994].

It is important to distinguish this method from the traditional feedforward backpropagation ANN (used in every software package), which requires stochastic training to select/optimize the number of hidden nodes. With the exception of feedforward backpropagation, the majority of the 50 or so existing ANN algorithms found in the literature perform better when large numbers of data are available. This custom counterpropagation algorithm does not suffer from some of the limitations associated with the feedforward backpropagation in that it cannot be over-trained. The hidden layer is a Kohonen self-organizing-map (SOM) used to cluster the data. The output layer maps the clusters to a known a priori classification (turning the Kohonen unsupervised ANN into a supervised ANN). The more training data available, the better the classifications/predictions. This is true of most ANN algorithms, and explains why the majority of most commercial and proprietary data mining applications now use ANNs.

4.2 CounterpropagationPreferred over RecurrentBackpropagation

Comparison of the root-mean-square (RMS) error values between canopy surface wetness data (both visual observations and measurements), $W_{\text{obs}}$, and the ANN predictions, $\hat{W}_{\text{obs}}$, using the two ANNs indicates slightly better predictions using CP over RBP for the moisture events shown in this work (Figures 4 and 5). However, comparisons of predictions to observations (both visual and sensor) over 28 of the moisture events comprising two of the four datasets in Geneva N.Y., indicate no significant advantage in terms of the prediction capability of either network over the other. A significant difference, however, was observed in the time required for training each of the ANNs. RBP takes on average 12 to 13 times longer to train the same number of training patterns to the required RMS error value of $1 \times 10^{-6}$.

[Note: For our largest training set, consisting of 2450 records of multiple types of weather data, and corresponding sensor data measurements at approximately 10 minute intervals, the recurrent backpropagation network never did converge to the required RMS error value. As a result, only predictions using the counterpropagation model will be presented in this paper.]
4.3 Site and Data

A variety of datasets were used for training, testing, and validation. Roger Magarey, while at Cornell University, collected an extensive dataset from four grape cultivars growing at the Climatological Reference Station (NWS) in Geneva, New York. Drs. Seem and Magarey have graciously provided the data needed to train the ANN models for predictions of canopy surface wetness (see Figure 3).

<table>
<thead>
<tr>
<th>Site</th>
<th>Site ID</th>
<th>Cultivar</th>
<th>Data Type</th>
<th>Data Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geneva</td>
<td>GV97</td>
<td>Chardonnay</td>
<td>Ag. Station†; Visual</td>
<td>1 hr</td>
</tr>
<tr>
<td>LO98</td>
<td></td>
<td>Concord</td>
<td>Ag. Station; Visual</td>
<td>10 min</td>
</tr>
<tr>
<td>NWS</td>
<td></td>
<td></td>
<td>Ag. Monitoring Site</td>
<td>1 hr</td>
</tr>
<tr>
<td>Branchport</td>
<td>DE97</td>
<td>Concord</td>
<td>Ag. Station; Visual</td>
<td>1 hr</td>
</tr>
<tr>
<td>Romulus</td>
<td>SW97</td>
<td>Cabernet Franc</td>
<td>Ag. Station; Visual</td>
<td>1 hr</td>
</tr>
<tr>
<td>Syracuse</td>
<td></td>
<td></td>
<td>Weather Station†</td>
<td>1 hr</td>
</tr>
</tbody>
</table>

†Ag station = Temperature, relative humidity, wind speed, LW sensors and net radiometers with measurements of each at separate elevations.

††Weather Station = Typical of data collected at a regional weather station, i.e., wet and dry bulb temperature, relative humidity, wind speed, and direction.

Figure 3. a) Approximate location of the four grape canopy data sets and (b) description of cultivars and types of data collected at each New York data site.

