Perceptions of Risk, Risk Aversion, and Barriers to Adoption of Decision Support Systems and IPM
Presented at the Annual Meeting of The American Phytopathological Society July 31 to August 5, 2009, Portland, OR

Decision Aids for Multiple-Decision Disease Management as Affected by Weather Input Errors

W. F. Pfender, D. H. Gent, W. F. Mahaffee, L. B. Coop, and A. D. Fox

First and second authors: U.S. Department of Agriculture-Agricultural Research Service (USDA-ARS), Forage Seed and Cereal Research Unit, 3450 SW Campus Way, Corvallis, OR 97331; first, second, third, and fourth authors: Oregon State University Botany and Plant Pathology, Corvallis; third author: USDA-ARS Horticultural Crops Research Unit, Corvallis, OR; and fifth author: Fox Weather LLC, Fortuna, CA.

ABSTRACT

Pfender, W. F., Gent, D. H., Mahaffee, W. F., Coop, L. B., and Fox, A. D. 2011. Decision aids for multiple-decision disease management as affected by weather input errors. *Phytopathology* 101:644-653.

Many disease management decision support systems (DSSs) rely, exclusively or in part, on weather inputs to calculate an indicator for disease hazard. Error in the weather inputs, typically due to forecasting, interpolation, or estimation from off-site sources, may affect model calculations and management decision recommendations. The extent to which errors in weather inputs affect the quality of the final management outcome depends on a number of aspects of the disease management context, including whether management consists of a single dichotomous decision, or of a multi-decision process extending over the cropping season(s). Decision aids for multi-decision disease management typically are based on simple or complex algorithms of weather data which may be accumulated over several days or weeks. It is difficult to quantify accuracy of multi-decision DSSs due to temporally overlapping disease events, existence of more than one solution to optimizing the outcome, opportunities to take later recourse to modify earlier decisions, and the ongoing, complex decision process in which the DSS is only one component. One approach to assessing importance of weather input errors is to conduct an error analysis in which the DSS outcome from high-quality weather data is compared with that from weather data with various levels of bias and/or variance from the original data. We illustrate this analytical approach for two types of DSS, an infection risk index for hop powdery mildew and a simulation model for grass stem rust. Further exploration of analysis methods is needed to address problems associated with assessing uncertainty in multi-decision DSSs.

Decision support systems (DSSs) have been developed increasingly to help optimize disease management decisions (7–9). For some pathosystems, management strategy is concerned with a single dichotomous decision. For example, a management decision can be taken as to whether or not to apply a spray to a fruit crop at bud break, or whether to apply a preplant cultural control procedure. There has been extensive consideration of optimizing dichotomous management decisions, through the use of receiver operating characteristic (ROC) analysis (17,32). This analytical method can be used to evaluate the probability of making the correct decision and to identify management action thresholds that minimize long-term average costs. McRoberts et al. (19) expand the scope of the analysis to include subjective perceptions of risk in calculating the conditions under which a DSS will be beneficial and/or adopted. For many diseases, however, management requires more than one decision during the course of the season, such as evaluating the need and timing for several fungicide sprays. One management approach is to use a standard schedule, spraying every 2 weeks for example, but such a calendar-based strategy may be suboptimal in terms of number and timing of applications to minimize disease-related costs. Therefore, DSSs have been developed with the goal of improving or optimizing management decisions to reduce disease-related costs (pesticide costs, yield loss) or to achieve other production goals. For multiple-decision management, it typically is not a question of whether to apply management (i.e., whether or not there will be need for treatment), but a question of the time course of disease development, and therefore optimal timing and frequency of management intervention. A multiple-decision case can sometimes default to simple treat/no-treat case, but this choice occurs as one among several possible choices and at an undefined time during the season, unlike a dichotomous decision that will be taken at a predetermined time in the season or in response to a single risk assessment. In multi-decision situations there are degrees to the correctness of the management actions (not just correct and incorrect) and degrees to the favorability of outcomes from the DSS.

Corresponding author: W. F. Pfender; E-mail address: pfenderw@onid.orst.edu

doi:10.1094/PHYTO-05-10-0131

This article is in the public domain and not copyrightable. It may be freely reprinted with customary crediting of the source. The American Phytopathological Society, 2011.

Accuracy and precision of DSSs. Accuracy of a DSS could be defined as its probability of making the optimal management decision to achieve a desired outcome, such as minimizing cost related to management inputs and yield loss; precision is the repeatability of the outcome as indicated by the (reciprocal of) the variance. Estimating accuracy and precision of a DSS is important for determining its usefulness and likelihood of adoption by crop managers (1,19). Other papers (6,32) have addressed a more comprehensive definition of accuracy for a dichotomous decision context, using ROC analysis, or have extended this use to discrete decisions in a multi-decision context (3). For many multi-decision disease management situations, however, it may be impossible or impractical to determine whether a given decision is correct or incorrect, for reasons elaborated later in this paper. For multi-decision cases we can propose a working definition for model accuracy: correctly choosing the set of season-long management actions that minimizes disease-related costs (management costs plus yield loss).

