P1.1 VALIDATION OF SITE-SPECIFIC ESTIMATION OF WEATHER VARIABLES IN THE UPPER MIDWEST AND APPLICATION TO DISEASE RISK ASSESSMENT

Stephen N. Wegulo, Mark L. Gleason*, Kwang S. Kim, and S. Elwynn Taylor
Iowa State University, Ames, Iowa

Weather plays a major role in the occurrence, severity, and spread of plant diseases. Weather variables that influence development of plant diseases include air temperature, relative humidity (RH), leaf wetness, rainfall, and wind. These variables can be used in disease management strategies such as disease-warning systems to reduce pesticide sprays and increase grower profits. Many growers, however, have been reluctant to adopt disease-warning systems because the equipment required for on-site measurements is expensive and it takes time and expertise to set up the equipment and download data (Campbell and Madden, 1990). Site-specific estimation of weather variables is a potentially attractive alternative because it avoids or minimizes the expense and inconveniences associated with on-site measurements.

Site-specific weather estimates (up to 1 km^2 resolution) are obtained using site-specific weather simulators (Magarey et al, 2001). SkyBit, Inc. (Bellefonte, PA) utilizes raw data from the National Weather Service (NWS), the Federal Aviation Administration (FAA), and the U.S. military (Magarey et al, 2001) and delivers site-specific weather estimates to end users by electronic mail or fax.

The objectives of this study were to (i) validate site-specific estimates of temperature, duration of periods with RH > 90%, leaf wetness duration, and rainfall amount and duration provided by SkyBit, Inc. against on-site measurements in the Midwestern U.S., (ii) spatially compare accuracy of site-specific commercial estimates of weather variables to ground-station measurements, and (iii) assess the impact of weather estimation errors on simulated performance of various disease-warning systems.

MATERIALS AND METHODS

On-site measurement of weather variables. Air temperature, RH, and rainfall were recorded hourly at 15 automated weather stations in Iowa (IA), Illinois (IL), and Nebraska (NE) (Fig. 1) from May to September of 1997, 1998, and 1999. At each station, air temperature, RH, and rainfall were measured using a thermistor, an electronic hygrometer, and a tipping bucket rain gage, respectively. Air temperature and RH were measured at a height of 1.5 m whereas rainfall was measured at a 1-m height. To record wetness, flat, printed-circuit electronic sensors (Model 237, Campbell Scientific, Logan, UT) were deployed at a 45^o angle at each of the 15 weather stations in IA, IL, and NE. The sensors were 0.3 m above the ground and faced north at level and unobstructed sites on managed turf grass. Sensor surfaces were covered with latex paint of proprietary composition (R. Olson, Savannah, GA, personal communication) in order to increase sensor sensitivity to small water droplets and to simulate emissivity of plant leaves (Lau et al, 2000). The sensors were determined by field calibration to be accurate within 1 h/day. An hour was scored as wet ("1") when the sensors detected wetness for ≥ 30 min or dry ("0") when wetness was detected for < 30 min.

Site-specific estimation of weather variables. SkyBit, Inc. (Bellefonte, PA) provided, by electronic mail, model-derived hourly estimates of air temperature, RH, leaf wetness duration (LWD), rainfall, and wind speed for the 15 weather stations in IA, IL, and NE from May to September in 1997, 1998, and 1999. RH data from both on-site measurements and site-specific estimates were used to determine duration of periods with RH > 90% because this index is used in several disease-warning systems, for example BLITECAST (Krause et al, 1975).

Validation of site-specific estimates against on-site measurements. To avoid splitting dew periods between two days, a day was considered to run from 1200 noon to 1100 the following morning. On-site measurements of weather variables were assumed to approximate reality. Deviations of SkyBit estimates from on-site measurements were assumed to be errors. Bias of SkyBit estimates was represented by mean error (ME) and accuracy was represented by mean absolute error (MAE) (Wilks, 1995). Mean error was calculated as

$$ME = \frac{1}{n}\sum_{k=1}^{n}(e_k - m_k) \tag{1}$$

and mean absolute error was computed as

$$MAE = \frac{1}{n}\sum_{k=1}^{n}|e_k - m_k| \tag{2}$$

in which n is the number of 24-hour periods during the 3-year study, e is the sum (discrete variables) or average (continuous variable = temperature) of hourly SkyBit estimates in a 24-hour period, m is the sum (discrete variables) or average (continuous variable = temperature) of hourly on-site measurements in a 24-hour period, and (e_k, m_k) are e and m in the kth 24-hour period.

