한국농림기상학회지, 제4권 제1호(2002)
Korean Journal of Agricultural and Forest Meteorology, Vol. 4, No. 1, (2002), pp. 23~28

Optimal Weather Variables for Estimation of Leaf Wetness Duration Using an Empirical Method

K. S. Kim[1], S. E. Taylor[1], M. L. Gleason[2] and K. J. Koehler[3]
[1]*Department of Agronomy, Iowa State University, Ames 50011, USA*
[2]*Department of Plant Pathology, Iowa State University, Ames 50011, USA*
[3]*Department of Statistics, Iowa State University, Ames 50011, USA*
(Received December 21, 2001; Accepted February 18, 2002)

결로시간 예측을 위한 경험모형의 최적 기상변수

김광수[1] · S. E. Taylor[1] · M. L. Gleason[2] · K. J. Koehler[3]
[1]미국 아이오와 주립대학 농학과, [2]미국 아이오와 주립대학 식물병리학과
[3]미국 아이오와 주립대학 통계학과
(2001년 12월 21일 접수; 2002년 2월 18일 수락)

ABSTRACT

Sets of weather variables for estimation of LWD were evaluated using CART(Classification And Regression Tree) models. Input variables were sets of hourly observations of air temperature at 0.3-m and 1.5-m height, relative humidity(RH), and wind speed that were obtained from May to September in 1997, 1998, and 1999 at 15 weather stations in Iowa, Illinois, and Nebraska, USA. A model that included air temperature at 0.3-m height, RH, and wind speed showed the lowest misidentification rate for wetness. The model estimated presence or absence of wetness more accurately (85.5%) than the CART/SLD model (84.7%) proposed by Gleason *et al*. (1994). This slight improvement, however, was insufficient to justify the use of our model, which requires additional measurements, in preference to the CART/SLD model. This study demonstrated that the use of measurements of temperature, humidity, and wind from automated stations was sufficient to make LWD estimations of reasonable accuracy when the CART/SLD model was used. Therefore, implementation of crop disease-warning systems may be facilitated by application of the CART/SLD model that inputs readily obtainable weather observations.

Key words : **Disease warning system, integrated pest management, weather models**

1. INTRODUCTION

Leaf wetness duration(LWD) is a key environmental parameter for assessing the likelihood of outbreaks of many economically important crop diseases (e.g., Timmer *et al*., 2000; Carisse *et al*., 2000). Consequently, LWD estimates are frequently used as inputs for disease-warning systems, which advise growers on efficient timing of fungicide sprays or other management tactics (Gleason, 2000). Even though electronic sensors can measure LWD indirectly, estimating it with models may be more convenient since models free growers from the need to install and maintain the sensors and data loggers.

Wetness on a leaf can be caused by irrigation, mist, rain, guttation, or dew. Dew duration has been modeled using energy balance equations or empirical approaches (Pedro and Gillespie, 1982a, 1982b; Gleason *et al*., 1994). Pedro and Gillespie (1982a, 1982b) built a model utilizing energy balance theory that estimated LWD during dew period within 1 h per night. Empirical models that used statistical procedures, such as

Corresponding Author : Kwang Soo Kim(madk@iastate.edu)

CART(Classification and Regression Tree), are also able to estimate dew-derived LWD within 1 h per night (Gleason *et al*., 1994).

When an empirical approach is used to build a LWD model, a key step is to determine which variables should be included in the model. In general, input variables of empirical models have been selected from the variables monitored at conventional weather stations, such as air temperature, RH, and wind speed (Huber and Gillespie, 1992). However, energy balance analysis concerning thermal radiation on a surface suggests that cloud cover has an important influence on dew formation, which is a key component of LWD. Air temperature at the crop canopy level may also influence LWD because dew is a by-product of heat exchange near the canopy. Although these variables may be useful for accurate estimation of LWD, they are seldom available from standard weather stations. Few empirical LWD estimation models, therefore, have used either of these variables. In this study, models were developed using air temperature near the surface and cloud cover for estimation of LWD to find an optimal set of weather variables with CART technique.