There are several aspects of DSSs that affect their accuracy and precision, including the intrinsically variable and stochastic nature of biological systems as well as the accuracy of the disease model itself. An additional potential source of error in weather-based disease management DSSs is inaccuracy in the weather data inputs. Most DSSs use weather variables as major inputs, and the underlying disease models are commonly developed with reference (whether explicitly or implicitly) to highly accurate weather data such as would be obtained from on-site measurement. But on-site weather data may not always be available for implementing the DSS in practice, due to the cost of acquiring, maintaining, and managing data from weather stations in each production field. Therefore estimates or proxies of on-site weather data commonly are used (11,12), and these weather data will be inaccurate to some degree. In this paper, we address the effects of error in weather data on the performance of disease management DSSs. Although measurement error can be a significant problem, we limit our consideration in this paper to errors due to estimation of weather data from off-site sources.

Effects of weather data input errors on DSSs. The magnitude of effects on DSS due to weather data error depends partly on the magnitude of the errors, but also on the context in which the data are used. The nature of the underlying disease model may make any particular model more sensitive to a given weather variable than to others, and some models more sensitive than other models to any particular weather variable. For example, a particular disease index used as a DSS may be driven primarily by temperature with moisture inputs as a modifying variable, whereas a different disease model may be driven primarily or exclusively by rainfall or some other moisture measurement. Typically, temperature can be estimated more accurately than moisture. Also, the use of thresholds in a disease model can interact with weather data errors to determine the magnitude of DSS error. The disease model may be a rule-based index, for example, with a temperature threshold for declaring disease risk, i.e., the threshold value divides the outcome into two categories (+ or -) for disease development. If the actual temperature is near this threshold value, then a relatively small error in the weather data can cause the risk to be declared in the incorrect category, whereas an error of similar magnitude would be inconsequential when the actual temperature is not near the threshold. Two aspects of error in weather data are bias (i.e., a mean value for the weather component that is greater or less than the true value) and random variability around the measured mean. The effects of bias and variance (i.e., the importance of accuracy and precision) in the input weather data may differ from one another and may differ among disease models.

For dichotomous decision management, the ROC analysis context (17) can be used to evaluate the effects of weather data

errors. A bias will cause the calculated estimator to be greater (positive bias) or less than (negative bias) its true value, thereby over-predicting or under-predicting the need for action. The bias may be uniform, or it may be asymmetrical (e.g., biased at high temperatures but unbiased at intermediate or low temperatures). If the bias is known (both magnitude and sign), the error can be corrected by adjusting the input weather data, the indicator calculation and/or the threshold value. If the sign (+ or -) of the bias is unknown, however, the error cannot be corrected, and the result is that the error will decrease the probability of true positive and true negative decisions. If weather data errors are random, such that there is no bias, there will be no directional effect on predicting the need for management action, but there will be a decrease in the true positive proportion and true negative proportion (decrease in precision). That is, nonbiased input errors decrease the likelihood ratio for a positive decision and increase the likelihood ratio for a negative decision (17). As noted previously, the effect on model outcome will depend on how the particular predictor is calculated and whether the true conditions are near or far from the threshold value.

Characteristics of multi-decision DSSs. Before addressing the effect of weather data errors on multiple-decision DSSs, a consideration of the characteristics of multiple-decision disease management may be instructive. In multiple-decision management, a crop manager is dealing with a series of biological processes that are interconnected through time. The outcome of a particular biological event (e.g., infections occurring on a given day) is manifested only some time after the event, and is commonly obscured and influenced by other similar events that occur earlier and later (infections occurring on previous and subsequent days). This temporal overlapping of multiple events makes it difficult to assess the accuracy of a model's declaration (e.g., that many infections really did occur on a particular day as indicated by the index or model) and the effectiveness of the management action taken with respect to that event. Another characteristic of multiple-decision management is the opportunity for the manager to consider near-term and later actions, a situation referred to as "multi-stage problem with recourse" (14,31). Unlike dichotomous decision management, in which a decision is taken irrevocably and later assessed for its correctness (e.g., it may have been a true positive or a false positive decision), multiple-decision management allows for subsequent action that could, for example, mitigate an incorrect decision. An incorrect decision to take no action (false negative decision) can later be compensated for by taking action; an incorrect decision to take action on a given day (false positive decision) may result in a delayed need for the next required action. Furthermore, this opportunity to repeatedly reevaluate management requirements in light of new information becomes an essential part of determining an optimal management strategy. One must consider that the optimal strategy for today may be to wait for tomorrow's, or next week's, information before taking a management action. This "multi-stage problem with recourse" nature of the process leads directly to another characteristic, which is that there may be not a *single* optimal strategy but rather *several* different strategies that achieve the same result of minimizing disease-related costs. One consequence of nonunique solutions on the assessment of DSS accuracy is the difficulty of identifying the optimal outcome to use as the standard for comparison when assessing the outcome of the particular management strategy implemented. Approaches to the problem of quantifying accuracy or optimizing strategies for multistage processes have been described (14,21). A third characteristic, not unique to multi-decision management but often a significant complicating factor, is the use of nonquantitative (ordinal) outputs in many such DSSs. Some DSSs produce an output such as "high disease hazard" or "moderate disease

hazard", which prescribes an action from a graded series of responses such as shortening (or lengthening) the interval between successive fungicide applications, or choosing a particular type of fungicide over another. When the nondichotomous nature of these ordinal outputs is considered in combination with the occurrence of overlapping biological events, it is particularly difficult to evaluate the accuracy of the model. For example, what is the expectation (measurable disease quantity) if the model is correct in declaring three days of high disease hazard followed by 2 days of moderate disease hazard, and how can one ascertain the accuracy of the particular grade of response indicated? In some other areas of crop management, procedures have been suggested for incorporating ordinal information in evaluating the accuracy of decision making (28), but we are unaware of applications in plant disease management.

