USING CROSS-VALIDATION TO EVALUATE CERES-MAIZE YIELD SIMULATIONS WITHIN A DECISION SUPPORT SYSTEM FOR PRECISION AGRICULTURE

K. R. Thorp, W. D. Batchelor, J. O. Paz, A. L. Kaleita, K. C. DeJonge

ABSTRACT. *Crop growth models have recently been implemented to study precision agriculture questions within the framework of a decision support system (DSS) that automates simulations across management zones. Model calibration in each zone has occurred by automatically optimizing select model parameters to minimize error between measured and simulated yield over multiple growing seasons. However, to date, there have been no efforts to evaluate model simulations within the DSS. In this work, a model evaluation procedure based on leave-one-out cross-validation was developed to explore several issues associated with the implementation of CERES-Maize within the DSS. Five growing seasons of measured yield data from a central Iowa cornfield were available for cross-validation. Two strategies were used to divide the study area into management zones, one based on soil type and the other based on topography. The decision support system was then used to carry out the model calibration and validation simulations as required to complete the cross-validation procedure. Results demonstrated that the model's ability to simulate corn yield improved as more growing seasons were used in the cross-validation. For management zones based on topography, the average root mean squared error of prediction (RMSEP) from cross-validations was 1460 kg ha^{-1} when two growing seasons were used and 998 kg ha^{-1} when five years were used. Model performance was shown to vary spatially based on soil type and topography. Average RMSEP was 1651 kg ha^{-1} on zones of Nicollet loam, while it was 496 kg ha^{-1} on zones of Canisteo silty clay loam. Spatial patterns also existed between areas of higher RMSEP and areas where measured spatial yield variability was related to topography. Changes in the mean and variance of optimum parameter sets as more growing seasons were used in cross-validation demonstrated that the optimizer was able to arrive at more stable solutions in some zones as compared to others. Results suggested that cross-validation was an appropriate method for addressing several issues associated with the use of crop growth models within a DSS for precision agriculture.*

Keywords. *Corn, Crop model, Cross-validation, Decision support system, Precision agriculture, Spatial variability, Yield.*

The basic premises of precision agriculture are that the characteristics and processes of agricultural systems vary over space and time and that a better understanding of this variation can lead to improvements in management of crop production inputs. In the past decade, process-oriented crop growth models are one of many tools that have been implemented to understand the spatial and temporal variability that occurs naturally within agricultural systems. Fundamentally, crop models have been used to identify the factors that limit crop yield and cause observable spatial yield variability (Batchelor et al., 2002). Specific yield-limiting factors that have been studied using crop models include water stress, nitrogen stress, soybean (*Glycine max* (L.) Merr.) cyst nematode, and weeds (Paz et al., 1998; Paz et al., 1999, Paz et al., 2001b, Paz et al., 2002). With an understanding of the causes of spatial yield variability, crop modeling techniques can be extended to develop variable-rate management strategies that optimize either crop yield or producer economic returns spatially across the field. Methodologies for precision management based on crop model simulations have been developed for plant population and variety selection in soybeans (Paz et al., 2001a; Paz et al., 2003) and for nitrogen fertilizer applications (Paz et al., 1999) and irrigations (DeJonge et al., 2007a) in corn (*Zea mays* L.). In the special case of precision nutrient management, crop model simulations are also useful for understanding the spatial and temporal roles of soil nutrients in both crop production and environmental quality. Using the results of spatial simulations for yield and nitrogen left behind over many seasons of historical weather, Thorp et al. (2006) developed precision nitrogen management strategies for reducing the environmental impacts of corn production. Other applications of crop growth models in precision agriculture include yield forecasting (Hodges et al., 1987; Liu et al., 1989), yield gap analysis (Paz et al., 2004), and simulation of crop response to climate change or genetic modification (Boote et al., 1996).

Submitted for review in September 2005 as manuscript number BE 6097; approved for publication by the Biological Engineering Division of ASABE in June 2007. Presented at the 2005 ASABE Annual Meeting as Paper No. 053002.

The authors are **Kelly R. Thorp, ASABE Member Engineer,** Agricultural Engineer, USDA-ARS Arid-Land Agricultural Research Center, Maricopa, Arizona; **William D. Batchelor, ASABE Member Engineer,** Professor and Head, Department of Agricultural and Biological Engineering, Mississippi State University, Mississippi State, Mississippi; **Joel O. Paz, ASABE Member Engineer,** Public Service Assistant, Department of Biological and Agricultural Engineering, University of Georgia, Griffin, Georgia; **Amy L. Kaleita, ASABE Member Engineer,** Assistant Professor, Department of Agricultural and Biosystems Engineering, Iowa State University, Ames, Iowa; and **Kendall C. DeJonge, ASABE Member Engineer,** Hydraulic Engineer, U.S. Army Corps of Engineers, Omaha, Nebraska. **Corresponding author:** Kelly R. Thorp, USDA-ARS, 182 ALARC, 21881 N Cardon Lane, Maricopa, AZ 85238; phone: 520-316-6375; fax: 520-316-6330; e-mail: kthorp@uswcl.ars.ag. gov.

Recent work at Iowa State University has focused on the use of the DSSAT crop growth models (Jones et al., 2003) to answer precision agriculture questions at the sub-field-level scale (Batchelor et al., 2002). Utilization of crop growth models in this way has posed many interesting questions and challenges in regards to the details of the implementation, the feasibility of the approach, and the validity of the simulation results. Fields must first be divided into subunits or management zones, each representing an independent region for implementation of the one-dimensional crop growth model. However, it is unclear what criteria are appropriate for dividing the field into zones. Several different agricultural system characteristics, including soil type (Franzen et al., 2002), topography attributes and soil electrical conductivity (Fraisse et al., 2001b), and yield history (Hornung et al., 2006), have been explored as the basis for management zone delineation. A second challenge of applying crop growth models in precision agriculture has been the need for an appropriate calibration strategy for tuning the model to zone-specific conditions. To facilitate this goal for corn in Iowa, Garrison et al. (1999) incorporated a tile drainage routine into the CERES-Maize crop growth model (Jones and Kiniry, 1986), and Paz et al. (1999) described a calibration approach for adjusting two water balance parameters, the saturated hydraulic conductivity (K_{SAT}) of the bottom soil layer and the effective tile drainage rate, to estimate spatially the effects of soil water dynamics on corn yield. Third, a strategy is needed for evaluating crop model simulations in each management zone using datasets independent of those used for calibration. To date, efforts to use crop models for precision agriculture applications in Iowa have been performed using only the calibration strategy developed by Garrison et al. (1999) and Paz et al. (1999). No strategies for model evaluation using independent datasets have yet been investigated. A fourth challenge involves the feasibility and impracticality of using a crop model for precision agriculture when the user is required to manually calibrate, evaluate, and apply the model independently within each management zone. Depending on the strategy for delineation of management zone boundaries, a single field may contain hundreds or thousands of zones that, under the design constraints of our approach, would each require independent model simulations. In a practical sense, this limitation would greatly reduce the ability to generate timely prescriptions for precision management of crop production inputs. Thus, a decision support system (DSS) has been designed to automate the processes of management zone delineation, model calibration, model evaluation, and application of the model to answer questions in the area of precision agriculture (Batchelor et al., 2004; DeJonge et al., 2007b). The DSS incorporates geographic information system (GIS) tools for management zone delineation, algorithms to develop model input files for spatial simulations (Thorp et al., 2005), an optimization routine for implementing a model calibration strategy, and algorithms for evaluating and applying the model across management zones.

To continue the development of crop growth modeling for applications in precision agriculture, the overall objective of this study was to investigate strategies for evaluating crop model simulations within the framework of our DSS. A common strategy for evaluating agricultural systems models in other research has been to partition the measured dataset into two groups: one group for model fitting during the calibration phase, and the other for model testing during the evaluation phase (Garrison et al., 1999; Zhao et al., 2000; Bakhsh et al., 2001). In these more classical studies, modelers typically adjust several parameters manually, and an important aspect of model evaluation is the modeler's subjective comparison of measured and simulated data. In addition, in these kinds of studies, models have usually been applied to plot-level sites, and researchers have rigorously controlled and measured various aspects of the agricultural system to allow for a thorough evaluation of all model components. Our application of crop growth models is different from this approach in several ways. First, we intend to apply the model in production fields where measured data for describing many aspects of the agricultural system, especially the soil system, will be limited or nonexistent, and our simulations will therefore be subject to several uncertainties. Second, to compensate for these uncertainties, we have developed a model calibration strategy that implements an optimization routine to automatically adjust several soil hydrologic parameters by minimizing error between measured and simulated crop yield over multiple growing seasons (Paz et al., 1999). Thus, beyond baseline adjustment of the model for Iowa conditions (Garrison et al., 1999), we do not rely on any other manual adjustment of model parameters based on subjective comparisons of measured and simulated data at a field site. Third, because we are optimizing the model based only on historical measurements of crop yield, we assume that the model will be better calibrated for a study site if we include as many seasons of yield information as possible in the optimization. Furthermore, we assume that, when only a few seasons of yield information are available, leaving one or two seasons out of the optimization may have significant effects on the final optimization result. On the other hand, if we use all the yield measurements in the optimization, then there exists no independent dataset for an unbiased model evaluation. Because of the above issues, our modeling application requires an approach that is different from the "two-group partition method" typically used for evaluating agricultural systems models.