2. MATERIALS AND METHODS

2.1. Data acquisition

Hourly measurements of air temperature, RH, and wind speed were collected from May to September in 1997, 1998, and 1999 at 15 sites in Iowa (IA), Illinois (IL), and Nebraska (NE), USA (Fig. 1). Wind speed was measured at a height of 3 m (IA and NE) or 10 m (IL). Air temperature and RH were measured at a height of 1.5 m. A thermocouple thermometer (Model 107, Campbell Scientific) was installed to measure air temperature at 0.3-m height, the same level as the wetness sensors. Cloud cover estimates for each site were obtained for 1999 from SkyBit, Inc. (Bellefonte, PA).

Fig. 1. Locations where wetness was measured from 1997 to 1999 (May-September). Data from the sites indicated by triangles, squares, and circles were used as training, validation, and prediction sets, respectively.

Leaf wetness was measured using electrical wetness sensors (Model 237, Campbell Scientific, Logan, UT) deployed 0.3 m above managed turfgrass. Each sensor was mounted on the end of a segment of PVC pipe and deployed facing north at an angle of 45^o to horizontal. The sensor surface was painted with latex paint in order to increase sensitivity to small water droplets and to approximate the emissivity of plant leaves (Davis and Hughes, 1970). The sensors were oven-dried overnight to remove moisture from the paint. Wetness sensors were interrogated by data loggers every 5 min, and hourly wetness data were summarized as the proportion of an hour when the sensor was wet. An hour was scored as 0 (“dry hour”) when the sensor was wet for ≤ 30 min and 1 (“wet hour”) when the sensor was wet for > 30 min.

2.2. Model development and validation

Air temperature at 0.3-m and 1.5-m heights, RH, wind speed, dew point depression, and cloud cover were used as input variables to build LWD estimation models for the dew-eligible period (20:00 to 9:00 the next morning). A classification and regression tree (CART) technique (Breiman *et al*., 1984) was used with S-plus (Math Soft Inc., Cambridge, MA; Venables and Ripley, 1997). The CART modeling consisted of three consecutive steps: growing, validation, and pruning. Data sets from randomly selected sites (Fig. 1) used for the growing and validation steps will be referred to as the training and validation data sets, respectively. Four trees were grown from training data, distinguished by whether or not either rainfall or cloud cover was included (Table 1). Due to the lack of cloud cover estimates in 1997 and 1998, trees that used cloud cover as a potential variable were fit only to the training 1999 data.

Table 1. Data sets prepared for CART analysis

Data Set	Years Included	Days with measured rain[a]	Cloud Cover
A	1997-1999	Included	No cloud data included
B	1999	Included	Cloud data from SkyBit
C	1997-1999	Excluded	No cloud data included
D	1999	Excluded	Cloud data from SkyBit

a. $\geq$ 0.25 mm/day

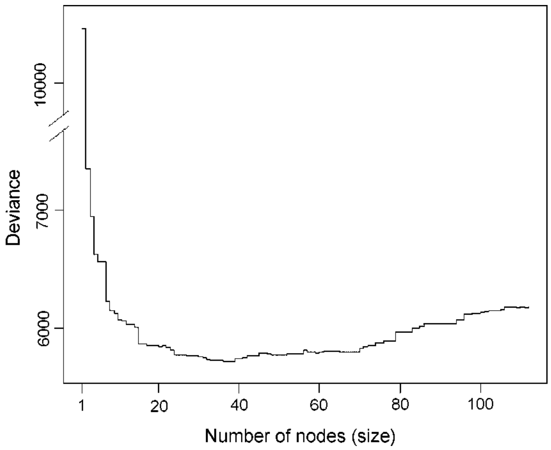

Fig. 2. Plot to determine optimal size of a tree in set A. Data from the training set were used for validation. Smaller deviance values (Y axis) indicate better prediction of response variable, which is either wetness or dryness in the data. The size of the tree, which is the number of terminal nodes, is give by the X axis.