Another consideration that complicates the assessment of multi-decision DSSs is their dependence on management context. For a pathosystem to warrant management with multiple inputs, the crop must have sufficient economic value. Such high-value crops typically justify ongoing and repeated attention by the disease manager. In practice, a DSS in multi-decision management is used as only one of several information inputs, and the management strategy is repeatedly revised as informed by the DSS as well as by other information such as direct observations in the field (disease scouting), crop development (and associated considerations of ontogenetic resistance and vulnerability of the marketable product), and logistical constraints to management inputs and the farmer's objectives (20). As elaborated by McRoberts et al. (19), the context of the manager's subjective perception of disease hazard is another major determinant of whether and how the DSS is used. Therefore it is difficult to assess the accuracy of a DSS (its ability to produce the optimal strategy) apart from its context of actual use where it is a significant, but not isolated, decision input.

Weather data errors and multi-decision DSSs. Given the complexity of multi-decision management, the difficulty of quantifying DSS accuracy, and the range of DSS types and management situations, it is not possible to characterize the effects of weather data errors on performance of DSSs in general. One may be able to make general statements concerning the effect of weather data errors on a particular DSS for a given pathosystem. We are unaware of any reports of analysis to deduce input error effects on the very large population of possible scenarios (time courses in disease development, schedules of management intervention) for any given multi-decision disease management pathosystem. This objective is similar to the goal of error/uncertainty analysis in systems modeling, and although the problem has received little attention in plant pathology models it has been discussed in other disciplines such as ecology (16) and environmental engineering (2,29). The general problem is to quantify variation in model output caused by uncertainty in model inputs (16), and there are various approaches to address different facets of the problem. The procedures that have been described previously were applied to models constructed of continuous equations rather than logic rules. Sensitivity analysis is a procedure to quantify a model's relative sensitivity to its various input parameters (chosen to represent a defined range of values), and its results describe properties of the model rather than properties of the input data (16). First order error analysis (FOEA) uses variances of input parameters to estimate variance of the model output, with the assumption (not always applicable) that the functional relationships in the several model equations are linearly related to the output (2). FOEA can be useful if computing resources are limited, but with current availability of computing power it is less commonly used than Monte Carlo simulation (MCS) approaches (29). In MCS the model is run

many times, each run using different values of input parameters selected at random from a known probability distribution. With either FOEA or MCS the mean and variance (reflecting accuracy and precision) in model output can be characterized, and this information can be interpreted with respect to model outputs ("error analysis") and decisions taken ("decision analysis") (16). These approaches can be used for an analysis that is either forward-looking (what is the effect of a given input error on output accuracy?) or backward-looking (which input errors are responsible for a significant output inaccuracy?). The MCS can be made more applicable to a given modeling situation by including the actual uncertainty characteristics of the input parameter values used, rather than those from an arbitrary probability distribution.

The MCS approach is flexible and can be useful in plant disease DSS analysis, whether the DSS is a rule-based index or a simulation constructed of linked equations. The approach is an inductive method, in which one examines a large number of specific scenarios spanning a range of conditions, in order to discern patterns of input error effects on DSS performance. Inferring the performance of the DSS from a large number of specific cases can be most convincingly carried out in actual use, i.e., in field tests. One would compare outcomes (disease-related costs, including yield effects) in replicate treatments managed according to the DSS, operated either with on-site weather data or with weather data having known characteristics of error (bias, variance). One could assume that management using the high-quality input data produces the optimal outcome, or alternatively one could compare the outcomes using a different standard treatment such as a calendar-based spray schedule. Clearly, applying this approach to a large population of location-year cases would be prohibitively expensive. A different implementation of the inductive approach would be to use a modified Monte Carlo method: produce hypothetical scenarios of a pathosystem epidemic, derive the DSS-prescribed management actions that would be taken for each scenario given highest-quality weather data, then compare this output with management actions prescribed by the DSS using less-accurate weather data inputs. If we limit our question to the effect of input data errors on DSS performance, we can set aside the question of DSS accuracy and simply look for differences in management choices under the different data input situations. That is, we can assume that the DSS operated with the highest-quality weather data input provides the optimal management, then assess the extent to which lesser-quality weather data changes that management strategy (e.g., timing and number of fungicide applications). Large differences in management choices would be interpreted as large deviations from the optimal management strategy. If the DSS has the form of an index, it is possible to determine the number of management actions (fungicide applications) taken, but not necessarily the resultant disease severity or yield loss. If the DSS is based on a full-scale simulation model, it is possible also to quantify output differences in terms of disease severity and even yield loss if a damage function is known. In the following paragraphs we illustrate the inductive approach based on MCS with input values chosen from a distribution created to mimic plausible errors. We apply it to temperature inputs for two DSS examples that differ in the nature of their underlying disease models (an index and a simulation), and carry the process through to simulate management decisions.

Examples of temperature input error in DSSs. To produce the input weather data sets differing in fidelity to the original data, we applied selected levels of bias and variance to an actual data set collected during the growing season with an on-site datalogger (Campbell Scientific, Logan, UT). The on-site data set had hourly values for air temperature (1.5 m above grass). In the following discussion, we make the simplifying assumption that the measured data are error-free and representative values for temperature; al-

though the contribution of measurement errors is an important topic, we omit it in the context of this analysis.