Leave-one-out (LOO) cross-validation (Efron and Gong, 1983; Efron and Tibshirani, 1998) is a statistical procedure that is suitable for evaluating model simulations in the context of our applications in precision agriculture. Since collection of measured yield data is limited by the frequency with which a crop of interest is grown at a site, the number of seasons of measured data available for use in our model calibration procedure has typically been small. LOO cross-validation is especially useful in cases where measurements are limited, because it permits all available measured data to be iterative and exhaustively used for both model calibration and model validation. In addition, the resulting estimate of model predictive performance is more reliable than estimates from the "two-group partition method" and less biased than estimates derived from calibration-dependent datasets (Jones and Carberry, 1994). One drawback of using this technique to evaluate crop growth models is that it complicates the selection of model parameters to be used for post-evaluation applications. LOO cross-validation has been previously used for evaluation of crop growth models in the work of Jones and Carberry (1994) and Irmak et al. (2000). The main objective of this study was to develop a procedure that implements LOO cross-validation to evaluate model simulations within the context of our DSS for precision agriculture. Specifically, the evaluation procedure was developed to assess how model predictive performance responds to the use of additional yield measurements for model calibration, to compare model simulations spatially across management zones, and to understand the reliability of the optimum parameter sets generated by the model calibration procedure. Thus, the evaluation procedure can be used to characterize how many growing seasons of measured data are "enough" for an adequate model calibration, to identify causes of spatial variability in model performance across a study site, and to make inferences about what optimum parameter sets should be used for future applications of the model. A secondary objective was to demonstrate the model evaluation procedure using CERES-Maize within the DSS to simulate five seasons of corn production at a study site in central Iowa.

Materials and Methods

Simulation Environment

Researchers at Iowa State University have recently developed a DSS to facilitate the use of the DSSAT crop growth models for applications in precision agriculture (Batchelor et al., 2004; DeJonge et al., 2007b). One feature of the DSS is a set of GIS tools that can be used to establish management zone boundaries, extract information from soil surveys and yield monitor data, and format crop model input files for spatial simulations (Thorp et al., 2005). These GIS tasks are used to establish a simulation environment in which the one-dimensional DSSAT crop growth models are applied independently in each management zone, and the simulation output for each zone is used to develop prescriptions for site-specific management. After preparing the simulation environment for a field site of interest, an optimization routine within the DSS is used to calibrate the model. The optimization routine is based on the simulated annealing algorithm as described by Corana et al. (1987) and implemented by Goffe et al. (1994). Model parameters are adjusted in order to minimize the root mean square error (RMSE) between measured and simulated yield within each management zone over multiple growing seasons. Using the DSS, up to ten model parameters can be uniquely optimized for several hundred management zones at one time. After model calibration is complete, the DSS stores the optimized parameter values for each management zone in a database, where they can be retrieved when needed for model evaluation and application to questions in precision agriculture. Algorithms are available in the DSS for generating site-specific recommendations for nitrogen fertilizer (Paz et al., 1999; Thorp et al., 2006) and irrigation (DeJonge et al., 2007a) in corn and for plant population and variety selection in soybean (Paz et al., 2001a; Paz et al., 2003). Given the large amount of simulations necessary to apply crop growth models in precision agriculture, issues of practicality would have prevented these studies from occurring without the existence of this DSS.

Development of the DSS is on-going, and there are currently some noteworthy limitations in our applications of crop growth models in precision agriculture. First, the DSS is not currently able to alternate between different crop models, so the effects of crop rotations cannot be simulated. Second, the DSS does not currently simulate continuously over multiple growing seasons, so the model is reinitialized at the beginning of each year. Although this limitation does not allow the model state variables to be tracked over multiple years, independent simulation of growing seasons was a design constraint for development of the model optimization routine. It also facilitates the resampling of growing seasons for model calibration, as required to carry out LOO cross-validation. Third, since the DSS was intended to be used for studying precision management in a production setting, it has been designed with the understanding that detailed measurements of the agricultural system would be unavailable. We assumed that yield monitor data would be the only information measured locally at a study site, and even that has typically been limited to just a few growing seasons. Soil characteristics would be derived from a county survey. Weather data would be either measured locally or obtained from the nearest National Climatic Data Center (NCDC) site, and management information would be provided by the producer. Since measured data for making a detailed calibration of all model components are unavailable, we use our known quantity, yield, to automatically optimize a few select model parameters and artificially compensate for the remaining uncertainties in the model.

Crop Growth Model

Data used in this investigation came from a production cornfield in Iowa; thus, the CERES-Maize crop growth model (Jones and Kiniry, 1986) was implemented within the DSS for this study. CERES-Maize is a computer program that utilizes carbon, nitrogen, and water balance principles to simulate the processes that occur during the growth and development of corn plants within an agricultural system. The model calculates growth and development of corn plants within a homogeneous area on a daily time step, and the final crop yield is computed on the date of harvest. Inputs required for model execution include management practices (plant genetic coefficients, plant population, row spacing, planting and harvest dates, and fertilizer application amounts and dates), environmental factors (soil type, drained upper limit, lower limit, and saturated hydraulic conductivity), and weather conditions (daily minimum and maximum temperature, solar radiation, and precipitation). Garrison et al. (1999) calibrated CERES-Maize for conditions in Iowa and modified the model for simulating tile drainage in the water and nitrogen balances. This modification was necessary for simulating corn growth on the tile-drained soils of the Midwestern U.S., and it enabled the development of our strategy for optimizing model parameters across management zones. In addition, similar to the CROPGRO-Soybean model (Calmon et al., 1999a), the version of CERES-Maize used within the DSS was modified in-house for simulating the effect of saturated soil water conditions on root growth distribution in the soil profile, which is important for simulating crop growth on soils that sometimes exhibit perched water tables. CERES-Maize has been widely used to simulate the collective effect of plant genetics, management practices, weather, and soil conditions on the growth, development, and yield of corn plants, and the model has been shown to perform adequately on plot-level, field-level, and regional scales for a wide variety of corn hybrids, climatic conditions, and soil types around the world (Hodges et al., 1987; Carberry et al., 1989; Liu et al., 1989; Jagtap et al., 1993; Pang et al., 1998; Garrison et al., 1999; Paz et al., 1999; Fraisse et al., 2001a).

Model Calibration Strategy

The model calibration strategy used within the precision agriculture DSS has been developed and tested for both corn (Garrison et al., 1999; Paz et al., 1999) and soybeans (Paz et al.,

1998) in Iowa. The strategy is based on the hypothesis that wet spring weather leads to high soil moisture contents in the soils of Iowa. High soil moisture leads to higher water tables in some areas of the field, which restricts maximum crop rooting depth. In addition, upward redistribution of water may cause oxygen depletion in soil layers with existing roots, which leads to root senescence. Later in the summer, the water table recedes as precipitation diminishes. Crops in areas of the field where maximum rooting depth was limited now have difficulty finding water and thus exhibit water stress during the critical grain filling period. Variable water stress across the field then leads to spatial yield variability. The common Midwestern practice of tile drainage further complicates the water table dynamics in production fields. Tile drains have been installed throughout the agricultural sector in the Midwestern U.S. for 150 years (Urban, 2005), but only recently have the locations of newly installed tile drains been mapped and recorded (Allred et al., 2004). Thus, the existence, location, functionality level, and hydrologic impact of tile drains in many production fields in Iowa is largely unknown, which increases the difficulty of simulating this process.

Given these assumptions about the hydrology of agricultural fields in Iowa, a procedure has been developed for optimizing two model parameters, including the K_{SAT} of the bottom soil layer (190 to 210 cm in depth) and the effective tile drainage rate, to calibrate models for simulating hydrologic effects on crop yield in production fields (Paz et al., 1998; Paz et al., 1999). These parameters govern the movement of water through the soil profile and can be used to estimate the effect of impermeable soil layers and tile drainage on water table, soil water, and nitrogen dynamics (Garrison et al., 1999). Model calibration gives small values for the deep-layer K_{SAT} parameter (cm day^{-1}) when a soil exhibits poor drainage. This causes water to move more slowly through the bottom soil layer, and a water table may form and restrict root growth as it moves upward in the soil profile. If the soils within a grid cell are well-drained, then model calibration will result in a large value for the deep-layer K_{SAT} parameter. In this case, excess water is more quickly lost out the bottom of the profile and water tables are kept low or never form, allowing roots to grow deep in the soil profile. The effective tile drainage rate (day^{-1}) is simulated as the fraction of water per day that leaves the soil layer at the specified tile depth when a water table exists above the tile drain. Adjustment of this parameter allows for simulation of appropriate water contents in the root zone soil layers above the tile drain and attempts to account for the largely unknown component of tile drainage in the hydrologic balance.