Fig. 1. Locations of 15 weather stations in Iowa, Illinois, and Nebraska where on-site weather data were recorded, and for which SkyBit weather data were estimated from May to September of 1997-1999.

Standard error of the mean (SEM), another measure of accuracy, was calculated as

$$\text{SEM} = \frac{s}{\sqrt{n}} \tag{3}$$

where s is the square root of the average squared difference between SkyBit estimates and on-site measurements and n is the number of 24-hour periods in the 3-year study.

* *Corresponding author and address:* Mark L. Gleason, Iowa State University, Dept. of Plant Pathology, Ames, IA 50011; e-mail: mgleason@iastate.edu.

Spatial comparison of accuracy of site-specific estimates to ground-station measurements. Another way to assess errors of weather estimates is to compare these errors to those of measurements made by ground stations at various distances from a given site. Thus, a grower or pest manager may choose between commercial estimates for his/her farm and measurements made some distance away from the farm. Coefficients of determination (R^2) for all pairs of the 15 weather stations included in the study were calculated for air temperature and used to gauge accuracy of using measurements at one station to estimate temperature at another station (Hubbard 1994), and these values were compared to overall mean R^2 values for SkyBit temperature estimates. In the same manner, the critical success index (Schaefer, 1990) was used to determine accuracy for LWD, daily duration of periods with RH > 90%, and rainfall occurrence.

R^2 and CSI values were used to determine the mean distance from one weather station to another at which accuracy of station-to-station estimation was equivalent to accuracy of SkyBit estimates. R^2 and CSI values from ground stations were plotted against corresponding distances between stations and corresponding values from SkyBit were overlapped on the plot for comparison. The CART/SLD/Wind (Classification and Regression Tree/Stepwise Linear Discriminant/Wind) speed model developed by Kim et al (2002) was applied to LWD data to determine whether it enhanced spatial accuracy of SkyBit LWD estimation.

Impact of site-specific estimation errors on simulated performance of BLITECAST and the P-Days model. Site-specific estimates and on-site measurements of maximum and minimum air temperature, rainfall amount, number of hours per day with RH > 90%, and the maximum and minimum temperature during periods with RH > 90% were input into the disease management module of WISDOM, a computer program that helps potato growers in Wisconsin to make crop management decisions on a daily basis (Pscheidt and Stevenson 1986). The disease management module of WISDOM contains the BLITECAST (Krause et al 1975) disease-warning system for management of potato late blight caused by *P. infestans,* and a physiological days (P-Days) model (Pscheidt and Stevenson 1986) for management of potato early blight, caused by *Alternaria solani.* Late blight development in BLITECAST depends mainly on RH > 90% and temperature during periods with RH > 90%, whereas the P-Days model is entirely dependent on temperature. In BLITECAST, fungicide sprays are initiated when a threshold of 18 severity values (DSVs) is reached. In the P-Days model, a threshold of 300 P-Days signals the initial rise in airborne spore concentrations of *A. solani.* At this time weekly protectant fungicide sprays are initiated for management of early blight.

Impact of site-specific estimation errors on simulated performance of a dew-dependent model. On-site measurements and SkyBit estimates of temperature and LWD were input into TOM-CAST (Gleason et al, 1995), which utilizes LWD and temperature as inputs (Madden et al, 1978). An 85-to-96-day growing season was simulated. The chlorothalonil option for fungicide application was used. The CART/SLD/Wind (Classification and Regression Tree/Stepwise Linear Discriminant/Wind) speed model developed by Kim et al (2002) was applied to SkyBit LWD estimates to determine its impact on simulated performance of TOM-CAST. The TOM-CAST program was written in Microsoft Visual FoxPro (Microsoft Corporation, Redmond, WA). The impact of the CART/SLD/Wind speed model on simulated performance of TOM-CAST was evaluated by comparing the number of sprays and mean absolute deviations, from on-site values, of SkyBit and model-corrected SkyBit cumulative severity values (DSVs).

