Modeling for precision agriculture: how good is good enough, and how can we tell?

E.J. Sadler[1], J.W. Jones[2] and K.A. Sudduth[1]
[1]*USDA-ARS, Columbia, MO 65211, USA*
[2]*University of Florida, Gainesville, FL 32611, USA*
sadlerj@missouri.edu

Abstract

During the development of precision agriculture technology, prior existence of crop simulation models prompted their application to modeling spatial variation in yield. On the face of it, extending a fairly mature 1-D model of crop growth and yield appeared to be a matter of developing spatial suites of input parameters and running a model for each set. For many models, extensive literature had already reported independent tests in multiple combinations of variety, soils and climate which was generally considered substantial validation of the performance of the models. However, most prior literature tests had as objectives, the evaluation of model performance in simulating mean yields across multiple plots in yield trials, which represented the majority of yield data before yield monitors. Precision agriculture requires not just simulation of the mean, but also a simulation of spatial variation. No real consensus has emerged regarding exactly how to test model performance or of what performance constitutes success. In some cases, success simulating inter-annual variability has been asserted as proof of simulating spatial variability. Further, common measures of goodness of fit suffer from dependence on the range of variation in the independent variable. When multiple sources of variation, for example inter-annual and spatial, are combined in a test, commonly used performance measures may fail to support the hypotheses represented in a paper's objectives. We outline several issues relevant to the topic, specifically (1) fundamental differences between simulating means and simulating variation and how these results can be evaluated, (2) the need to link performance measures to stated objectives, (3) an example of performance, isolating sources of variation and model performance toward simulating each source, and (4) a discussion of potentially preferable performance measures. By synthesis and example, we provide guidelines and structure for future precision agriculture modeling efforts.

Keywords: model evaluation, validation, spatial variation, temporal variation

Introduction

The information-intensive nature of precision agriculture practically invites the use of process-level computer simulation models of crop growth and yield. Used retrospectively, they can help build understanding, and used prospectively, they can predict effects of precision agriculture recommendations. However, for reasons described below, precision agriculture poses several challenges to both models and modelers. These challenges are not generally recognized, requiring some examination of exactly what we need models to achieve. Further, little consensus has developed about our expectations for models, requiring some discussion of what constitutes success. Finally, the special circumstances of precision agriculture must be considered during model evaluation. Therefore, we examine these three issues, discussing philosophy, theory, and practical issues surrounding modeling for precision agriculture, and illustrate these issues with representative datasets from our experience.

What is 'modeling for precision agriculture'?

Ironically, although the answer to this question defines almost everything discussed in this paper, the question itself has gone largely unasked. A recent book on agricultural systems models (Ahuja *et al.*, 2002b) includes examples spanning several different approaches. Kiniry *et al.* (2002) discuss models applicable to three spatial scales: individual plants, whole fields, and whole drainage basins. Models corresponding to the first of those scales have been parameterized with site-specific inputs, and multiple representations have been run independently to simulate site-specific crop growth and yield. In one such example, Kersebaum *et al.* (2002) applied a one-dimensional model at multiple points within two fields. They further illustrated the combination of that type of model results with autoregressive state-space analysis. Sadler *et al.* (2002) discussed conventional approaches with one-dimensional models, plus driving the models with remote sensing inputs, and performing objective parameterization in an inverse modeling approach. Ahuja *et al.* (2002a) briefly discussed topographic analysis, scaling, and modeling to assess both temporal and spatial variability across landscapes. In separate work, McBratney *et al.* (1997) described several geostatistics approaches to quantifying variability of measured yield in spatial and temporal dimensions, and in their combination. There are many unanswered questions about modeling for precision agriculture and a number of useful approaches. In this paper, we restrict our discussion to the spatial application of models to obtain estimates of crop yield that vary spatially and temporally. To do this, we must examine assumptions made in models that are applied to precision agriculture.

It would appear that researchers have operated under the modeling paradigms developed concurrent with model development, either concluding that no change is needed or not recognizing differences that, although subtle, may have profound implications on use of models and interpretation of their results. Ultimately, modeling requirements are defined by the objectives of modeling experiments, and under the needs of precision agriculture, these objectives are fundamentally different than for much of prior modeling research. In short, most prior modeling objectives required simulating yield for a homogenous area. When replications across a field were used for parameterization, models functioned at the field or larger scale. On the other hand, precision agriculture requires simulation of point yields at many places within a field. Thus, not only is there a requirement to accurately simulate the field mean, there is also a requirement to accurately simulate the variation in the field (see Sadler and Russell, 1997 for a broad description of this topic). This is a fundamentally different requirement than simulating the mean, but this distinction appears to not be recognized in many research articles. Maximum performance requires point-wise accuracy, meaning the highs and lows must be accurately matched.