With limited measured data to describe the agricultural system and limited time for manual calibration of the model across management zones, simulated annealing optimization has been implemented within the precision agriculture DSS to automatically carry out the calibration procedure. During optimization, the two model parameters are adjusted independently for each management zone to minimize the RMSE between measured and simulated crop yield over multiple growing seasons. For our crop modeling applications, RMSE serves as the objective function to be minimized during simulated annealing optimization and can be defined as:

$$\text{RMSE}_i = \left(\frac{1}{n} \cdot \sum_{j=1}^{n} \left(Ym_{i,j} - Ys_{i,j} \right)^2 \right)^{0.5} \tag{1}$$

where $Ym_{i,j}$ is the measured yield and $Ys_{i,j}$ is the simulated yield in the ith management zone for the jth of the n seasons of measured yield data used to calibrate the model.

Other researchers have used similar strategies to calibrate crop growth models. For example, Liu et al. (1989) adjusted phenological coefficients for maize until the simulated dates for silking and maturity closely matched observed dates. Jones and Carberry (1994) adjusted potential kernel number and potential kernel growth rate to minimize error between measured and simulated corn yield. A strategy similar to ours has also been employed to calibrate crop models for agricultural systems having highly restrictive claypan soil layers (Fraisse et al., 2001a). Other researchers have also implemented optimization algorithms that calibrate crop models by solving for a parameter set that maximizes or minimizes an objective function (Jones and Carberry, 1994; Calmon et al., 1999b; Irmak et al., 2001). Typically, the RMSE, or related error statistics, is the objective function to be minimized during the optimization (Kobayashi and Salam, 2000; Gauch et al., 2003).

Model Evaluation Strategy

Because the number of growing seasons of measured yield information available for use in the model calibration procedure has typically been small, efforts to evaluate the model with calibration-independent data have to date been neglected. To address this issue, a model evaluation strategy based on the LOO cross-validation statistical technique (Efron and Gong, 1983; Efron and Tibshirani, 1998) was developed to assess model performance within the context of our DSS for precision agriculture. Three important issues need to be addressed by the evaluation. First, since multiple growing seasons of measured yield information are required to calibrate the model, it is necessary to evaluate how the model performance changes as additional growing seasons of yield information become available for inclusion in the optimization. Second, since the DSS was designed to apply the model spatially across management zones, it is interesting to evaluate how model performance changes between zones and to identify what is causing the model to perform better in some zones as compared to others. Third, since the overall goal of model calibration and validation is to determine what parameters should be used for future applications of the model, it is necessary to understand the reliability of the optimized parameter sets generated during model calibration. To address these issues, measured data from a study site in central Iowa was used to develop and test a model evaluation strategy based on LOO cross-validation.

Cross-Validation

Given a set of measured yield data from n growing seasons, LOO cross-validation requires the crop model to be calibrated and independently validated n times. For the jth of n growing seasons, the model calibration procedure using simulated annealing optimization is performed using data from the n – 1 other growing seasons, leaving out the data from the jth growing season. After optimizing the model with the jth growing season left out, the calibrated model is used to simulate the jth growing season. This process is repeated until data from all n growing seasons has been left out and used for model validation one time. The LOO cross-validation estimate of model predictive performance is then calculated as the root mean square error of prediction (RMSEP) between measured and simulated yield for the n independent model validation simulations (fig. 1). In the context of this study, RMSEP can be defined as:

$$\mathrm{RMSEP}_i = \left(\frac{1}{n} \cdot \sum_{j=1}^{n} \left(Ym_{i,j} - Ys_{i,-j} \right)^2 \right)^{0.5} \qquad (2)$$

where $Ym_{i,j}$ is the measured yield value for the ith management zone in the jth of the n growing seasons used in LOO cross-validation, and $Ys_{i,-j}$ is the simulated yield value in the ith management zone obtained using the optimum parameters from a calibration with the jth growing season left out. By repeating the LOO cross-validation procedure in each management zone, a map of model performance can be generated for the study area.

Evaluation Procedure

Assuming management practices remain constant from year to year, climate variability, especially precipitation differences, are known to drive temporal yield variability in corn and soybeans (Jaynes and Colvin, 1997). Since our model calibration procedure utilizes temporal yield measurements over multiple growing seasons to optimize soil hydrologic parameters within management zones, we hypothesize that the ability of the calibrated model to simulated independent growing seasons will depend on the degree to which the climatic conditions for growing seasons used in the calibration are representative of the possible range of conditions at the site. Furthermore, since each growing season of measured yield data provides a unique example of crop response to climate, model performance should generally improve as the number of growing seasons available for calibration increases. To investigate this issue, a simple strategy was designed that utilizes LOO cross-validation to analyze all possible subsets of measured data from available growing seasons. The strategy implements LOO cross-validation within a framework of combinatorial statistics. According to statistical theory, the number of combinations of n distinct objects taken r at a time, or "n choose r," is:

$$\binom{n}{r} = \frac{n!}{r!(n-r)!} \qquad (3)$$

for $r = 0, 1, 2, ..., n$. Applying this theory to our modeling application, n represents the total number of available growing seasons of measured data, and r represents the number of growing seasons with which the model will be tested using LOO cross-validation. To evaluate how the model performs when only two growing seasons of data are available, LOO cross-validation is first applied for all possible combinations of measured data from two growing seasons, that is, "n choose 2" times. This gives "n choose 2" unique values for RMSEP, each of which represents the estimate of model performance when only one growing season is available for model calibration and one for model validation. The same analysis can then be carried out for all possible combinations of three growing seasons, that is, "n choose 3" times. Since three growing seasons are used now, each LOO cross-validation procedure follows the example shown in figure 1. This results in "n choose 3" unique values for RMSEP, each of which represents the LOO cross-validation estimate of model performance when two growing seasons are used for calibration and one for model validation. This process can then be repeated, incrementing r by one each time, up to n. At that point, "n choose n" is one, the LOO cross-validation procedure can only be applied one time, and only one estimate of RMSEP can be generated (table 1). This procedure allows for an examination of how model performance changes as additional growing seasons of measured yield data become available for model optimization. In particular, it examines how the soil parameters generated by the calibration procedure affect the simulation of calibration-independent growing seasons as additional growing seasons are used to calibrate the model.

Applying the LOO cross-validation procedure spatially allows for an analysis of model performance across management zones. We hypothesize that spatial variability in model performance will result from two main limitations in our attempt to optimize parameters to simulate water dynamics in the soil system. First, values for the parameters governing soil water retention and conductivity are obtained from soil survey information and research literature rather than being measured directly. Thus, the model may have greater difficulty simulating yield on the soil types where the soil survey or literature values do not approximate actual soil properties very well. Second, because of the model's one-dimensional nature, it may have greater difficulty simulating spatial yield variation that is due to field topography. For example, although spatial redistribution of water according to topography has been linked to spatial yield variability (Marques da Silva and Alexandre, 2005), the hydrologic balance of our model does not simulate any surface or subsurface run-on or

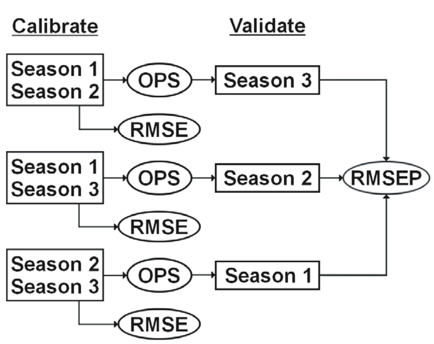

Figure 1. Example of leave-one-out cross-validation as applied with three growing seasons of measured data. Optimization of the model three separate times leaving out one season each time gives three optimum parameter sets (OPS), which are used to validate the model three times using the growing season left out.

Table 1. Possible combinations of growing seasons when five seasons of measured yield data are available. Leave-one-out cross-validation is applied separately to each combination.[a]

	2 Years	3 Years	4 Years	5 Years
1	GS 1 & 2	GS 1, 2, & 3	GS 1, 2, 3, & 4	GS 1, 2, 3, 4, & 5
2	GS 1 & 3	GS 1, 2, & 4	GS 1, 2, 3, & 5	
3	GS 1 & 4	GS 1, 2, & 5	GS 1, 2, 4, & 5	
4	GS 1 & 5	GS 1, 3, & 4	GS 1, 3, 4, & 5	
5	GS 2 & 3	GS 1, 3, & 5	GS 2, 3, 4, & 5	
6	GS 2 & 4	GS 1, 4, & 5		
7	GS 2 & 5	GS 2, 3, & 4		
8	GS 3 & 4	GS 2, 3, & 5		
9	GS 3 & 5	GS 2, 4, & 5		
10	GS 4 & 5	GS 3, 4, & 5		

[a] GS = growing seasons.

run-off between neighboring management zones. In addition, since weather conditions are assumed to be uniform across the field, the model does not simulate any spatial yield variability that may arise from hillshade effects on incoming solar radiation (Bennie et al., 2006). If model error is mainly due to issues with simulating the soil system, then an appropriate strategy for assessing model performance spatially would be to divide the field into management zones based on either soil type or topography and to implement the model evaluation procedure within each zone. The resulting spatial patterns of RMSEP would be useful for identifying the locations where the model has the greatest difficulty simulating yield and for determining whether the locations relate spatially to soil type and/or topography.