Impact of site-specific estimation errors on simulated performance of a rain-dependent model. Model prediction of post-bloom fruit drop of citrus caused by *C. acutatum* (Timmer and Zitco 1996) is based on the current number of infected flowers and total rainfall for the previous five days. We chose this model because rainfall is its sole input. The equation is

$$y = -7.15 + 1.2\sqrt{TD} + 0.44\sqrt{100R} \qquad (4)$$

in which *y* is the percentage of infected flowers at the next disease assessment, *TD* is the total number of diseased flowers on 20 trees, and *R* is total rainfall (mm) during the past 5 days. Fungicide applications are made if all of the following criteria are met: (i) *y* > 20%, (ii) a significant portion of the crop (> 75%) is in bloom, and (iii) no fungicide application has been made the previous 14 days.

The model was simulated by inputting on-site and SkyBit rainfall data from 15 weather stations in the Midwest over a 3-year period (1997-1998). The mean percentage of diseased flowers on a single assessment date, the number of fungicide sprays, and the time from the first disease assessment to the first fungicide spray were compared for on-site versus SkyBit data. The computer program used in simulations was written in Microsoft FoxPro (Microsoft Corporation, Redmond, WA).

RESULTS

Validation of site-specific estimates against on-site measurements. SkyBit estimated daily mean temperature within 0.2°C, but overestimated hourly daytime temperature by up to 0.4°C and underestimated hourly nighttime temperature by up to 0.4°C (Table 1; Fig. 2). Mean hourly temperature was overestimated by up to 1.4°C when wetness or RH > 90% was measured, especially during daytime and pre-dawn hours.

SkyBit underestimated duration of periods with RH > 90% by an average of 4.2 h/day (Table 1). Underestimation was consistent across sites, ranging from 1.3 to 7.5 h (Table 1). SkyBit underestimated duration of periods with RH > 90% at all times of day; this underestimation was greatest during the night and least in the afternoon (Fig. 2).

Overall, SkyBit underestimated mean LWD by 1.3 h/day (Table 1). In a 24-hour period, SkyBit overestimated LWD during the day and underestimated it at night (Fig. 2). Overestimation peaked at 09:00, whereas underestimation peaked at 01:00 (Fig. 2).

On days when rain was either measured or estimated, or both, SkyBit overestimated rain duration by 2.6 h/day and rain amount by 2.5 mm/day. SkyBit misidentified non-rain days as rain days on 612 of 5,236 days or 12% of the time. In a 24-h period, SkyBit consistently overestimated hourly rainfall duration (by up 0.052 hr/hr) and amount (by up 0.07 mm/hr).

Spatial comparison of accuracy of site-specific estimates to on-site measurements. SkyBit estimated daily mean temperature with greater accuracy than ground-station measurements made 71 km from a given site, the minimum site-to-site distance for which the comparison was made (Fig. 3). Accuracy of ground-station measurements of daily mean temperature declined rapidly with distance (Fig. 3). SkyBit also detected daily rain occurrence with greater accuracy than ground-station measurements made at all distances evaluated (Fig. 3). Accuracy of ground-station measurements of duration of periods with RH > 90% was greater than that of mean SkyBit estimates up to 585 km. Beyond this distance, accuracy of SkyBit estimates of duration of periods with RH > 90% exceeded that of ground-station measurements (Fig. 3). Accuracy of ground-station measurements of hourly wetness duration was greater than that of SkyBit estimates at all distances evaluated. When the CARTD/SLD/wind speed model of Kim et al (2002) was applied to SkyBit wetness estimates, accuracy of SkyBit exceeded that of ground-station measurements at all distances evaluated (Fig. 3).

Impact of site-specific errors on performance of simulated disease-warning systems. SkyBit data (duration of periods with RH > 90%) resulted in 11 compared ($P < 0.0001$) to 85 severity values for on-site data in simulated performance of the BLITECAST (potato late blight) disease-warning system. SkyBit data delayed fungicide spray thresholds by up to 68 days. In the physiological days (P-Days) model for predicting the likelihood of occurrence of potato early blight, SkyBit data (daily mean temperature) were as accurate as on-site data in predicting fungicide spray threshold dates.