Perversely, while we add this need to simulate spatial variation, we simultaneously remove three of the four types of variation in model inputs. Models generally simulate crop growth and yield in response to weather, crop, management, and soil inputs. In the precision agriculture context, there is usually only one weather station and one cultivar, and often only one management (assuming uniform culture, which is the case in many published model tests). Thus, the source of variation in model outputs is by definition restricted to variation in the soil inputs. Unfortunately, spatial soil inputs are particularly difficult to obtain, which has prompted a number of inverse modeling studies to determine best-fit parameters (see Ferreyra *et al.* 2006). Finally, processes involving physical, chemical, and biological variation in real soils are not always fully represented in models of crop growth and yield. These latter two issues are discussed both broadly and with examples by Sadler *et al.* (2002).

Bearing these issues in mind, we propose to identify common types of precision agriculture research objectives, and establish the type of modeling objectives that are needed to meet the research objectives. Sadler *et al.* (1998, 2000) discussed modeling for precision agriculture as needing to be capable of simulating the effect of soil parameters known to cause variation in the subject

context, and of candidates for variable-rate management. These requirements must be addressed with model structure.

Given that precision agriculture involves explicitly managing within-field variation, it would appear that almost all relevant research objectives involve estimating spatial crop growth and yield. For some objectives, the relevant area for which yield is simulated could be a management zone or a soil map unit of reasonably homogenous soil characteristics. Simulating yield for these conditions is a natural extension of prior modeling research, and the goal may be considered to be the map unit or management zone mean. For many other objectives, however, the requirement would appear to be the simulation of yield at all points in the field. Examples of such studies include spatial recommendations for on-the-go management, or feasibility studies to examine whether there would be economic benefit to precision agricultural management. For any case, if the interpretation depends on zone or point-wise accuracy in simulating yield, then the conclusions of the paper are only as good as the accuracy of the model.

How good is good enough?

General accuracy issues regarding modeling for precision agriculture were discussed by Sadler *et al.* (2000), who pointed out that accuracy requirements are as varied as model research objectives. Thus, there can be no definitive statement of required accuracy. Ideally, the model result would exactly match the corresponding measurement at all points in the field. However, sub-ideal results can provide sufficient information to meet some research objectives. For instance, qualitative accuracy, in which the direction of the effect of some management change is simulated correctly, can indicate what management might be recommended in some cases. If the simulated high and low yields properly indicate the areas of the field where the high and low yields occur, management zones could be delineated from the information. Target yields for zones or map units may require only the accuracy of the mean.

However, any research objective that depends on the extremes or range of yields expected would suffer if these were not quantitatively accurate. Any objective depending on the sensitivity of the model, such as optimizing variable-rate management, would need to have accuracy of both the mean and of the derivative with respect to the managed input (Sadler *et al.*, 2000). Risk analysis probably puts even more emphasis on the model's ability to simulate well the tails of the distributions. These latter imply the variation is also simulated accurately. The need for unbiased yield estimates, or accuracy of simulating the mean, is generally recognized. The need for accurate simulation of the variation is not.

How can we tell?

Bearing these considerations in mind, how can models be tested and their performance be confirmed? Model tests from pre-precision agriculture literature typically included regression or correlation of simulated against observed values (or observed against simulated – more on that later), calculations of root mean square error, mean error (or bias), and in some cases, model efficiency as defined by Nash and Sutcliffe (1970). Most model tests in precision agriculture have used regression as the primary test. Further, there has been essentially no discussion of measurement error in published tests. This issue must eventually be considered, but it is beyond the scope of this work.

Simple linear regression

Simple linear regression of simulated yield as a function of observed yield is pre-programmed in most application software and therefore is probably most widely used. The interpretation of the coefficient of determination (r^2) as the fraction of variation in the measurements being explained by the model is intuitive as a performance measure. There is some difference of opinion whether to

regress simulated against observed or the reverse, but r^2 from both is numerically equivalent, and perfect agreement converges to the same coefficients, with intercept of zero and slope of one. Researchers using the regression approach generally conclude that a model produces useful results if the simulated output represents ~70-80% or more of the variation in the observed result. Although there has been less discussion of slope and intercept, it is not recommended to rely solely, or even primarily, on r^2 without due consideration being given to slope and intercept (Krause *et al.*, 2005).