Since LOO cross-validation is applied to all possible combinations of measured growing seasons, several unique sets of optimized parameters are generated during the calibration phase of the model evaluation procedure. As the number of growing seasons used for model calibration increases, the behavior of the resulting parameter values can provide further insight into model performance and the reliability of the calibration. In this study, optimized parameter sets in each management zone were grouped according to the number of growing seasons used in the calibration. The mean and variance of the parameter sets were then used to study how the parameters behaved as additional seasons of measured information were used to calibrate the model. We expected the mean parameter values to stabilize as the number of growing seasons increased, and we expected the variance between parameter sets to grow smaller. This result would indicate that the additional growing seasons were helping the optimizer to consistently converge on parameter sets having roughly the same value, and the calibration procedure would therefore be more reliable in that management zone. Since optimized parameters in this study existed as a set of two, multivariate statistical techniques were used to describe the parameter variation by computing bivariate confidence ellipses around the sample means for parameter sets in each management zone (Johnson and Wichern, 2002). The confidence region around the sample mean of the two-dimensional parameter set, as applied to the ith management zone in this two-parameter study, is the ellipse determined by all (μ_1, μ_2) such that:

$$n \begin{bmatrix} \bar{x}_{i,1} - \mu_1 \\ \bar{x}_{i,2} - \mu_2 \end{bmatrix}^T \mathbf{S}_i^{-1} \begin{bmatrix} \bar{x}_{i,1} - \mu_1 \\ \bar{x}_{i,2} - \mu_2 \end{bmatrix} \le \frac{p(n-1)}{(n-p)} F_{p,n-p}(\alpha) \quad (4)$$

where p is the number of parameters, n is the sample size of parameter sets, $\left(\bar{x}_{i,1}, \bar{x}_{i,2}\right)$ is the set of sample means for the two parameters, $F_{p,n-p}(\alpha)$ is the value of the F-distribution at α on p and $n - p$ degrees of freedom, and $\mathbf{S}_i^{-1}$ is the inverse of the 2×2 covariance matrix for the parameter estimates in the ith management zone. Given n sets of parameters, $\mathbf{S}_i$ is defined as:

$$\mathbf{S}_i = \frac{1}{n-1} \begin{bmatrix} \sum_{j=1}^{n} (x_{i,j,1} - \bar{x}_{i,1})^2 & \sum_{j=1}^{n} \begin{pmatrix} (x_{i,j,1} - \bar{x}_{i,1}) \\ \times (x_{i,j,2} - \bar{x}_{i,2}) \end{pmatrix} \\ \sum_{j=1}^{n} \begin{pmatrix} (x_{i,j,1} - \bar{x}_{i,1}) \\ \times (x_{i,j,2} - \bar{x}_{i,2}) \end{pmatrix} & \sum_{j=1}^{n} (x_{i,j,2} - \bar{x}_{i,2})^2 \end{bmatrix} \quad (5)$$

where ($x_{i,j,1}$, $x_{i,j,2}$) is the jth set of two parameter estimates in the ith management zone. As a measure of the variation between the bivariate parameter sets, the generalized variance was computed, which is simply the determinant of the covariance matrix $\mathbf{S}_i$. All confidence ellipses in this study were computed using an α level of 0.05.

Application

Measured data for testing the model evaluation strategy were available within a 20.25 ha study area of a production cornfield near Perry, Iowa (41.93080° N, 94.07254° W). To apply the strategy within the context of our DSS for precision agriculture, the GIS component of the DSS was first used to develop management zones for the study area. Two separate management zone maps were created, one based on soil type and the other based on topography. Management zones based on soil type were developed from a digitized soil survey of the study site. Five primary soil types, Canisteo silty clay loam, Clarion loam, Nicollet loam, Harps loam, and Okoboji silty clay loam, were present across the study area, and 20 unique management zones were established based on these soil types (fig. 2a). In each zone, the soil profile was simulated in 15 layers to a depth of 210 cm, and estimates of the physical properties for each soil type in each layer were obtained from two sources. The saturated hydraulic conductivity (K_{SAT}, cm day^{-1}), bulk density (BD, g cm^{-3}), soil pH, and soil texture in each soil layer were obtained from the county soil survey (USDA-SCS, 1981), and Ratliff et al. (1983) provided estimates for drained upper limit (DUL, cm^3 cm^{-3}) and lower limit (cm^3 cm^{-3}) based on soil texture. Saturated moisture content (SAT, cm^3 cm^{-3}) was calculated from BD using:

$$\text{SAT} = 0.92 * \left(1 - \frac{\text{BD}}{2.65}\right) \quad (6)$$

Management zones based on topography were developed from an interpolated contour of elevation data collected with a real-time kinetic (RTK) global positioning system (GPS) receiver at the site. Elevation ranged from 299 m to 306 m above sea level, with a maximum slope of approximately 2%. A 1 m contour of elevation data resulted in 14 unique management zones based on topography (fig. 2b). The DSS GIS component was used to determine which soil type covered the largest area of each topographic management zone, and soil properties were assigned to each zone based on its dominant soil type.

Since soil properties were not measured at the site, appropriate initial conditions for soil water content and nutrient levels were assumed and assigned uniformly to each management zone. Initial soil water content was set to 0.3 cm^3 cm^{-3}, a value just below the DUL for the soils in the field. Initial nutrient levels were set arbitrarily to 0.1 g elemental N, P, and K per Mg soil. For the purpose of this study, it was assumed that the soil profile contained only a negligible amount of nutrients at the beginning of the season, and that spring-applied fertilizer applications served to raise the nutrient concentrations to levels that would support plant growth. Since growing seasons are simulated independently within the DSS, these initial conditions were used to start the simulation of every growing season. Plant population was assigned uniformly to each management zone based on the average of population measurements that were collected in the 1996 growing season, and default cultivar coefficients for

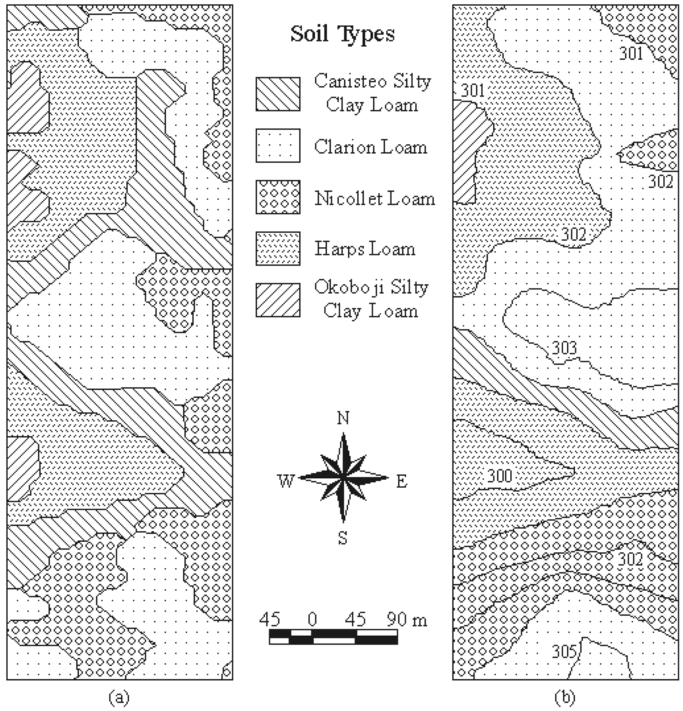

Figure 2. Management zones were established based on (a) soil type from a county soil survey and (b) topography from measured elevation data (m). For model simulations, dominant soil types were assigned to topographic management zones based on the type that covered the largest area in the zone.

a 2650 to 2700 growing degree day corn variety were used. Model inputs for management practices, including planting date, harvest date, and fertilizer application rates and dates, were set according to the producer's actual production practices. Weather data, including solar radiation, maximum and minimum daily temperature, and precipitation amount, were collected daily from a weather station directly at the site.