The four cultivar data sets contain temperature, relative humidity, wind speed, canopy net radiation, rainfall, soil heat flux, soil moisture, and surface wetness (both visual and sensor data). Leaf surface wetness measurements were collected at three vines within a canopy. Each vine supported five sensors. Visual canopy surface wetness, \( W_{obs} \), was estimated from the proportion of observed wet leaves. Visual measurements were replicated for three leaves observed at five canopy positions for three vines per site totaling 45 visual measurements at any given time. Sensor measurements of \( W_{obs} \) were also estimated from the proportion of wet sensors. Painted and unpainted Campbell scientific sensors were placed in the same five canopy positions along the same 3 vines. The visual estimates of \( W_{obs} \) provide validation of the forecasts of canopy surface wetness, \( \hat{W}_{obs} \), from the ANN-based models. Observations of canopy surface wetness included both rain and dew moisture events. Temperature and relative humidity were collected at four elevations within the canopy. (Note: Surface wetness measurements (both visual and sensor) were collected at three of these elevations (bottom, middle, and top of the canopy.).) The weather stations logging the local data at all sites (4 sites in total) were located away from the edge of the field, near the middle of a row. Analysis of the \( W_{obs} \) data, (both visual and sensor), indicates a significant dependence on elevation within the canopy, an observation echoed by Jacobs et al. [2005] and Batzer et al. [2008]. As a result of these analyses, it was decided that 1) visual observations (rather than sensor data) would be considered “ground truth” and used throughout this research time frame to train the ANNs, 2) \( W_{obs} \) should not be averaged over the three canopy elevations, and 3) dew events and rains events would be treated separately (i.e., the ANN will be a function of elevation and moisture event-type).
4.4 Methodology

The CP was developed and trained using ten classes of surface wetness ranging from 0 to 1 in increments of 0.1. For the preliminary results shown in Figure 1, only 12 visual observations (along with corresponding measurements of temperature, relative humidity, wind speed, and canopy net radiation) were used as training patterns. Although not discussed here, the selection of training data is crucial to whether or not the network “learns” a particular task: how well a network “learns” depends on the examples presented during training. For this particular moisture event consisting of only 12 training patterns, the visual canopy surface wetness data (each point being an average over all vines, and elevations) contained no instances/classifications of 0.2 or 0.4. As a result, during the testing phase, when input vector patterns to the network consist of temperature, relative humidity, wind speed, and canopy net radiation collected at 10-minute intervals, it is not possible for the ANN to classify the output as 0.2 or 0.4 (i.e., the CP ANN cannot predict outside the data range that it has been trained on). One of the first datasets used for training, testing and validation was Geneva N.Y. (LO98). The growing season consisted of 173 days and contained 116 visual W_{obs} measurements. This training dataset used the visual measurements of W_{obs} from the first 8 moisture events (75 records) spanning 40 days (in addition to temperature, relative humidity, wind speed, and net canopy radiation) as input training patterns, and then predicted \( \hat{W}_{obs} \) over the remaining four moisture events (23 days). Comparisons of the ANN predictions to the observed visual data (along with the corresponding RMS error values and sensor measurements for the same time frames) are plotted in Figure 4. Sensor measurements are plotted on the figure, but only to indicate the beginning and end of moisture events. Similar results were obtained for the top and bottom canopy elevations; however, only the middle canopy elevations are shown.

Initially, a sensitivity analysis was performed on an energy-based surface wetness model (SWEB28) developed by Drs. Magarey and Seem at Cornell University. The results of this analysis indicate that of the four parameters tested (temperature, relative humidity, wind speed, and net canopy radiation), the physics-based model is the most sensitive to small changes in relative humidity, followed by net canopy, and, to a much less extent, wind speed and temperature.

A suite of sensitivity analyses were also performed on the ANN-based models using the two Geneva datasets to determine the most influential parameters needed for forecasting W_{obs} using the ANN. Analysis revealed that two of these 6 inputs (net canopy radiation and leaf area index) did not significantly improve predictions of W_{obs}. Comparison of the RMS error values over all moisture events indicates no significant increase in the RMS error values when net canopy radiation and leaf area index are omitted from the list of variables. In fact, in some cases the error was reduced. As a result, both were removed from the input training sets. A review of the literature corroborates these findings [Kim et al., 2006; Sentelhas et al., 2008].

**Figure 4.** ANN predictions (squares) of canopy surface wetness and visual W_{obs} measurements (large triangles) obtained using moisture event records of local weather data for the Geneva, N.Y. (LO98) data.
4.5 Results and Scalability

The overall RMS error for all moisture events when 75 visual records (8 moisture events) are used in training is 0.2512. The overall RMS error for all moisture events when trained on 2450 records (15 sensor events) is 0.1981. Similar reductions in the RMS error values were found for the GV97 data set (RMS = 0.3014 and 0.1613 respectively). Therefore, significant decrease in the overall RMS error value is found when the training data are increased. Unfortunately, it was only possible to perform this test using the sensor data as the training target outputs; the collection of visual data over the same number of moisture events and time intervals is too labor intensive and companion data sets do not exist.