Root mean square error

Many researchers have reported the root mean square error (RMSE) as a performance measure. It has useful characteristics in that it approaches zero with perfect performance and penalizes large error with the commonly used square function.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(O_i - S_i)^2}{n}} \quad (1)$$

Where O = observed value, S= simulated value (formally, 'predicted' is not rigorous because it does not exist concurrently with observed values), and n is the number of values.

Model efficiency

The hydrologic disciplines often employ a model efficiency developed for river forecasting by Nash and Sutcliffe (1970).

$$E_{NS} = 1 - \frac{\sum_{i=1}^{n}(O_i - S_i)^2}{\sum_{i=1}^{n}(O_i - \bar{O})^2} \quad (2)$$

Where $\bar{O}$ represents the mean observed value.

The E_{NS} statistic approaches 1 for perfect model performance, and a value of 0 indicates that the mean value is as good a predictor as the model (Krause *et al.* 2005). In the hydrologic interpretation, E_{NS} of 0.5 or more is generally considered sufficient to begin interpreting the model results as representative. However, that threshold is more than likely specific to the hydrologic discipline and would need to be determined for other modeling disciplines through experience.

Additional challenges with multiple-year data

Any of the above techniques should be safe to apply to data from tests in which single sources of variation (i.e. temporal or spatial in precision agriculture) exist. However, depending on the relative contribution of the two sources, simple application of any of the above techniques may cause misinterpretation of the statistical results. In many cases, inter-annual variation of the mean field yield greatly exceeds within-field variation of point values during the year. Under these conditions, it can be shown that models capable of simulating field means but demonstrably incapable of simulating spatial variation can still produce values of $r^2$and E_{NS} high enough to suggest performance adequate for general use.

As the foundation for much of the following depends upon the reader agreeing with the thesis that temporal and spatial variation must be considered separately, we offer two examples as proof. There are two cases in which a model explains none of the spatial variation in yield: one in which the model returns a constant, and one in which the model returns a number that is random relative to the observed value. These cases are easily constructed and demonstrated. We started with observations from a 7-yr record of soil-map-unit-mean yield from Florence, SC, USA (Sadler *et al.*, 2000). We then created two datasets for which performance of spatial yield simulation was zero, but temporal

yield simulation was perfect or nearly so. For the constant case, we set the 'simulated' yield to equal the observed annual mean (Figure 1). For the random case, we generated a pseudo-random number using the random normal function in SAS (SAS Institute, 2006) with mean equal to the observed annual mean and coefficient of variation (CV) of 10% (Figure 2). Thus, in both cases, the temporal

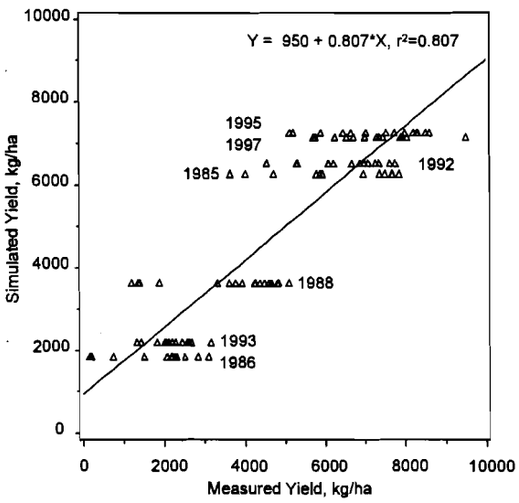

Figure 1. Synthesized data to illustrate zero spatial performance with perfect temporal performance. Simulated output was the annual mean of the measured values (measured data from Sadler *et al.*, 2000).