Five growing seasons of corn yield information were available at the study site. Corn yield was measured using a yield monitor on a grain combine at the conclusion of the 1994, 1996, 1998, 2000, and 2002 growing seasons. The GIS component of the DSS was used to clip the measured yield information for each year according to management zone, to compute the average yield in each zone, and to prepare the yield input files necessary for spatial simulations using CERES-Maize within the DSS. Given that five growing seasons of measured information were available, the proposed model evaluation strategy was carried out by applying LOO cross-validation for all the combinations of five growing seasons, as shown in table 1. An extra calibration procedure was also carried out to determine the optimum parameter sets when all five seasons were used to calibrate the model, meaning none were left for validation. Based on the number of ways to combine the measured growing seasons (eq. 3), model calibration procedures required for LOO cross-validation resulted in five, ten, ten, five, and one optimum parameter set(s) for combinations of one, two, three, four, and five growing seasons, respectively. The LOO cross-validation procedures resulted in five, ten, ten, and five RMSEP estimates of model performance for combinations of two, three, four, and five growing seasons, respectively. Optimum parameters sets and RMSEP values for each combination were obtained across all 20 management zones based on soil type and all 14 management zones based on topography.

Results and Discussion

Model Performance with Additional Seasons

As the number of growing seasons of measured data used for LOO cross-validation increased from two to five, the average RMSEP for the various growing season combinations generally tended to decrease. Management zones based on topography showed the greatest range in average RMSEP, with a value of 1460 kg ha^{-1} when only two growing seasons were used for LOO cross-validations and 998 kg ha^{-1} when five growing seasons were used (table 2). For management zones based on soil type and topography, the coefficients of determination (r^2) between number of growing seasons used for cross-validations and average RMSEP were 0.86 and 0.73, respectively. Average RMSEP trends were linearly decreasing as the number of growing seasons available for LOO cross-validations increased (fig. 3). These results demonstrate that as more growing seasons of measured data become available for use in a cross-validation procedure, the ability of the model to simulate growing seasons independent of the calibration generally improves. By the definition of LOO cross-validation (fig. 1), increasing the number of growing seasons available for the procedure means that more growing seasons will be simulated in the optimization procedures used to calibrate the model. When the parameters from those optimizations are then used to validate the model and the resulting RMSEP is lower than that from using fewer growing seasons in the procedure, it can be concluded that the model's ability to simulate yield in independent growing seasons has improved.

Table 2. Root mean square error of prediction (RMSEP) results for cross-validations based on two through five seasons of measured data across soil type and topography management zones.[a]

	Soil Type Zones		Topography Zones	
Seasons	Average (kg ha^{-1})	SD (kg ha^{-1})	Average (kg ha^{-1})	SD (kg ha^{-1})
Two	1378	1468	1460	1421
Three	1392	1293	1244	1011
Four	1271	958	1334	1120
Five	1171	630	998	539

[a] SD = standard deviation.

In addition to the average, the variation of RMSEP from cross-validations also tended to decrease as the number of growing seasons included in the procedure increased from two to five. For management zones based on soil type, the standard deviation of RMSEP across zones was 1468 kg ha^{-1} when two growing seasons were used in LOO cross-validations, and it was 630 kg ha^{-1} when five growing seasons were used in the procedure (table 2), and a similar trend was seen across management zones based on topography. In addition to the standard deviation, the inter-quartile range and total range of RMSEP were visibly shown to decrease as the number of growing seasons used in LOO cross-validation procedure increased (fig. 3). This demonstrates that, as more growing seasons of measured information become available for LOO cross-validation and more seasons are therefore used in the optimization routine, the resulting calibrated model is able to simulate yield in independent growing seasons within a narrower expected range. When only two growing seasons of measured information are available for LOO cross-validation, and only one season is therefore used to calibrate the model, the optimization procedure may overfit the parameters to the conditions of the calibration seasons. The calibrated model may then simulate yield very poorly in the independent seasons, as indicated by the upper-bound RMSEP over 7000 kg ha^{-1} for the two growing season case (fig. 3). However, when five growing seasons are used for LOO cross-validation, the model is calibrated based on the conditions of four growing seasons, and thus it is more reliably able to simulate yield in independent seasons within a narrower range of expected RMSEP values. An interesting result was also that the lower-bound RMSEP was slightly greater as more growing seasons were used for LOO cross-validation. When only two seasons are used, chances are greater that two seasons having similar weather patterns may be encountered, and thus the calibrated model for one season is able to simulate the other season with low error between measured and simulated yield. However, as more seasons are used for cross-validation, it is likely that weather patterns between seasons will be more different. The increased lower limit for RMSEP then arises due to the imperfections in the model as it is used to simulate a wider range of conditions.

Reducing the average and standard deviation of RMSEP as additional growing seasons were used in the cross-validation procedure was an expected result. However, the results were only able to provide partially conclusive evidence about how many growing seasons are "enough" for an adequate model calibration. With the understanding that the model is a simplified approximation of what actually happens in the agricultural system, model simulations will never be perfect. Thus, RMSEP cannot decrease indefinitely,

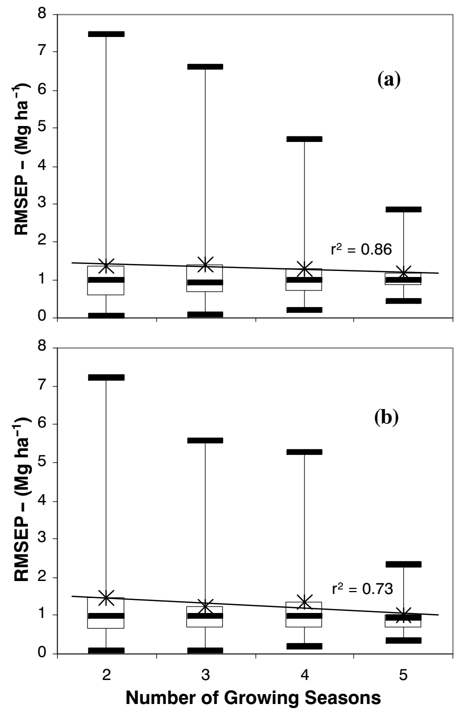

Figure 3. Box plot of values for root mean square error of prediction (RMSEP) as the number of seasons of measured data available for cross-validation increases. Data were aggregated across management zones by (a) soil type and (b) topography.

and we might expect the average and standard deviation of RMSEP values to stabilize as the number of growing seasons used in the model evaluation procedure increases beyond five. At that point, the number of growing seasons required for an adequate calibration could be determined. Based on the results from the data currently available, it can be concluded that using LOO cross-validation to calibrate and validate the model based on five growing seasons results in a better performing model than using two, three, or four growing seasons. However, it is possible that using more than five seasons of measured information could provide a model that continues the trend of decreasing RMSEP and better performance in simulating calibration-independent datasets.

Model Performance Spatial Variability

Since the LOO cross-validation procedures were performed independently for each management zone, maps of RMSEP show how the model performance varied spatially depending on soil type (fig. 4a) and topography (fig. 4b). Since zones of higher and lower RMSEP tend to cluster together, there is evidence that spatial patterns in model performance exist across the field. Since RMSEP in each zone represents the ability of calibrated models to simulate yield in multiple independent growing seasons (eq. 2), an understanding of the spatiotemporal pattern of yield variability at the site, as it relates to soil type and topography, is necessary to provide insight into why the model performance varied as it did across these management zones.

For the LOO cross-validation procedure involving measured datasets from all five growing seasons, the mean RMSEP values for zones having the Nicollet, Okoboji, and Clarion soil types were 1651, 1079, and 1059 kg ha^{-1}, respectively. Lower mean RMSEP values of 926 and 496 kg

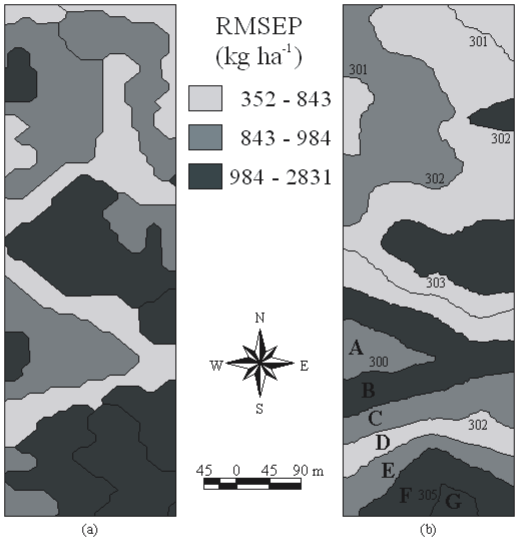

Figure 4. Spatial variability of RMSEP across management zones by (a) soil type and (b) topography for the cross-validation procedure using all five growing seasons.

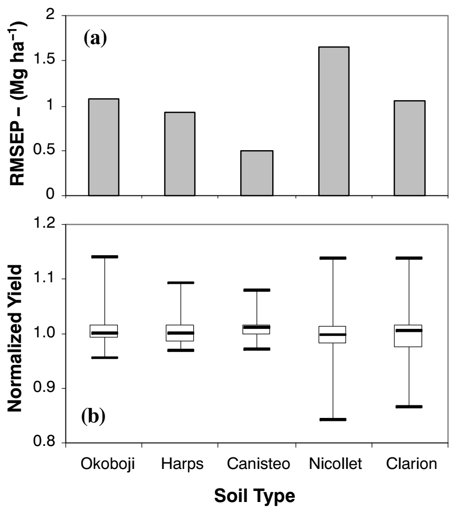

Figure 5. (a) Mean RMSEP from cross-validation on five growing seasons across soil types and (b) variability of measured yield, normalized by the annual field-level average yield, across growing seasons and soil types are both greater on Nicollet, Okoboji, and Clarion soils than on Harps and Canisteo soils.