To select levels of bias and variance to apply to the temperature data, we examined the accuracy of temperature estimates produced using previously developed online virtual weather data statistical evaluation tools (<http://uspest.org/virt/eval>) (5). To serve as a guide for plausible error levels to use in the present study, we determined temperature error rates for two types of estimated weather: site-specific weather forecasts downscaled from the National Weather Service Global Forecasting System using a mesoscale model and Fox Weather LLC's MtnRT custom software, and distance-elevation weighted regression data (5). For three locations near hop and grass seed growing areas in the Willamette Valley of Oregon, overall mean differences between observed and estimated hourly temperature ($^{\circ}\text{C}$) were approximately normally distributed and ranged from -0.7 to $+0.9$, with standard deviations of differences ranging from 1.5 to 6.1. From this information, we selected means (bias) ranging from -1 to $+1$, and standard deviations in the range 2 to 8 for the present error analysis. We also plotted observed and estimated temperature through time and noted that the bias of the estimates was not strictly systematic, and variance was not random. Rather, daily time course of estimated temperatures had the same general pattern as actual temperatures (rising from a daily minimum at about sunrise to a daily maximum in mid to late afternoon) with day-to-day variability in the accuracy of maximum, minimum and rate of change. To emulate such errors we started with the actual data and the estimated data for the same location. For each hour of the day (for example, 10 a.m., 11 a.m., etc.), the average signed difference over the whole season (estimated temperature minus actual temperature) was determined and added to the respective hourly actual temperatures for any given day to produce a smooth curve. Each reduced-fidelity data set was produced by applying a multiplier to the set of hourly differences for each half-day (10 a.m. to 10 p.m.) during the season. Multipliers were randomly selected from normal distributions with specified mean and variance. After producing a data set with modifications derived from a specific distribution, the population of hourly differences between original and modified data was analyzed for mean and variance. Distribution parameters for the randomly selected multipliers were adjusted as needed to achieve final data sets that met our requirements for mean and standard deviation of errors (differences from temperature values of the original data). Data sets were produced with the following hourly-difference characteristics (mean, standard deviation): 0.2; 0.4; 0.8; 1.2; 1.4; 1.8; -1.2 ; -1.4 ; -1.8 . Ten random-draw realizations (MCS) were produced for each characterization. Frequency histograms of the season-long population of hourly differences were similar in

appearance to those observed for the previously cited estimation (5).

To evaluate effects of bias and variance of temperature input data on the example DSSs, each DSS was run with the original on-site data as well as the 90 data sets with modified temperature (10 realizations for each of 9 bias $\times$ variance characterizations). This analysis is not intended to be exhaustive, nor definitive for all levels of error, but to illustrate an approach to the problem of characterizing effects of weather input errors using two example DSSs.

Hop powdery mildew index. The hop powdery mildew index, HOPS, is a modification of the Gubler-Thomas grape powdery mildew index (13), which has been used since 1994 in various modifications on grapes in California and other areas throughout the world (30). The HOPS index was initially developed from growth chamber experiments for temperature effects on infection severity, host susceptibility, and latent period, and adjusted empirically based on field observations taken in hop yards over several years in Oregon, Idaho, and Washington (18). The index uses inputs of temperature and rainfall (Box 1) to estimate a daily infection risk and infer speed of epidemic development (ordinal outputs of low, medium, and high risk). Based on the infection risk category, the index prescribes application intervals specific to different fungicides (based on mode of action) or biological control agents. In the Pacific Northwestern United States, 5 to 10 (or more) fungicide applications typically are made per season to manage hop powdery mildew, depending on cultivar susceptibility and geographic region. Mahaffee et al. (18) noted that growers who reported using the index made on average 1.2 fewer fungicide applications per season and had substantially less diseased hop cones at harvest compared with growers that did not report using the index.

This DSS presents some challenges for assessing accuracy, whether with high-quality or less-accurate weather data. It is difficult to translate the ordinal output of epidemic speed to a quantity that can be measured to assess accuracy. It is difficult to assess whether a given day's hazard estimate is correct, because epidemiological events occurring on that day will be modified and integrated with similar events occurring on previous and subsequent days. However, it is possible to apply the inductive approach described previously to examining the effect of weather data errors on index performance, by comparing index output from less-accurate weather data with the index output from the original on-site weather data.

The hop powdery mildew index recommended 10 fungicide applications when run with the original weather data set (Box 1; Fig. 1). The index outputs and the associated management recommendations were relatively insensitive to positive or negative

BOX 1

Hop powdery mildew index. The hop powdery mildew risk index begins for the season when shoots are 15 to 30 cm tall on at least 50% of the hills (groups of one or more plants). At 6 a.m. each day, the index uses hourly temperature and rain data from the previous 24 h period to calculate a series of five hierarchical rules:

- (i) If there are greater than six continuous hours above 30°C , then subtract 20 points.
- (ii) If there are greater than 2.5 mm rainfall, then subtract 10 points.
- (iii) If there are greater than six continuous hours above 30°C on the previous day, then no change in the index.
- (iv) If there are at least six continuous hours of temperatures between 16 and 27°C , then add 20 points.
- (v) If none of the above rules apply, then subtract 10 points.

If a rule is true, then none of the subsequent rules takes effect. Index values accumulate over time, with minimum and maximum values of 0 and 100. Index values of 0 to 30, 40 to 60, and 70 to 100 indicate conditions of low, moderate, or high infection risks, respectively, with corresponding treatment interval recommendations for a given fungicide or biological control agent.

temperature bias of the magnitude tested ($\pm 1^\circ\text{C}$) (Table 1). When the standard deviation was 2°C , temperature bias resulted in a similar (average difference of ≤ 0.3) number of fungicide applications recommended compared with the original weather data. The relatively small mean absolute difference in daily index values in these cases, as well as inspection of graphs of the index values over the season, indicated that trends in the index were generally similar over time. That is, seasonal effects on index outputs due to bias in temperature data were not apparent.