In simulated performance of TOM-CAST, SkyBit data (LWD and temperature) resulted in an average of 3.7 compared ($P = 0.03$) to 4.7 fungicide sprays for on-site data during the simulated growing season. The threshold for the first fungicide spray from SkyBit data, however, preceded that from on-site data in eight of 10 site-years. Application of the CARTD/SLD/Wind speed model of Kim et al (2002) to SkyBit data corrected the number of fungicide sprays from 3.7 to 5.

In simulated performance of the rain-dependent model developed by Timmer and Zitko (1996) for management of post-bloom fruit drop of citrus, SkyBit data resulted in 20.8% disease incidence (DI), 2.3 sprays (S), and 3.7 days from the first disease assessment to the first spray (DTFS) compared ($P < 0.0001$) to 16.7%, 1.7, and 10.7 DI, S, and DTFS, respectively, for on-site data.

DISCUSSION

Our study is the first comprehensive evaluation of the accuracy of site-specific weather estimation technology and its potential impact on performance of disease-warning systems. We focused on all of the most commonly used weather inputs to disease-warning systems, whereas previous reports (Gleason et al, 1997, 2002) focused primarily on LWD.

SkyBit estimated daily mean temperature more accurately than duration of periods with RH > 90%, wetness duration, and rainfall duration and amount. This accuracy is attributable to the property of temperature as a continuous variable, which can be estimated more accurately than a discrete (0 or 1) variable.

Spatial analysis of on-site and site-specific weather variables revealed that SkyBit estimates of daily mean temperature for a given location can be used reliably as temperature predictors for locations within a distance of ≤ 71 km. In the Midwest, therefore, SkyBit temperature estimates for a given site could exceed accuracy of measurements made at ground-stations further than 71 km from the site. Errors in simulated performance of BLITECAST and TOM-CAST using SkyBit data were mostly attributable to errors in estimation of hours of RH > 90% (BLITECAST) and LWD (TOM-CAST). The delayed spray thresholds when SkyBit data were input into BLITECAST resulted mainly from underestimation of hours of RH > 90%.

Underestimation of LWD by SkyBit was reflected in simulated performance of TOM-CAST, in which SkyBit data recommended, on average, fewer sprays than on-site data. Work on application of the CARTD/SLD/Wind speed model (Kim et al 2002) to reduce deviations of SkyBit DSVs from on-site DSVs in simulated TOM-CAST is ongoing. Preliminary results suggest that the model has potential in improving accuracy of wetness-dependent disease-warning systems.

SkyBit's consistent overestimation of rainfall amount, which resulted in higher disease incidence and more fungicide sprays compared to on-site measurements in simulated performance of Timmer and Zitko's (1996) rain-dependent model for prediction of post-bloom fruit drop of citrus, suggests that there is potential for economic loss if growers use SkyBit rainfall data in rain-dependent disease-warning systems.

In this study, we have demonstrated a high level of accuracy in SkyBit estimation of daily mean temperature for a given site and on a spatial scale. Errors in SkyBit estimation of duration of periods with RH > 90%, rainfall duration and amount, and LWD suggest a need to identify the sources of these errors and to refine algorithms for estimation of these variables. The CARTD/SLD/Wind speed model (Kim et al 2002) developed by our group is a significant step towards improvement of SkyBit accuracy in estimation of surface wetness duration. Development of similar models for correction of SkyBit errors in estimation of duration of periods with RH > 90% and occurrence, duration, and amount of rainfall would greatly enhance utility of SkyBit's site-specific weather estimates in disease-warning systems. This could in turn spur wider adoption and implementation of disease-warning systems with a concomitant increase in grower profits and environmental health resulting from reduced pesticide use.

LITERATURE CITED

Campbell, C. L., and Madden, L. V. 1990. Introduction to Plant Disease Epidemiology. John Wiley & Sons, New York.