Hewitson and Crane [1996] describe techniques and applications of climate downscaling. Magarey et. al [2002] describe some emerging technologies, the limitations, and the prospects for future improvements for obtaining site-specific weather data without on-site sensors. Weather data may not be available at local spatial and temporal scales and the errors associated with spatial interpolation may be too large to provide accurate estimates of SWD using local models. To test and validate the ANN for local forecast errors (associated with the use of weather data that has been interpolated or downscaled from regional weather station data), the following test was performed: Crude local estimates of temperature, relative humidity, and wind speed data for Branchport, N.Y. were obtained using the ordinary kriging package available in ArcGIS Spatial Analyst and regional data from three weather stations located in Geneva, Romulus, and Syracuse, N.Y. (see Figure 3a). The data was not adjusted for elevation; the intent was simply to test the robustness and errors associated with the ANN’s ability to forecast local canopy surface wetness given weather data interpolated from regional weather stations. Validation of the ANN forecast was achieved using visual data gathered at the local scale over a trial prediction period.

For training purposes, we used locally gathered Branchport temperature, relative humidity, wind speed measurements, and corresponding leaf wetness sensor measurements provided at 1-hour intervals during the month of July 1997. Estimates of canopy surface wetness were forecast by the ANN for the month of August 1997. Comparison of the ANN forecasts trained on both locally recorded and spatially interpolated weather data (see Figure 5, squares and diamonds respectively) indicate a significant decrease in the RMS error values when local weather data are used. We note that other companies such as SkyBit Technologies, Inc., have experience in interpolating spatial data such as that gathered at weather stations, and deliver weather and disease forecasts in the form of data or maps at the crop scale. Improving local weather estimates of temperature, relative humidity, wind speed, and net canopy radiation using more sophisticated downscaling interpolation methods and shorter prediction intervals will greatly improve ANN forecasts and reduce error estimates. However, since there was no significant decrease in the prediction capabilities of the ANN when the net canopy radiation was omitted from the input training set, downscaling of local weather variables may not be warranted since temperature, relative humidity, and wind speed are relative easy to measure by local growers.
Figure 5. Trained on 3 rain events and 12 dew events LO98 sensor data, interpolated on 5 rain events, where visual and sensor data exist.

5. CONCLUSIONS

The work described above leads to several conclusions:

1. Comparison of the ANN forecasts using both the visual and sensor $W_{\text{obs}}$ measurements as target outputs shows a significant increase in the accuracy of prediction when the number of training examples increases. That is, the ANN is better able to learn or generalize the relationship between input patterns (locally measured weather data) and output patterns (locally observed $W_{\text{obs}}$ measurements) when the number of training patterns increases. One advantage of incorporating the statistical power of ANNs into existing risk disease models is that they “learn” the empirical relationships directly from the measured data. These relationships may be nonlinear and a neural network may model physical relationships with high degrees of nonlinearity.

2. We expect the ANN methodology to be easily transferable to other crops. However, the validation of model predictions with sensor and visual observations over a range of crop types and climates using standard protocol will be mandatory.

3. A standard for surface wetness measurements is required to validate models used to predict $W_{\text{obs}}$ and SWD. Without standardized measurements to accomplish proper validation, these models will face barriers to entry into disease management strategies.

4. Visual observations and sensor measurements of leaf wetness represent point measurements of wetness and, as a result, cannot measure the proportion of leaves wet in a canopy ($W_{\text{obs}}$). Measurements are therefore averaged to provide information that is appropriate for the larger, canopy scale.
5. There was no significant increase in the prediction capabilities of the ANN when net canopy radiation and leaf area index were omitted from the input training set, and in some cases, the error was actually reduced. This is a favorable outcome, as net radiation measurements and estimates are not as commonly available as wind, rain, relative humidity, and temperature measurements.

6. Local collection of temperature, relative humidity, and wind speed data (as opposed to obtaining those weather variables from more complicated downscaling interpolation models) may be warranted.

ACKNOWLEDGEMENTS

We are grateful to Dr. Robert Seem, Professor of Plant Pathology at Cornell University and the New York State Agricultural Experiment Station, and Dr. Roger Magarey, Plant Pathologist at USDA-APHIS and North Carolina State University, for fostering a productive collaboration and for providing an extensive dataset and preliminary copy of their canopy water budget model (SWEB Version 0.28).

REFERENCES


Hewitson, B.C., and R.G. Crane, Climate downscaling: Techniques and application, Climate Research, 7, 85-95, 1996.
Schuepp, P. H., Microstructure, density and wetness effects on dry deposition to foliage, Agricultural and Forest Meteorology, 47(2-4), 179-198, 1989.