The first step in fitting a tree to the training data is to generate a very large tree. Using this very large tree to predict LWD for every hour in the validation data set, a tree size, the number of nodes, that provides the most accurate predictions, is selected. As the size of the tree fit to the training data is increased, it eventually begins to react to random variation. At that point, predictions for the validation data become less accurate. This is shown in Fig. 2, where including the first 15 nodes in the tree fit to the training data provides reasonable accuracy. Only slight improvement is obtained by increasing model complexity to 38 nodes. After the optimal size for a tree was determined, the tree was pruned to the indicated number of nodes. The predictions for the validation data are used to estimate misclassification rates (i.e., number of wet hours per day misclassified as dry or vice versa) for hours of leaf wetness.

The misclassification rate was used to select the best pruned trees for each of the four validation data sets. For each data set, the tree with the smallest misclassification rate was designated as the "Model" given the name of the data set. For analysis of model performance, measurements from wetness sensors were assumed to be "true", and deviations of model estimates from wetness sensor measurement were considered to be errors. The Critical Success Index(CSI) (Schaefer, 1990) was used to evaluate accuracy in predicting wetness events by calculating the proportion of hours in which the occurrence of wetness was estimated correctly as a proportion of the total hours in which either sensors or the model identified wetness. The model's no-alarm rate (proportion of hours in which wetness was measured by sensors but not estimated) and false-alarm rate (proportion of hours in which wetness was estimated but not measured) were calculated daily and averaged over the study period. Differences between measured and model-estimated LWD for the dew-eligible period were averaged and designated as the Mean Error(ME). Overall accuracy, defined as the percentage of hours in which each model identified wetness or dryness correctly for a night, was calculated over the study period.

3. RESULTS AND DISCUSSION

3.1. Model selection

The input variables selected for the most accurate model are shown in Table 2 for each data set. Air temperature at 0.3-m height was an input variable in each model. Wind speed and dew point depression (DPD) were selected in the analyses of the 1997-1999 data sets for which no cloud cover information was available. Cloud cover was also included in each analysis

Table 2. Models that provided the lowest misclassification rate for each data set

Data Set	Total hours	Input variables (Size[a])	False alarm rate[b]	Misclassification rate[c]
A	18,432	RH, WSPD[d], ATMP30[d] (15)	0.171	0.207
B	7,476	RH, ATMP30, Cover[d] (20)	0.248	0.205
C	13,260	RH, WSPD, ATMP30, DPD[d] (15)	0.213	0.211
D	5,736	RH, ATMP30, Cover (15)	0.348	0.191

a. Number of terminal nodes

b. Proportion of hours that were not scored as wet by wetness sensors but out of the hours scored as wet by the model.

c. Proportion of hours in which a model identified incorrectly either wetness or dryness.

d. WSPD, ATMP30, and Cover correspond to wind speed (m s^{-1}), air temperature at the height of 0.3 m (^{o}C), and cloud cover, respectively.

of the 1999 data for which cloud cover information was available.

Both models that included cloud cover among the input variables classified wet hours more accurately on nights in which rainfall was measured than nights in which no rainfall was measured (data not shown). These models, however, identified dry hours more correctly than wet hours when nights with measured rainfall were excluded from the data set; the remaining date constituted 88% of the study period. Models that classified dry hours more accurately than wet hours appeared to be more accurate when all hours in the data set were considered, because the data sets contained more dry hours than wet hours. Those models, nevertheless, may not be useful since, for disease warning applications, accurate estimation of wetness duration is more important than dryness duration. For example, Model D has the lowest overall misclassification rate, which considers both wet and dry, but has the highest false-alarm rate, which is based on hours when wetness was predicted by a model but not measured by the sensors.