Variability in the input temperature had a greater effect on index performance than bias did, and these effects increased as the standard deviation in the weather data sets increased from 4 to 8°C . Under the highest variability scenario, the mean number of fungicide applications recommended was decreased by 1.9 to 2.5 compared with the original data set, and was associated with mean absolute difference in index values of >30 . The latter indicates substantial differences in trends in the index during certain times of the season, which had the overall effect of

FIGURE 1

Representative examples of outputs of the hop powdery mildew risk index when calculated with original on-site weather data (black line) or degraded temperature data (red line). In **A**, the erroneous temperature data had bias = 1°C and standard deviation = 2, and in **B**, it had bias = 0 and standard deviation = 8. The lines for the original and degraded temperature inputs are superimposed up to day of year 134 in A. Recommended fungicide application dates are shown in color-coded triangles at the bottom of each figure assuming a 14-, 10-, or 7-day application interval for index values of 0 to 30, 40 to 60, and 70 to 100, respectively.

making the index more conservative (i.e., fewer recommendations to spray). These differences in index performance were particularly evident on days when the temperature in the original data set was near or slightly above 30°C for 4 to 6 h per day, the high temperature threshold specified in Rules 1 and 3 of the index.

As discussed previously, hypothetical experiments with this type of index do not permit analysis of disease development under the various scenarios. And hypothetical experiments obviously do not allow for adjusting management strategy based on continual

observation of disease development as would be the practice in actual use.

Stem rust simulation model. The DSS for stem rust of perennial ryegrass seed crops (Box 2) is based on a simulation model. The model was constructed from field and controlled-environment experiments quantifying the effects of temperature and moisture on probability of infection (23) and duration of latent period (22). The model runs on a daily time-step and includes components for host plant growth, intra-plant spread (24), ontogenic resistance and fungicide effects (26). The DSS has an action threshold for estimated total infections (visible plus latent). It has been in development and testing for approximately 8 year and is implemented as a publicly-available internet site. In 5 years of extension field trials at two different sites/year it always produced an outcome at least as favorable (lower disease-related cost, including yield loss) as the standard practice, and produced a more favorable outcome than the standard practice in 7 of 10 location-years (27). The epidemic model produces a quantitative output (disease severity each day) which can be compared directly with measured disease to compute a quantitative value for accuracy. The disease severity output can also be used to estimate yield loss, based on a damage function relating seed yield to healthy area duration for the 3 weeks centered on anthesis (25).

The stem rust DSS prescribed three fungicide applications (on days-of-year 118, 126, and 141) when run with the original weather data (Box 2: Table 2 and Fig. 2A). With the low-variance temperature data, three fungicide applications were recommended regardless of the bias of the data set. With higher-variance input data, there were some data sets for which the DSS recommended two or four applications, but three applications were recommended in most cases (Table 2). The timing of the sprays did differ, with average spray date diverging increasingly from the original-data days as the season progressed. The bias in the input data set had no consistent effect on mean or range of spray date, but the range of dates among the 10 realizations of each data set types did increase (i.e., precision decreased) with the highest-variance inputs (Table 2, analysis not shown).

Because the stem rust DSS is based on a simulation model, it was possible to simulate the effects of altered spray dates on disease development. To produce these analyses, the fungicide application decisions that were derived from each altered weather data set were applied to the epidemic simulated with true weather data. That is, we simulated epidemic development that would occur with observed weather when fungicide treatments were applied according to the DSS run with erroneous weather data. Ninety such simulations were made (10 realizations of each of nine types of degraded weather data), and the mean and range of disease severities in these simulations were calculated (Fig. 3A).

TABLE 1

Management recommendations and deviations associated with the hop powdery mildew risk index for a single location-year, using temperature inputs that vary from measured temperature observations

Weather data (bias, standard deviation) ^a	Number of applications ^b	Mean interval ^b	Mean absolute index difference ^c
0,0	10	7.5	...
0,2	10.2	7.4	8.5
1,2	9.7	7.8	12.6
-1,2	9.8	7.7	11.6
0,4	9.5	7.9	16.7
1,4	9.3	8.1	17.3
-1,4	9.6	7.8	14.2
0,8	7.8	9.8	34.9
1,8	7.5	10.2	40.7
-1,8	8.1	9.4	31.8

^a The original on-site weather data are labeled 0,0 (bias = 0, standard deviation = 0). Season-long hourly temperatures in other weather data sets differ from the original in average value and/or variability. There were 10 randomly created realizations of each data set type.

^b The number of fungicide applications and mean interval was determined for 15 April to 14 July, assuming a 14-, 10-, or 7-day application interval for index values of 0 to 30, 40 to 60, and 70 to 100, respectively.

^c Mean absolute index difference was calculated from daily deviation between the index values obtained from the original temperature data and the index values obtained from the degraded weather data sets.

BOX 2

Stem rust simulation model.

Weather input: Hourly temperature and leaf wetness

Calculations: Daily infection probability (temperature × wetness duration for nighttime and early morning)
Fungus latent period, daily plant growth (daily heat unit accumulation)
Fungicide effects (function of time and plant growth stage)

These calculations are components of a simulation model that estimates disease severity from several interacting processes of pathogen activity, plant growth and fungicide effects.