Felland, C. M., Travis, J. W., Russo, J. M., Kleiner, W. C., and Rajotte, E. G. 1997. Validation of site-specific weather data for insect phenology and disease development in Pennsylvania apple orchards. PA. Fruit News, **77,** 10-17.

Gleason, M. L., MacNab, A. A., Pitblado, R. E., Ricker, M. D., East, D. A., and Latin, R. X. 1995. Disease-warning systems for processing tomatoes in eastern North America: are we there yet? Plant Dis., **79,** 113-121.

Gleason, M. L., Parker, S. K., Pitblado, R. E., Latin, R. X., Speranzini, D., Hazzard, R. V., Maletta, M. J., Cowgill, W. P., Jr., and Biederstedt, D. L. 1997. Validation of a commercial system for remote estimation of wetness duration. Plant Dis., **81,** 825-829.

Gleason, M. L., Taylor, S. E., Auraz, L. F., and Kim, K. S. 2002. Assessment of the accuracy of site-specific estimates or rainfall, air temperature, relative humidity, and wetness duration in Guanacaste Province, Costa Rica. Agronomia Mesoamericana, **in press.**

Gleason, M. L., Wegulo, S. N., Petit, E., and Batzer, J. C. 2000. Evaluation of site-specific weather data as inputs to disease-warning systems on watermelons and apples. (Abstr.) Phytopathology, **90,** S27.

Hubbard, K. G. 1994. Spatial variability of daily weather variables in the high plains of the USA. Agric. For. Meteorol., **68,** 29-41.

Kim, K. S., Taylor, S. E., Gleason, M. L., and Koehler, K. J. 2002. Model to enhance site-specific estimation of leaf wetness duration. Plant Dis., **86,** 179-185.

Krause, R. A., Massie, L. B., and Hyre, R. A. 1975. BLITECAST: a computerized forecast of potato late blight. Plant Dis. Reptr., **59,** 95-98.

Lau, Y. F., Gleason, M. L., Zriba, N, Taylor, S. E., and Hinz, P. N. 2000. Effects of coating, deployment angle, and compass orientation on performance of electronic wetness sensors during dew periods. Plant Dis., **84,** 192-197.

Madden, L., Pennypacker, S. P., and MacNab, A. A. 1978. FAST, a forecast system for Alternaria solani on tomato. Phytopathology, **68,** 1354-1358.

Magarey, R. D., Seem, R. C., Russo, J. M., Zack, J. W., Waight, K. T., Travis, J. W., and Oudemans, P. V. 2001. Site-specific weather information without on-site sensors. Plant Dis., **85,** 1216-1226.

Pscheidt, J. W., and Stevenson, W. R. 1986. Comparison of forecasting methods for control of potato early blight in Wisconsin. Plant Dis., **70,** 915-920.

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

Timmer, L. W., and Zitko, S. E. 1996. Evaluation of a model for prediction of postbloom fruit drop of citrus. Plant Dis. 80:380-383.

Wilks, D. S. 1995. Statistical Methods in the Atmospheric Sciences. Academic Press, San Diego.

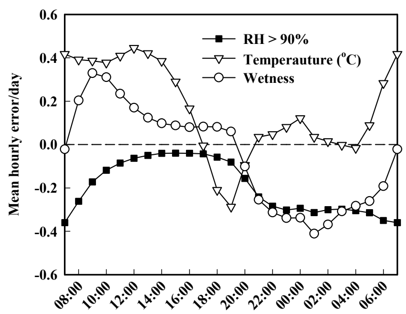

Fig. 2. Mean hourly SkyBit errors/day (SkyBit minus on-site) for three weather variables at 15 stations in Iowa, Illinois, and Nebraska (May to Sep 1997-1999).

Opposite: Fig. 3. Accuracy with which weather variable predictions were made for a given location based on SkyBit estimates or on-site measurements at a location within a distance of 1296 km. **A.** Temperature. **B.** Rainfall occurrence. C. RH > 90%. **D.** Leaf wetness duration.