Table 3. Measured corn yield according to soil type and growing season.[a]

Soil Type	Measured corn yield (kg ha^{-1})					
	1994	1996	1998	2000	2002	SD
Okoboji	11204	8820	9755	7571	10685	1458
Harps	10937	8941	9706	7716	10457	1275
Canisteo	10898	9051	9928	7541	10471	1331
Nicollet	10481	9114	9945	7554	10535	1241
Clarion	10802	9091	9928	7502	10330	1297
SD	261	122	113	82	129	

[a] SD = standard deviation.

ha^{-1} were found for zones having the Harps and Canisteo soil types (fig. 5). Thus, the calibrated model had more difficulty simulating yield on the Nicollet, Okoboji, and Clarion soils than on the Harps and Canisteo soils. An analysis of the spatiotemporal patterns of measured corn yield variability across the study site helps improve the understanding of these results. Averaging corn yield measurements across zones of the same soil type for each of the five growing seasons demonstrated that temporal yield variability across growing seasons was more significant than spatial yield variability across soil types in a single season. The standard deviation of measured corn yield across growing seasons ranged from 1241 to 1458 kg ha^{-1} depending on soil type, and the standard deviation of measured corn yield across soil types ranged from 82 to 261 kg ha^{-1} depending on the growing season (table 3). Thus, to be able to appropriately compare temporal yield variability spatially between zones, the effect of year–to-year variability first needed to be removed. To do this, corn yield measurements for each soil type management zone in each season were normalized by the field-level average yield for that season. The resulting normalized yield represents the relative departure of zone-level yield from the field-level average yield and allows yield to be better compared across growing seasons and across management zones. Results demonstrated that variability in normalized yield across growing seasons and soil types was greater on Nicollet, Okoboji, and Clarion soil types than on Harps and Canisteo soils. Thus, the measured yield on Nicollet, Okoboji, and Clarion soils tended to deviate more from the field-level average yield than that on Harps and Canisteo soils. This suggested a direct relationship between the variation of normalized yield across growing seasons and mean RMSEP for each soil type (fig. 5). The soils types having measured yield values that deviated more greatly from the field-level average were also the soil types on which RMSEP from cross-validation indicated that the calibrated model was having greater difficulty simulating yield in independent growing seasons.

These results make more sense when considering the topologic placement of these soil types on the Iowa landscape. According to the county soil survey (USDA-SCS, 1981), Clarion loam is a well-drained soil found on the convex hilltops and knolls of uplands. Nicollet loam is similar to Clarion; however, it is found on lower knolls and exhibits well-drained to somewhat poor drainage patterns. Okoboji silty clay loam is a very poorly drained soil and is located at the bottom of concave upland depressions. Harps loam soils are poorly drained and exist on the rims of the upland depressions surrounding the Okoboji soils and on the low ridges between depressions. Canisteo silty clay loam soils are poorly drained and lie on the low flats between upland hilltops and depressions. Thus, for the soil types present at our study site, there is a generally increasing topologic trend moving from Okoboji to Harps to Canisteo to Nicollet to Clarion soil types. The southern portion of the study area provides a perfect example of this relationship between soil type and topography for these soils (fig. 2). Existing in the west central portion of the study area is an area of Okoboji silty clay loam at an elevation just less than 300 m. Harps loam surrounds this topologic depression at a slightly higher elevation. Toward the southeast, the elevation increases to more than 305 m, and Clarion loam is found on the hilltop. On the lower knolls between Clarion loam is found Nicollet loam, and Canisteo is found winding between the hilltop and the depression.

To better understand the effects of the soil system on model performance, variability in measured yield was

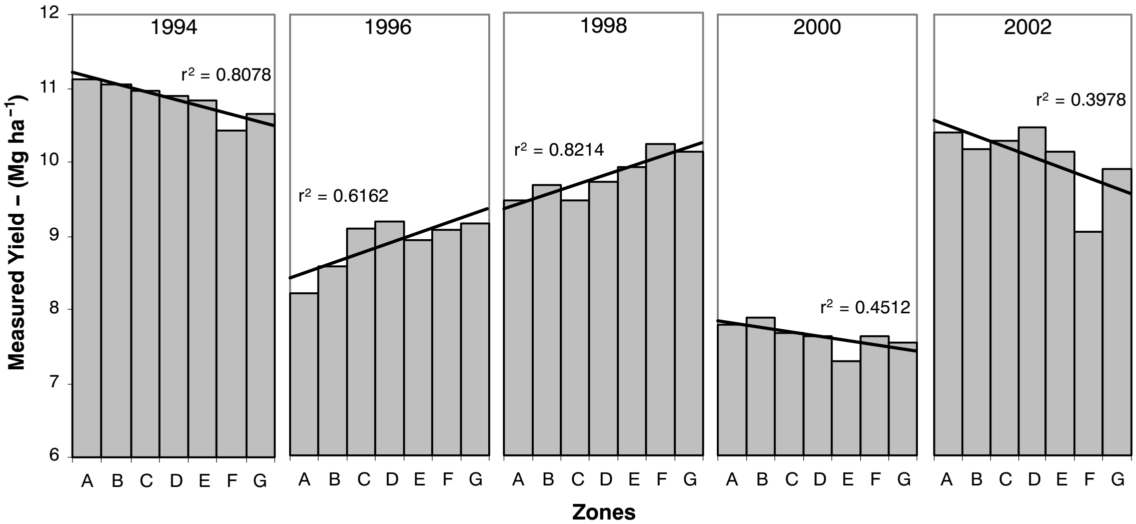

Figure 6. Measured yield exhibited trends along zones of increasing topography from zone A to zone G, and the direction of the trend varied based on the growing season.

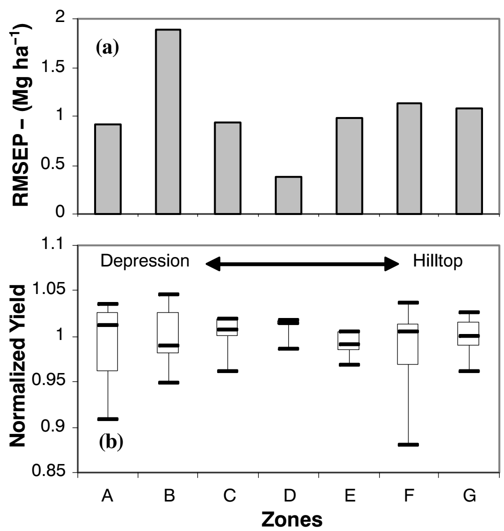

Figure 7. (a) RMSEP from cross-validation on five growing seasons and (b) variability of measured yield, normalized by annual field-level average yield, across growing seasons are both greater in the depression and on the hilltop.

investigated along seven management zones associated with the topographic rise at the southern portion of the study area. The zones of interest are labeled A through G in figure 4b, where zone A is at the lowest elevation in the topological depression and zone G is at the highest elevation on the hilltop. Plotting the measured corn yield for these zones in each growing season demonstrated that yield trends existed along the topography of this region. Furthermore, the direction of the trend varied depending on the growing season (fig. 6). In 1996 and 1998, yield was higher on the hilltop and lower in the depression, with a decreasing yield trend down the slope. In 1994, 2000, and 2002, yield was higher in the depression and lower on the hilltop, with a decreasing yield trend up the slope. When measured yield values for these zones were normalized by the field-level average yield for the season, the results showed that measured yield across growing seasons deviated most significantly from the field-level average yield for zones located at the topological extremes. Zones A and B, located in the depression, and zones F and G, located on the hilltop, all tended to exhibit greater variation in normalized measured yield, while zones lying on the sideslope between the topological extremes tended to exhibit less variation (fig. 7). Finally, similar to the zones based on soil type, the values for RMSEP in topography-based zones exhibit the same pattern as the variation in normalized yield across these zones. RMSEP tends to be greater in the zones at the topological extremes in the depressions and on the hilltops (figs. 4b and 7).