3.2. Comparison of Model A with the CART/SLD model

Model A was selected as the best model to compare with the CART/SLD model because it had the lowest

Table 3. Accuracy in identifying either wet or dry hours in which wetness was measured during the hours 20:00-9:00

Location	Total (h)	Model A			CART/SLD		
		False alarm rate[a]	No alarm rate[b]	CSI[c]	False alarm rate	No alarm rate	CSI
Lewis, IA	2,580	0.172	0.072	0.424	0.137	0.091	0.416
Crawfordsville, IA	3,696	0.121	0.058	0.653	0.113	0.072	0.638
Belleville, IL	3,300	0.218	0.018	0.634	0.200	0.030	0.624
Monmouth, IL	4,032	0.159	0.079	0.547	0.068	0.178	0.464
Red Cloud, NE	2,424	0.097	0.120	0.486	0.069	0.196	0.383
West Point, NE	2,388	0.046	0.083	0.704	0.023	0.163	0.608
All six sites	18,420	0.141	0.069	0.582	0.105	0.118	0.532

a. Proportion of hours in which sensors measured no wetness but the model estimated wetness.
b. Proportion of hours in which sensors measured wetness but the model estimated dryness.
c. Critical Success Index. Proportion of hours in which the occurrence of wetness was estimated correctly as a proportion of the total hours in which either sensors or the model identified wetness.

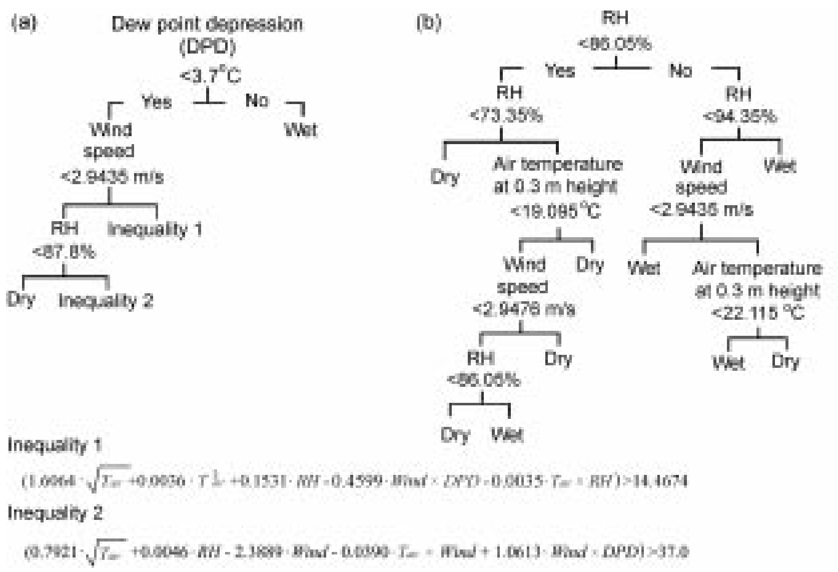

Fig. 3. (a) The structure of the CART/SLD model. If either inequality 1 or 2 is met, the hour is classified as a wet hour where T_{air} is air temperature at 1.5 m (b) The structure of Model A.

false alarm rate (Table 2). Higher CSI values indicated that Model A identified wet hours more accurately than the CART/SLD model (Table 3). The no-alarm rates for Model A were lower than that of the CART/SLD model for all sites. Overall, no-alarm rates were 0.069 and 0.118 for Model A and the CART/SLD model, respectively. Model A, however, misclassified a higher proportion of hours in which no wetness was measured, resulting in higher false-alarm rates than CART/SLD.

Inclusion of air temperature at 0.3-m height may have enabled Model A to identify dew periods at this height more accurately than the CART/SLD model, which uses air temperature at 1.5-m height, by accurately representing the microenvironmental conditions under which dew was likely to occur. When wetness was measured at lower RH values, the difference between 0.3-m and 1.5-m temperatures was relatively large (Fig. 4). During clear and calm nights, when net radiation is usually negative, the temperature of leaf or sensor surfaces is lower than that of air. This favors dew formation by promoting condensation of water vapor on the surface, implying that measuring air temperature near crop surfaces may help to estimate LWD more accurately. In addition, air temperature near the surface of leaves may provide a more accurate representation of the thermal environment of fungal spores on relatively low-canopy crops, such as strawberries and melons, for use with disease-warning systems (Huber and Gillespie, 1992).