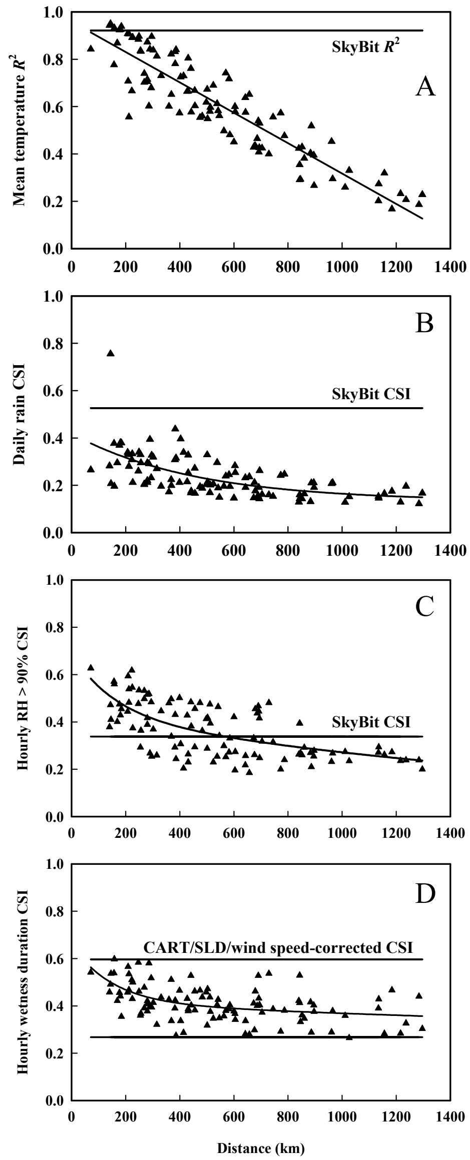

Table 1. Mean errors between measured (on-site) and SkyBit-estimated values for temperature, RH ≥ 90%, and leaf wetness duration from 15 sites in the Midwestern U.S., 1997-1999

	Temperature (°C) (daily mean)			Duration of periods with RH ≥ 90% (h/day)				Wetness duration (h/day)			
City	n[w]	ME[x] (°C)	SEM[y]	n	ME	SEM	MAE[z]	n	ME	SEM	MAE
Ames, IA	390	0.25	0.04	227	-3.19	0.24	3.66	358	-1.28	0.35	7.68
Lewis, IA	428	0.61	0.04	197	-1.27	0.29	3.03	381	1.65	0.28	6.69
Nashua, IA	429	0.16	0.03	253	-4.73	0.23	4.75	361	0.35	0.32	6.77
Sutherland, IA	428	0.32	0.04	293	-5.06	0.24	5.11	381	-1.69	0.33	7.15
Crawfordsville, IA	430	0.45	0.04	304	-4.38	0.23	4.46	366	-1.78	0.37	7.81
Bellville, IL	400	0.29	0.04	343	-7.51	0.20	7.55	267	-1.67	0.48	8.02
Bondville, IL	363	0.46	0.10	342	-5.37	0.21	5.46	327	-2.50	0.42	8.56
Dixon Springs, IL	354	1.15	0.09	319	-3.85	0.25	4.56	326	-1.91	0.32	7.25
Monmouth, IL	359	0.58	0.09	314	-6.76	0.28	6.79	322	-1.41	0.41	7.57
St. Charles, Il	358	0.79	0.13	314	-5.17	0.22	5.23	320	-2.12	0.39	7.49
Red Cloud, NE	458	-1.32	0.04	204	-1.92	0.17	2.04	389	-1.98	0.33	6.96
Gordon, NE	460	-0.67	0.04	366	-2.12	0.20	2.66	389	-2.27	0.34	6.98
O'neill, NE	406	-0.05	0.05	202	-3.55	0.28	3.65	321	-2.00	0.37	7.03
Sidney, NE	459	-0.12	0.04	401	-1.83	0.16	2.12	387	0.16	0.26	4.86
West Point, NE	461	0.29	0.04	307	-3.92	0.21	3.99	247	-2.98	0.36	7.77
All 15 sites	**6183**	**0.18**	**0.02**	**4386**	**-4.15**	**0.06**	**4.42**	**5142**	**-1.34**	**0.09**	**7.18**

[w]Number of 24-hour periods in the analysis.
[x]Mean error.
[y]Standard error of the mean
[z]Mean absolute error.