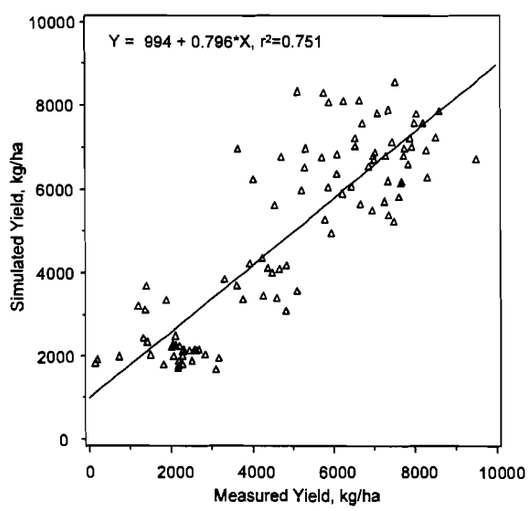

Figure 2. Synthesized data to illustrate zero spatial performance with perfect temporal performance. Simulated output was random values about the annual mean of the measured values with coefficient of variation (CV) of 10% (Sadler *et al.*, 2000).

variation was simulated very well by design, but there was no spatial simulation accuracy at all. As seen in the two figures, the r^2 values obtained were 0.81 and 0.75. However, by definition, these two cases have no value at all except in estimating the mean yield for the year. It is very difficult to argue that either case would contribute information useful to precision agriculture.

While it is immediately apparent that neither of the test cases just discussed were capable of simulating spatial yield, such is not usually the case with real data. In some cases, the model appears capable of simulating both temporal and spatial yield variation relatively well. For this situation, a method is needed to objectively analyze the data. We propose a method to separate the temporal and spatial components of variation, somewhat analogous to Kobayashi and Salam (2000), who separated mean squared deviation into its components. Our method uses linear regression to test temporal performance by comparing annual means of simulated and observed values, and then uses linear regression to test spatial performance by comparing the residuals from those means for the entire dataset. The residuals are computed using the following equations for each data value.

$$R_{Oy} = O - \overline{O}_Y \tag{3}$$

$$R_{Sy} = S - \overline{S}_Y \tag{4}$$

where the subscripts Y indicate the annual mean for observed and simulated values.

This procedure is illustrated using soybean yield data from Wang *et al.* (2003). For the purposes of this illustration only, their calibration and validation datasets were combined (and one apparent outlier that they identified was deleted), providing 13 sites in 3 years overall. The simple linear regression of S on O (Figure 3) indicates remarkably good fit to the 1:1 line, with $S=-228+1.09*O$, $r^2=0.98$, $E_{NS}=0.96$, $RMSE=116$, and $bias=3.69$. These measures all compare quite favorably with the best results these authors have seen. However, as shown in Figure 4, regression of the annual means indicates $S=-449+1.19*O$, $r^2=1.00$, suggesting that there is a slight underestimation of low yields in one year. When the regression was performed on the residuals from the means (Figure 5), the relationship was $R_{Sy}=0.898*O$, with $r^2=0.96$ (the intercept is zero by definition, but the regression was not constrained). This result, unanticipated from prior analyses of the combined data, illustrates additional interpretation that may be possible once the temporal and spatial performances

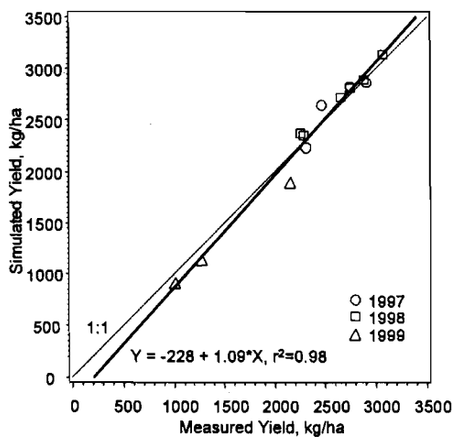

Figure 3. Simulated and observed soybean yields for three years from Wang *et al.* (2003). Their data point B3, which they identified as an outlier, was deleted. The data shown are the calibration and validation data combined.

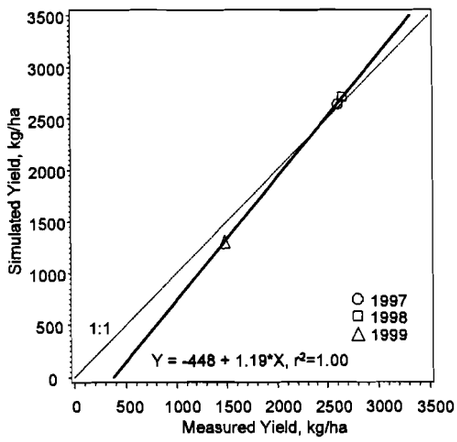

Figure 4. Simulated and observed annual mean soybean yields for three years from Wang *et al.* (2003).