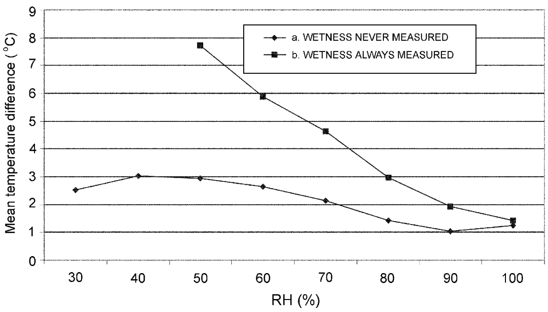

Fig. 4. The mean temperature difference (ΔT) between 1.5 m and 0.3 m heights under calm conditions (wind speed < 3 m/sec) during dew-eligible period (20:00 to 9:00). Wetness was measured at given relative humidity (RH) when ΔT was large.

The overall mean differences between measured and estimated wetness duration were 0.9 h and -0.2 h per night for Model A and the CART/SLD model, respectively (Table 4). Some models based on energy balance have reported <1 h/day error for LWD during dew-eligible periods (Pedro and Gillespie, 1984). Variability of LWD estimation, as indicated by the standard error of the mean (SEM), differed little between the models. The CART/SLD model estimated less LWD than Model A for all six sites. Gleason *et al*. (1994) reported that the CART/SLD model tended to underestimate dew duration in the midwestern US by 0.8 h/day.

The relatively slight improvement in LWD estimation accuracy provided by Model A compared to the CART/SLD model suggests that use of air temperature, RH, and wind speed with the CART/SLD model is operationally preferable, since all of these inputs are obtainable from standard weather stations, whereas Model A requires canopy temperature data that are not routinely gathered. Alternatively, it is also possible to make reasonably accurate site-specific predictions of LWD for a chosen locality or farm using the CART/SLD model with commercially available, site-specific estimates for the input variables (Kim *et al*., 2001). Linking the CART/SLD model to site-specific weather data deserves further attention as a convenient technique of providing LWD data to drive disease-warning systems on high-value crops.

Table 4. Mean error and accuracy for Model A and the CART/SLD model

Location	Nights	Mean Error[a] (h / night)		Accuracy[b] (%)	
		Model A	CART/SLD	Model A	CART/SLD
Lewis, IA	215	1.2 (0.24)	0.6 (0.23)	84.0	85.3
Crawfordsville, IA	308	0.8 (0.16)	0.5 (0.17)	87.7	87.4
Belleville, IL	275	2.4 (0.17)	2.0 (0.18)	84.2	83.4
Monmouth, IL	336	1.0 (0.21)	-1.3 (0.21)	82.3	82.5
Red Cloud, NE	202	-0.3 (0.24)	-1.5 (0.27)	86.7	84.0
West Point, NE	199	-0.5 (0.15)	-1.7 (0.17)	89.6	86.6
All Stations	1,535	0.9 (0.10)	-0.2 (0.11)	85.5	84.7

a. Mean of differences between model-estimated and measured (estimated - measured) LWD per day.
b. [Hours in which presence or absence of wetness estimated correctly/total hours in data set] × 100.

4. CONCLUSIONS

An empirical model using air temperature at 0.3-m height, wind speed, and RH identified wet hours during the period 20:00 to 9:00 more accurately than the CART/SLD model. Values of the Critical Success Index, which indicate accuracy of identifying wet hours, were 0.582 and 0.532 for Model A and the CART/SLD model, respectively. Model A, however, tended to misidentify more dry hours as wet, resulting in little overall improvement in LWD estimation accuracy. Mean errors for Model A and the CART/SLD model were 0.9 h and -0.2 h per night, respectively. This study demonstrated that, using the CART/SLD model, routine measurements of temperature, humidity, wind, and precipitation from automated stations were sufficient for LWD estimation within 1 h per night. Utilization of disease-warning systems in agriculture may be facilitated by application of the CART/SLD model that inputs readily obtainable weather measurements or site-specific weather estimates.