Outputs: Daily total disease (includes latent infections as well as visible disease)
Fungicide spray dates (intersection of total disease with action threshold)
Daily visible disease and area under disease progress curve for damage-critical time period

Figure 2B is an example of one of the most divergent outputs among the 90 cases (a data set with bias = 0 and standard deviation = 8). Three fungicide applications, the third very late, were prescribed when the DSS was run with this degraded weather data (Fig. 2B). When these three application dates (based on inaccurate weather data) were used in the epidemic simulated with the original (correct) weather data inputs, simulated disease exceeded the action threshold, i.e., was greater than would be desired for best management (Fig. 2C).

The 90 outcomes were further used to estimate disease-related costs (Fig. 3B), by applying the damage function (25), the price of the agricultural product (grass seed) and costs of fungicide application. An average bias of $\pm 1^\circ\text{C}$ in data with low variance had relatively little effect on economic outcomes of using the DSS, whereas most data set types with the medium or high levels of variance had more instances of suboptimal performance as measured by disease area and disease-related costs. It is noteworthy, however, that the mean disease level and costs did not differ greatly from the high-quality weather data case for any of the degraded data set types. As with the powdery mildew index, the ability to adjust management strategy based on continuing observation of disease development in the field is necessarily missing from this simulation analysis. In actual use, such adjustments would be an integral part of DSS use and would tend to mitigate management errors to some extent.

In these particular cases of DSSs and temperature input data (Boxes 1 and 2), variance in the data produced greater differences in outcome than bias did, at least for the ranges of variance and bias that we used based on what we considered to be typical of errors in air temperature estimation. These examples are meant to be illustrative of an approach; data sets with different biases and variances could be used instead. Also, we note that this analysis does not address the accuracy of the disease models themselves (which requires field observations and methods for quantitative comparison between estimated and observed disease), but rather the influence of weather data input errors on the performance of the model and DSS. The approach could also be applied to errors

in other weather inputs, such as leaf wetness, rainfall, or humidity. Moisture variables are typically more difficult than temperature to measure or estimate accurately, and we would thus expect larger uncertainty effects with moisture-driven models than with temperature-driven models.

Summary and conclusions. For dichotomous decision systems, analytical methods previously have been developed and refined to estimate probability of making correct decisions, to optimize this probability and to incorporate economic analysis based on this information (17,19). Assessing probability of correct decisions in multiple-decision systems of plant disease management adds another layer of complexity. It is possible to apply dichotomous decision ROC analysis to component decisions in a multi-decision management problem if assumptions are made regarding ability to observe consequent effects of current events (3,10). Carisse et al. (3) used ROC analysis to evaluate DSS indicators of onion botrytis blight for congruence with disease development 2 to 3 days later, in deciding the first or succeeding fungicide application. If the DSS is based on a model that produces quantitative outputs, the multi-decision process may be amenable to optimization using genetic algorithms (21), a stochastic optimization procedure that computationally mimics random recombination and fitness selection of genes on a chromosome to optimize a sequence of events (15).

One major difficulty in conducting a realistic analysis of multi-decision DSSs is that, in actual use, the DSS commonly is taken as only one (albeit an important one) of several tools to aid in making the management decision. Furthermore, the DSS calculations are commonly subject to revision based on new information as the season progresses. For these reasons, assessment of multi-decision DSS accuracy probably requires field testing under actual-use conditions, with a somewhat arbitrary choice of management strategy selected as the standard for comparison.

Despite these problems it is possible to assess the impact of weather data input errors on performance of the DSS, setting aside the question of DSS accuracy and analyzing simply

TABLE 2
Spray dates specified by rust simulation model operated for a single location–year,
using temperature input data sets that vary from original data

Weather data (bias, standard deviation) ^a	Mean number of applications (min, max) ^b	Mean date of application (min, max) ^c		
		First application	Second application	Third application
0,0	3	118	126	141
0,2	3.0 (3,3)	115 (105,118)	130 (119,140)	158 (140,172)
+1,2	3.0 (3,3)	116 (114,118)	135 (132,138)	163 (156,172)
–1,2	3.0 (3,3)	115 (111,118)	134 (122,140)	158 (151,170)
0,4	3.0 (3,3)	114 (111,118)	132 (125,137)	158 (144,171)
+1,4	2.8 (2,3)	115 (114,117)	133 (126,139)	163 (152,169)
–1,4	3.0 (3,3)	114 (111,119)	132 (122,140)	157 (139,171)
0,8	3.0 (3,3)	115 (113,117)	133 (126,141)	161 (149,172)
+1,8	3.0 (2,4)	114 (107,120)	133 (123,138)	163 (144,177)
–1,8	2.9 (2,3)	112 (109,115)	132 (126,173)	160 (144,173)

^a The original on-site weather data are labeled 0,0 (bias = 0, standard deviation = 0). Season-long hourly temperatures in other weather data sets differ from the original in average value and/or variability. There were 10 randomly created realizations of each data set type.

^b Number of fungicide applications prescribed by the model as run with the indicated weather data type. Average, minimum, and maximum of 10 runs.