Clearly, these results show that there is a topographic effect on spatial yield variability and that the effect manifests itself in opposite ways depending on the growing season (fig. 6). In some years, yield on the Okoboji and Harps soils in the topological depressions are higher than yield on the Nicollet and Clarion soil types on the hilltops. In other years, the opposite is true. Across growing seasons, the corn yield measurements on the Okoboji and Harps depressions and on the Nicollet and Clarion hilltops tend to deviate more significantly from the field-level average than yield found between the topological extremes (figs. 5 and 7). From this, we can conclude that topography has the greatest impact on yield at the extreme topographic locations. It is not exactly clear how or why topography affects measured yield in this way, but the data clearly indicate that an effect is there. The significance of this result is that the spatial patterns of model performance exhibited a similar trend. RMSEP values from LOO cross-validation using all five years of measured data indicated that the calibrated model simulated yield more poorly on both the Okoboji and Harps soils in the depressions and on the Nicollet and Clarion soils on the hilltops (figs. 4, 5, and 7). The calibrated model was always best able to simulate yield on the sideslopes between the topological extremes, typically having the Canisteo soil type. Since the model does not simulate any processes by which topography might affect yield, such as the lateral redistribution of water down the sideslope or hillshade effects on incoming solar radiation, we conclude that these limitations of the model contribute to the higher simulation error at the locations were topography affects yield the most.

Parameter Behavior

Optimized parameter sets for effective tile drainage rate and K_{SAT} of the bottom layer from zones E and F (fig. 4b) demonstrate how the mean and variance of parameters sets obtained from the model calibration procedures can be used to further evaluate the model performance and the reliability of parameters in each management zone. Drawing 95% confidence ellipses around the parameter means demonstrated that the optimizer tended to provide a more stable solution for zone E (fig. 8a) than for zone F (fig. 8b). For zone E, the generalized variance, computed from the covariance matrix for parameter sets from calibrations using one, two, three, and four growing seasons, was 2.9E-4, 4.9E-5, 2.3E-5, and 9.3E-6, respectively. Although the generalized variance is a unitless quantity, the magnitudes of the parameter set variances demonstrated a decreasing trend as the number of growing seasons used to calibrate the model increased. Therefore, with increased availability of yield measurements, the optimizer tended to converge on parameter values that were more similar. The 95% confidence ellipses around the parameter means for zone E tended to shrink in size as additional growing seasons were used to calibrate the model (fig. 8a). Narrowing the confidence limits means that we can be increasingly certain about the parameter values that should be used in future applications of the model. One exception is the ellipse for parameter sets based on four seasons of measured data, which is slightly larger than the ellipses for parameter sets based on two and three seasons of data. Although the ellipse is larger, the generalized variance for four-year parameters sets was less than that for three-year and two-year parameter sets. Thus, it is expected that the larger confidence ellipse for four-year parameter sets is more dependent on a difference in sample sizes; recombining the measured data for model calibrations results in ten parameter sets for two-year and three-year combinations but only five for four-year combinations. In addition to the size of the confidence ellipses, the relative position of the mean values gives another indication of the optimizer's ability to generate stable solutions for parameters in a management zone. For zone E, the parameter mean values are held relatively tightly together, and each point falls within the area of intersection for all confidence ellipses (fig. 8a). Even the parameter set generated by the model calibration using all five years of measured data falls within the expected confidence limits determined from calibrations using fewer seasons. With decreasing parameter variance and a stable mean value as more growing seasons are added into the calibration, we conclude that the optimizer is relatively stable and reliable in zone E.

Different patterns of behavior are evident in the parameter sets for zone F (fig. 8b). Confidence ellipses tended to shrink in size, as expected. However, the generalized variance was actually larger for parameter sets based on two years of data than that for one year of data. Generalized variance for parameter sets from calibrations based on one, two, three, and four seasons of measured data in zone F was 3.2E-4, 9.9E-4, 2.1E-4, and 3.6E-6, respectively. As the number of growing seasons used to calibrate the model increased, the mean values for the optimum parameters also underwent a substantial shift, especially the effective tile drainage rate parameter. Mean parameter values from calibrations based on more growing seasons also tended to drift beyond the confidence limits for parameter means based on fewer growing seasons. The parameter set from the calibration using all five growing seasons lies within the confidence limits for parameters based on four seasons of data, but the five-year parameter set does not intersect any other confidence regions. Thus, even with five years of measured data, the optimizer was incapable of providing stable parameter values as the number of growing seasons used in the calibration increased. Therefore, we conclude that the optimizer is less stable in zone F, and we have less confidence in deciding which parameter values should be used for future applications of the model. Additional seasons of measured information would be required to determine a range of parameter values around which the model begins to stabilize. Interestingly, zone F is located on the hilltop (fig. 4b), where the model had difficulty simulating yield due to topographic effects on spatial yield variability (fig. 7). On the other hand, zone E, which had the more stable optimizer solutions, lies more on the sideslope between the hilltop and the depression, an area were model performance was better.

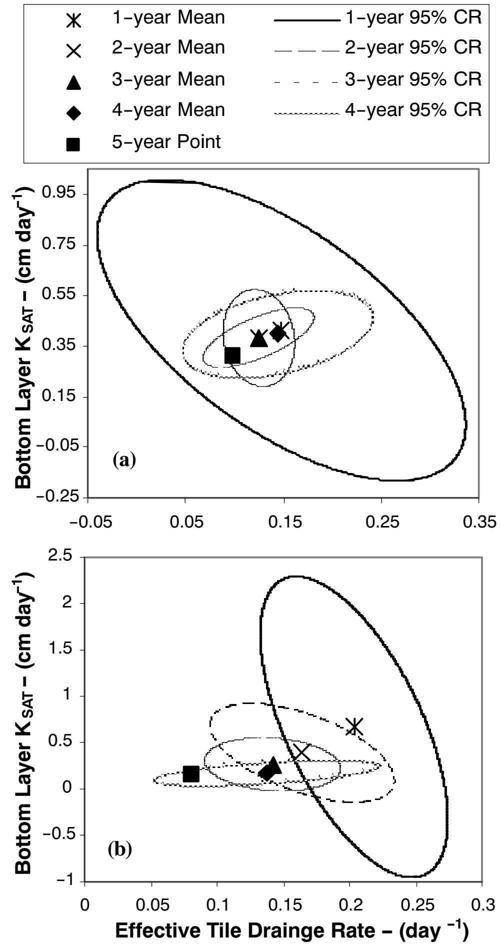

Figure 8. Parameter means and 95% confidence ellipses for optimum parameter sets from model calibrations using combinations of one, two, three, four, and five growing seasons for (a) zone E and (b) zone F.

Conclusions

Performance of crop growth simulations within the precision agriculture DSS used for this study were dependent on the number of growing seasons of measured yield information available for use in the model calibration procedure. The ability of the model to simulate yield in growing seasons independent of the calibration was shown to improve as the number of growing seasons used for model optimization increased. In this study, LOO cross-validation using five growing seasons of measured information generally resulted in a better performing model than that for two growing seasons; however, results suggested that there may be opportunity for further improvement in model performance as the number of available seasons of measured data is increased beyond five. This appears to be one of the most difficult limitations to overcome in using this approach to apply crop growth models in precision agriculture; it takes a long time to collect enough yield measurements to ensure the stability and reliability of model optimization in each management zone. Furthermore, investigations into the spatial nature of model performance suggested that some management zones would require more measured yield information than others in order to arrive at a stable and reliable set of optimized parameters. Characteristics of the soil system, including soil type and topography, were shown to be a probable cause of spatial variability in model performance, and spatial patterns in measured yield were also shown to be related to these characteristics of the soil system. Similarities in the spatial patterns of model performance and measured yield support the idea that the model is limited by its inability to account for any effects of topography on crop yield. Exploration of remote sensing data assimilation techniques for updating vegetative growth state variables would be an appropriate way to address some of these limitations in using crop growth models for applications in precision agriculture. Remote sensing could also be used to obtain better estimates of spatial variability in soil properties and/or plant population across management zones and to perhaps identify the location of tile drains. Use of remote sensing in these ways could result in better LOO cross-validation model performance evaluations when only a few measured growing seasons were available for model calibration.

REFERENCES

Allred, B. J., N. R. Fausey, L. Peters, Jr., C. Chen, J. J. Daniels, and H. Youn. 2004. Detection of buried agricultural drainage pipe with geophysical methods. *Applied Eng. in Agric*. 20(3): 307-318.

Bakhsh, A., R. S. Kanwar, D. B. Jaynes, T. S. Colvin, and L. R. Ahuja. 2001. Simulating effects of variable nitrogen application rates on corn yields and NO_3-N losses in subsurface drain water. *Trans. ASAE* 44(2): 269-276.

Batchelor, W. D., B. Basso, and J. O. Paz. 2002. Examples of strategies to analyze spatial and temporal yield variability using crop models. *European J. Agron*. 18(2): 141-158.

Batchelor, W. D., J. O. Paz, and K. R. Thorp. 2004. Development and evaluation of a decision support system for precision agriculture. In *Proc. 7th Intl. Conference on Precision Agriculture*. Madison, Wisc.: ASA-CSSA-SSSA.

Bennie, J., M. O. Hill, R. Baxter, and B. Huntley. 2006. Influence of slope and aspect on long-term vegetation change in British chalk grasslands. *J. Ecology* 94(2): 355-368.