요 약

CART(Classification and Regression Tree) 모형을 이용해서 결로시간 예측에 필요한 기상변수들을 평가하였다. 입력 기상 변수들은 0.3 m와 1.5 m에서 측정된 기온, 상대습도, 풍속의 시간별 측정값으로서 이 관측값들은 1997년 부터 1999년 5월에서 9월 사이에 미국의 Iowa, Illinois 및 Nebraska 주에 위치한 15개 자동 기상 관측소에서 관측된 것이다. 0.3 m에서 측정된 기온, 상대습도, 그리고 풍속을 이용해서 얻어진 모형이 가장 높은 결로시간의 예측 적중율(85.5%)을 보였으며, 이 모형은 Gleason 등(1994)의 CART/SLD 모형의 적중율(84.7%) 보다 다소 높았다. 그러나 새로운 변수를 추가한 경우에 정확도의 향상이 다소 있었으나 CART/SLD 모형을 대체할 정도는 아니었다. 따라서, 기온, 상대습도, 풍속들의 종관 기상관측값들을 입력변수로 사용하는 CART/SLD 모형이 종관 기상관측 자료 이외의 추가적인 자료를 필요로 하는 모형으로 결로시간을 예측하는 것보다 합리적일 것으로 보인다.

ACKNOWLEDGEMENT

Funding was provided by grant #98-34104-6118 from the USDA North Central Region IPM Program. We thank Robert Scott and Stephen Hollinger of the Midwest Climate Center, Champaign, IL, Shellie Hanneman, Karl Blauvelt, and Kenneth Hubbard of the High Plains Climate Center, Lincoln, NE, and Kert Berns, Iowa State University, who cooperated by gathering and compiling weather data.

REFERENCES

Breiman, L., J.H. Friedman, R.A. Olshen, and C.J. Stone, 1984: *Classification and regression trees*. Wadsworth, Monterey. 2-6.

Carisse, O., G. Bourgeois, and J.A. Duthie, 2000: Influence of temperature and leaf wetness duration on infection of strawberry leaves by *Mycosphaerella fragariae*. Phytopathology **90**, 1120-1125.

Davis, D.R. and J.E. Hughes, 1970: A new approach to recording the wetting parameter by the use of electrical resistance sensors. *Plant Disease Report* **54**, 474-479.

Francl, L.J. and S. Panigrahi, 1997: Artificial neural network models of wheat leaf wetness. *Agricultural and Forest Meteorology* **88**, 57-65.

Gleason, M.L, S.E. Taylor, T.M. Loughin, and K.J. Koehler, 1994: Development and validation of an empirical model to estimate the duration of dew periods. *Plant Disease* **78**, 1011-1016.

Gleason, M.L., 2000: Disease-warning systems. In: *Encyclopedia of Plant Pathology. Vol I*. John Wiley & Sons. New York. 367-370.

Huber, L. and T.J. Gillespie, 1992: Modeling leaf wetness in relation to plant disease epidemiology. *Annual Review of Phytopathology* **30**, 553-577.

Kim, K.S., S.E. Taylor, M.L. Gleason and K.J. Koehler, 2002: A model to estimate site-specific estimation of wetness duration using a wind speed correction. *Plant Disease* **86**, 179-185

Pedro, M.J. and T.J. Gillespie, 1982: Estimating dew duration. I. Utilizing micrometeorological data. *Agricultural Meteorology* **25**, 283-296.

Pedro, M.J. and T.J. Gillespie, 1982: Estimating dew duration. II. Utilizing standard weather station data. *Agricultural Meteorology* **25**, 297-310.

Schaefer, J.T., 1990: The critical success index as an indicator of warning skill. *Weather and forecasting* **5**, 570-575.

Timmer, L.W., H.M. Darhower, S.E. Zitko, T.L. Peever, A.M. Ibnez, and P.M. Bushong, 2000: Environmental factors affecting the severity of Alternaria brown spot of citrus and their potential use in timing fungicide applications. *Plant Disease* **84**, 638-643.

Venables, W.N. and B.D. Ripley, 1997: Tree-based models. *Modern Applied Statistics with S-plus*. (2nd ed.) Springer-Verlag, New York. 413-430.