^c Application date expressed as day of the year (0 to 365). The simulation model was run 10 times for each weather data type; entries are the average, earliest, and latest dates prescribed by the model for each set of 10 runs.

the sensitivity of DSS output to variations in input. We have illustrated this approach with two different types of DSS. In these examples we observe that an average bias of -1 to $+1^{\circ}\text{C}$ (typical of estimated air temperature data) has a relatively minor effect on DSS performance, whereas typically observed variance around the average bias can cause significant degradation of DSS performance. More realizations of input data sets, and a greater range of biases and variances, would be needed for a thorough analysis. With this type of inductive approach to error analysis, relative importance of uncertainty among different weather variables (e.g., air temperature, leaf wetness duration, relative humidity, and rain-fall) for DSS performance could be determined. We have illustrated the error propagation effects with only a single set of weather data, and therefore cannot draw general conclusions about the effects of input bias and variance on DSS performance. We intend only to raise the issue of uncertainty analysis for DSSs and propose one method to explore it.

Although this approach provides information for any type of DSS, it is more informative for DSSs that produce quantitative output than for those that produce ordinal output. Also, a DSS with an output that is amenable to calculation of disease severity permits estimation of economic results, i.e., variance of economic outcomes due to weather data input errors (Box 2). This economic variance information is important in assessing the likelihood of adoption and use of a DSS because risk perception and avoidance may be related to the variance (precision) of economic

outcomes as much as to the average (accuracy) (1). Economic information allows also for quantitative analysis of optimal management (19).

Recognition of the problems associated with assessing accuracy in multi-decision DSSs may lead to further exploration of analysis methods by other researchers. One challenge is to apply concepts and analyses of perceived risk and economic cost ratio (19) to multi-decision systems. For example the cost ratio described in McRoberts et al. (19) as the cost of a false positive relative to cost of a false negative decision could be reframed as the relative cost of different treatment regimes, such as starting a multi-date fungicide treatment regime earlier than needed versus later than needed. The effect of errors in DSSs (whether due to inaccurate weather inputs or other variables) on the user's subjective perception of damage probability may be another fruitful area of investigation, and could contribute to our understanding of overall adoption patterns for DSSs (9). Another area for exploration is the effectiveness of DSSs to improve the alignment of growers' subjective beliefs (credence) in the chance of an epidemic for a particular *relative* epidemic severity with the true chance of that severity (see 19). As uncertainty analyses are developed in plant disease DSSs, it would be beneficial to consult discussions in other disciplines, and to consider some approaches that include qualitative (4) or Bayesian (29) treatments of uncertainty analysis. In multi-treatment pathosystems, understanding the timing and relative severity of an epidemic is instrumental in implementing

FIGURE 2

Stem rust simulation used as a decision support system. **A**,

Simulation run with original on-site weather data shows estimated active disease and total infections (visible plus latent infections). Simulation includes effects of fungicide applications that are made on days the total infection line meets the threshold. The value of healthy area duration between days 160 and 180 is correlated with yield damage to the crop. **B**,

Simulation run with one of the 10 weather data sets having bias = 0 and standard deviation = 8. Fungicide decisions (arrows) based on this input data were different from those of the original data simulation. **C**, Fungicide application dates determined from the inaccurate weather (Fig. 3B) were applied to a simulation in a run with the original (correct) weather data. This procedure simulates the outcome expected when spray decisions based on inaccurate weather data are used. The diseased area between days 160 to 180 is larger in this case than it is in the simulation with decisions based on accurate weather data.

FIGURE 3

Outcomes of using management decisions (fungicide application dates) derived from degraded weather data in a simulated stem rust epidemic. The outcome of the simulation with original data is labeled 0,0. Other bars are labeled with the season-long bias and standard deviation of weather data sets from which the fungicide dates were derived. **A**, Area under the disease progress curve (AUDPC) for the yield critical period of days 160 to 180 in simulated stem rust epidemics. **B**, Disease-related costs (yield loss and fungicide application costs) derived from the simulations. Box plots show the median (line in open boxes), middle 50% of the data (open box), extremes of the data points (whiskers), and outliers (solid circles).

the most beneficial timing and sequence of treatments consistent with the farmer's and societal objectives, which is the ultimate goal.

ACKNOWLEDGMENTS

The use of trade, firm, or corporation names in this publication is for the information and convenience of the reader. Such use does not constitute an official endorsement or approval by the United States Department of Agriculture or the Agricultural Research Service of any product or service to the exclusion of others that may be suitable. Financial support was provided by USDA-ARS CRIS projects 5358-22000-035-00 and 5358-21000-035-00. We thank N. McRoberts, L. Madden, and an anonymous reviewer for their suggestions on an earlier draft of the manuscript. We also thank C. Thomas, C. Daly, and other members of the Western IPM Center Work Group on Weather Systems for their insights, comments, and discussions on this topic.