Boote, K. J., J. W. Jones, and N. B. Pickering. 1996. Potential uses and limitations of crop models. *Agronomy J*. 88(5): 704-716.

Calmon, M. A., W. D. Batchelor, J. W. Jones, J. T. Ritchie, K. J. Boote, and L. C. Hammond. 1999a. Simulating soybean root growth and soil water extraction using a functional crop model. *Trans. ASAE* 42(6): 1867-1877.

Calmon, M. A., J. W. Jones, D. Shinde, and J. E. Specht. 1999b. Estimating parameters for soil water balance models using adaptive simulated annealing. *Applied Eng. in Agric*. 15(6): 703-713.

Carberry, P. S., R. C. Muchow, and R. L. McCown. 1989. Testing the CERES-Maize simulation model in a semi-arid tropical environment. *Field Crops Res*. 20(4): 297-315.

Corana, A., M. Marchesi, C. Martini, and S. Ridella. 1987. Minimizing multimodal functions of continuous variables with the "simulated annealing" algorithm. *ACM Trans. Mathematical Software* 13(3): 262-280.

DeJonge, K. C., A. L. Kaleita, and K. R. Thorp. 2007a. Simulating the effects of spatially variable irrigation on corn yields, costs, and revenue in Iowa. *Agric. Water Mgmt*. (accepted).

DeJonge, K. C., A. L. Kaleita, K. R. Thorp, W. D. Batchelor, and J. O. Paz. 2007b. Methodology for the use of DSSAT models for precision agriculture decision support. *Computers and Electronics in Agric.* (in review).

Efron, B., and G. Gong. 1983. A leisurely look at the bootstrap, the jackknife, and cross-validation. *The American Statistician* 37(1): 36-48.

Efron, B., and R. J. Tibshirani. 1998. *An Introduction to the Bootstrap.* Boca Raton, Fla.: CRC Press.

Fraisse, C. W., K. A. Sudduth, and N. R. Kitchen. 2001a. Calibration of the CERES-Maize model for simulating site-specific crop development and yield on claypan soils. *Applied Eng. in Agric*. 17(4): 547-556.

Fraisse, C. W., K. A. Sudduth, and N. R. Kitchen. 2001b. Delineation of site-specific management zones by unsupervised classification of topography attributes and soil electrical conductivity. *Trans. ASAE* 44(1): 155-166.

Franzen, D. W., D. H. Hopkins, M. D. Sweeney, M. K. Ulmer, and A. D. Halvorson. 2002. Evaluation of soil survey scale for zone development of site-specific nitrogen management. *Agronomy J*. 94(2): 381-389.

Garrison, M. V., W. D. Batchelor, R. S. Kanwar, and J. T. Ritchie. 1999. Evaluation of the CERES-Maize water and nitrogen balances under tile-drained conditions. *Agric. Systems* 62(3): 189-200.

Gauch Jr., H. G., J. T. G. Hwang, and G. W. Fick. 2003. Model evaluation by comparison of model-based predictions and measured values. *Agron. J*. 95(6): 1442-1446.

Goffe, W. L., G. D. Ferrier, and J. Rodgers. 1994. Global optimization of statistical functions with simulated annealing. *J. Econ*. 60(1): 65-99.

Hodges, T., D. Botner, C. Sakamoto, and J. Hays Haug. 1987. Using the CERES-Maize model to estimate production for the U.S. cornbelt. *Agric. Forest Meteorology* 40(4): 293-303.

Hornung, A., R. Khosla, R. Reich, D. Inman, and D. G. Westfall. 2006. Comparison of site-specific management zones: Soil-color-based and yield-based. *Agron. J*. 98(2): 407-415.

Irmak, A., J. W. Jones, T. Mavromatis, S. M. Welch, K. J. Boote, and G. G. Wilkerson. 2000. Evaluating methods for simulating cultivar responses using cross validation. *Agron. J*. 92(6): 1140-1149.

Irmak, A., J. W. Jones, W. D. Batchelor, and J. O. Paz. 2001. Estimating spatially variable soil properties for application of crop models in precision farming. *Trans. ASAE* 44(5): 1343-1353.

Jagtap, S. S., M. Mornu, and B. T. Kang. 1993. Simulation of growth, development, and yield of maize in the transition zone of Nigeria. *Agric. Systems* 41(2): 215-229.

Jaynes, D. B., and T. S. Colvin. 1997. Spatiotemporal variability of corn and soybean yield. *Agron. J*. 89(1): 30-37.

Johnson, R. A., and D. W. Wichern. 2002. *Applied Multivariate Statistical Analysis.* 5th ed. Patparganj, Delhi, India: Pearson Education.

Jones, C. A., and J. R. Kiniry. 1986. *CERES-Maize: A Simulation Model of Maize Growth and Development*. College Station, Tex.: Texas A&M University Press.

Jones, P. N., and P. S. Carberry. 1994. A technique to develop and validate simulation models. *Agric. Systems* 46(4): 427-442.

Jones, J. W., G. Hoogenboom, C. H. Porter, K. J. Boote, W. D. Batchelor, L. A. Hunt, P. W. Wilkens, U. Singh, A. J. Gijsman, and J. T. Ritchie. 2003. The DSSAT cropping systems model. *European J. Agron.* 18(3-4): 235-265.

Kobayashi, K., and M. U. Salam. 2000. Comparing simulated and measured values using mean square deviation and its components. *Agron. J*. 92(2): 345-352.

Liu, W. T. H., D. M. Botner, and C. M. Sakamoto. 1989. Application of CERES-Maize model to yield prediction of a Brazilian maize hybrid. *Agric. Forest Meteorology* 45(3): 299-312.

Marques da Silva, J. R., and C. Alexandre. 2005. Spatial variability of irrigated corn yield in relation to field topography and soil chemical characteristics. *Precision Agric*. 6(5): 453-466.

Pang, X. P., S. C. Gupta, J. F. Moncrief, C. J. Rosen, and H. H. Cheng. 1998. Evaluation of nitrate leaching potential in Minnesota glacial outwash soils using the CERES-Maize model. *J. Environ. Qual*. 27(1): 75-85.

Paz, J. O., W. D. Batchelor, T. S. Colvin, S. D. Logsdon, T. C. Kaspar, and D. L. Karlen. 1998. Analysis of water stress effects causing spatial yield variability in soybeans. *Trans. ASAE* 41(5): 1527-1534.

Paz, J. O., W. D. Batchelor, B. A. Babcock, T. S. Colvin, S. D. Logsdon, T. C. Kaspar, and D. L. Karlen. 1999. Model-based technique to determine variable rate nitrogen for corn. *Agric. Systems.* 61(1): 69-75.

Paz, J. O., W. D. Batchelor, and G. L. Tylka. 2001a. Method to use crop growth models to estimate potential return for variable-rate management in soybeans. *Trans. ASAE* 44(5): 1335-1341.

Paz, J. O., W. D. Batchelor, G. L. Tylka, and R. G. Hartzler. 2001b. A modeling approach to quantify the effects of spatial soybean yield limiting factors. *Trans. ASAE* 44(5): 1329-1334.

Paz, J. O., W. D. Batchelor, and D. G. Bullock. 2002. Procedure to use a crop model to identify water-stressed areas in soybean fields using on-farm data. *Applied Eng. in Agric*. 18(5): 643-646.

Paz, J. O., W. D. Batchelor, and J. W. Jones. 2003. Estimating potential economic return for variable soybean variety management. *Trans. ASAE* 46(4): 1225-1233.

Paz, J. O., W. D. Batchelor, and P. Pedersen. 2004. WebGro: A web-based soybean management decision support system. *Agron. J*. 96(6): 1771-1779.

Ratliff, L. F., J. T. Ritchie, and D. K. Cassel. 1983. Field-measured limits of soil water availability as related to laboratory-measured properties. *SSSA J*. 47(4): 770-775.

Thorp, K. R., W. D. Batchelor, and J. O. Paz. 2005. Programming ArcGIS to generate crop model input files for spatial simulations. ASABE Paper No. 053012. St. Joseph, Mich.: ASABE.

Thorp, K. R., W. D. Batchelor, J. O. Paz, B. L. Steward, and P. C. Caragea. 2006. Methodology to link production and environmental risks of precision nitrogen management strategies in corn. *Agric. Systems* 89(2-3): 272-298.

Urban, M. A. 2005. An uninhabited waste: Transforming the Grand Prairie in nineteenth century Illinois, USA. *J. Historical Geography* 31(4): 647-665.

USDA-SCS. 1981. Soil Survey of Boone County, Iowa. Ames, Iowa: USDA Soil Conservation Service, Iowa Agriculture and Home Economics Experiment Station Cooperative Extension Service, Iowa State University, and State of Iowa Department of Soil Conservation.

Zhao, S. L., S. C. Gupta, D. R. Huggins, and J. F. Moncrief. 2000. Predicting subsurface drainage, corn yield, and nitrate nitrogen losses with DRAINMOD-N. *J. Environ. Qual.* 29(3): 817-825.