LITERATURE CITED

- Annou, M. M., Wailes, E. J., and Thomsen, M. R. 2005. A dynamic decision model of technology adoption under uncertainty: Case of herbicide-resistant rice. *J. Agric. Appl. Econ.* 37:161-172.
- Bobba, A. G., Singh, V. P., and Bengtsson, L. 1996. Application of first-order and Monte Carlo analysis in watershed water quality models. *Water Resour. Manag.* 10:219-240.
- Carisse, O., McRoberts, N., and Brodeur, L. 2008. Comparison of monitoring- and weather-based risk indicators of Botrytis leaf light of onion and determination of action thresholds. *Can. J. Plant Pathol.* 30:442-456.
- Chen, C. F., Ma, H., and Reckhow, K. H. 2007. Assessment of water quality management with a systematic qualitative uncertainty analysis. *Sci. Total Environ.* 374:13-25.
- Coop, L., Fox, A., Mahaffee, W., Gent, D., Pfender, W., Daly, C., Thomas, C., and Jepson, P. 2009. Forecast and virtual weather driven plant disease risk modeling system. (Abstr.) *Phytopathology* 99(suppl.):S24.
- Dewdney, M., Biggs, A. R., and Turechek, W. W. 2007. A statistical comparison of the blossom blight forecasts of MARYBLTY and Cougarblight with receiver operating characteristic curve analysis. *Phytopathology* 97:1164-1176.
- DeWolf, E. D., and Isard, S. A. 2007. Disease cycle approach to plant disease prediction. *Annu. Rev. Phytopathol.* 45:203-220.
- Fabre, F., Plantagenest, M., and Yuen, J. 2007. Financial benefit of using crop protection decision rules over systematic spraying strategies. *Phytopathology* 97:1484-1490.
- Gent, D. H., DeWolf, E. D., and Pethybridge, S. J. 2011. Perceptions of risk, risk aversion, and barriers to adoption of decision support systems and integrated pest management: An introduction. *Phytopathology* 101:640-643.
- Gent, D. H., and Ocamb, C. M. 2009. Predicting infection risk of hop by *Pseudoperonospora humuli*. *Phytopathology* 99:1190-1198.
- Gent, D. H., and Schwartz, H. F. 2003. Validation of potato early blight disease forecast models for Colorado using various sources of meteorological data. *Plant Dis.* 87:78-84.
- Gleason, M. L., Duttweiler, K. B., Batzer, J. C., Taylor, S. E., Sentelhas, P. C., Monteiro, J. E., and Gillespie, T. J. 2008. Obtaining weather data for input to crop disease-warning systems: Leaf wetness duration as a case study. *Sci. Agric.* 65:76-87.
- Gubler, W. D., Rademacher, M. R., Vasquez, S. J., and Thomas, C. S. 1999. Control of powdery mildew using the UC Davis powdery mildew risk index. Published online. *APSnet Feature Story*. The American Phytopathological Society, St. Paul, MN.
- Hoganson, H. M., and Smith, E. L. 1989. Recognizing uncertainty and the sequential nature of decisions in forest management planning. Pages 338-342 in: *Proc. 1989 National Convention Society of American Foresters*. Society of American Foresters, Spokane, Washington.
- Holland, J. H. 1975. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*. University of Michigan Press, Ann Arbor, MI.
- Jager, H. I., and King, A. W. 2004. Spatial uncertainty and ecological models. *Ecosystems* 7:841-847.
- Madden, L. V. 2006. Botanical epidemiology: Some key advances and its continuing role in disease management. *Eur. J. Plant Pathol.* 115:3-23.
- Mahaffee, W. F., Thomas, C. S., Turechek, W. W., Ocamb, C. M., Nelson, M. E., Fox, A., and Gubler, W. D. 2003. Responding to an introduced pathogen: *Podosphaera macularis* (hop powdery mildew) in the Pacific Northwest. Published Online. *Plant Health Progress* doi:10.1094/PHP-2003-1113-07-RV.
- McRoberts, N., Hall, C., Madden, L. V., and Hughes, G. 2011. Perceptions of disease risk: From social construction of subjective judgments to rational decision making. *Phytopathology* 101:654-665.
- Norton, G. A. 1976. Analysis of decision making in crop protection. *Agro-Ecosystems* 3:27-44.
- Parsons, D. J., and Te Beest, D. 2004. Optimising fungicide applications of winter wheat using genetic algorithms. *Biosyst. Eng.* 88:401-410.
- Pfender, W. F. 2001. A temperature-based model for latent period duration in stem rust of perennial ryegrass and tall fescue. *Phytopathology* 91:111-116.
- Pfender, W. F. 2003. Prediction of stem rust infection favorability, by means of degree-hour wetness duration, for perennial ryegrass seed crops. *Phytopathology* 93:467-477.
- Pfender, W. F. 2004. Role of phenology in host susceptibility and within-plant spread of stem rust during reproductive development of perennial ryegrass. *Phytopathology* 94:308-316.

25. Pfender, W. F. 2009. A damage function for stem rust of perennial ryegrass seed crops. *Phytopathology* 99:498-505.
26. Pfender, W. F., and Eynard, J. 2009. Field assessment of a model for fungicide effects on intraplant spread of stem rust in perennial ryegrass seed crops. *Phytopathology* 99:696-703.
27. Pfender, W. F., Mellbye, M. E., Silberstein, T. B., and Gingrich, G. A. 2009. Five year field evaluation summary of the USDA stem rust model for perennial ryegrass. Pages 26-28 in: 2008 Seed Production Research at Oregon State University, USDA-ARS Cooperating. Department of Crop and Soil Science Ext/CrS. 128.
28. Plant, R. E. 1997. A methodology for qualitative modeling of crop production systems. *Agric. Sys.* 53:325-348.
29. Reckhow, K. H., and Chapra, S. C. 1999. Modeling excessive nutrient loading in the environment. *Environ. Pollution* 100:197-207.
30. Thomas, C. S., Skinner, P. W., Fox, A. D., Greer, C. A., and Gubler, W. D. 2002. Utilization of GIS/GPS-based information technology in commercial crop decision making in California, Washington, Oregon, Idaho, and Arizona. *J. Nematol.* 34:200-206.
31. Thompson, E. F., and Haynes, R. W. 1971. A linear programming probabilistic approach to decision-making under uncertainty. *For. Sci.* 17:224-229.
32. Yuen, J. E., and Hughes, G. 2002. Bayesian analysis of plant disease prediction. *Plant Pathol.* 51:407-412.