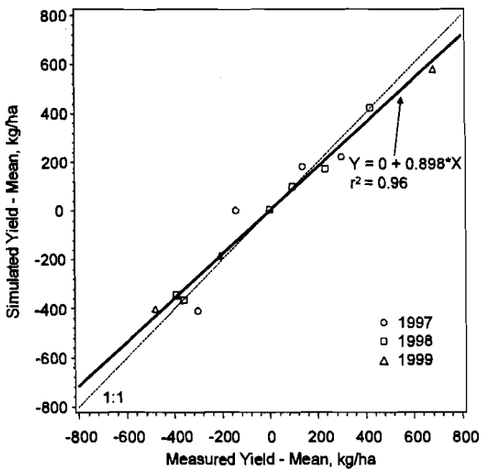

Figure 5. Simulated and observed residuals from annual mean soybean yields for three years from Wang *et al.* (2003).

are considered separately. Here, the slope less than unity seen in Figure 5 indicates a slight but systematic underestimation of the measured variation in yield. This was not apparent from the commonly used regression shown in Figure 3.

Conclusions

Tests of models should be chosen to match research objectives, in particular considering multiple sources of variation in the test data set. In precision agriculture, one would expect the primary goal to be the ability of a model to simulate spatial variation. A test combining year-to-year and

spatial variation demonstrated that good year-to-year performance masks spatial non-performance for typical precision agriculture data. A method to separate spatial from temporal variation and to test the separate performance was provided, and by example, the added value of separating the sources was shown.

References

Ahuja, L.R., Green, T.R., Erskine, R.H., Ma, L., Ascough II, J.C., Dunn, G.H., Shaffer, M.J. and Martinez, A. 2002a. Addressing spatial variability in crop model applications. In: Agricultural System Models in Field Research and Technology Transfer, eds. L.R. Ahuja, L. Ma, and T.A. Howell. Lewis Publishers, Inc., Boca Raton, FL, USA. Chapter 13, pp. 265-272.

Ahuja, L.R., Ma, L. and Howell, T.A. 2002b. Agricultural System Models in Field Research and Technology Transfer. Lewis Publishers, Inc., Boca Raton, FL, USA.

Ferreyra, R.A., Jones, J.W. and Graham, W.D. 2006. Parameterizing spatial crop models with inverse modeling: sources of error and unexpected results. Transactions of the ASABE 49 (5) 1547-1561.

Kersebaum, K.C., Lorenz, K., Reuter, H.L. and Wendroth, O. 2002. Modeling crop growth and nitrogen dynamics for advisory purposes regarding spatial variability. In: Agricultural System Models in Field Research and Technology Transfer, edited by L.R. Ahuja, L. Ma, and T.A. Howell. Lewis Publishers, Inc., Boca Raton, FL, USA. Chapter 11, pp. 229-252.

Kiniry, J.R., Arnold, J.G. and Yun, X. 2002. Applications of models with different spatial scales. In: Agricultural System Models in Field Research and Technology Transfer, edited by L.R. Ahuja, L. Ma, and T.A. Howell. Lewis Publishers, Inc., Boca Raton, FL, USA. Chapter 10, pp. 207-227.

Kobayashi, K, and Salam, M.U. 2000. Comparing simulated and measured values using mean squared deviation and its components. Agronomy Journal 92 345-352.

Krause, P., Boyle, D. P. and Bäse, F. 2005. Comparison of different efficiency criteria for hydrological model assessment. Advances in Geosciences 5 89-97.

McBratney, A.B., Whelan, B.M. and Shatar, T.M. 1997. Variability and uncertainty in spatial, temporal and spatiotemporal crop-yield and related data. In: Precision Agriculture: Spatial and Temporal Variability of Environmental Quality. Wiley, Chichester, UK (Ciba Foundation Symposium 210). pp.141-160.

Nash, J.E. and Sutcliffe, J.V. 1970. River flow forecasting through conceptual models part I — A discussion of principles. Journal of Hydrology 10 (3) 282–290.

Sadler, E.J. and Russell, G. 1997. Chapter 4. Modeling crop yield for site-specific management. In: The State of Site-Specific Management for Agriculture, eds. F.J. Pierce, and E.J. Sadler. ASA, Madison, WI, USA pp. 69-79.

Sadler, E.J., Busscher, W.J., Stone, K.C., Bauer, P.J., Evans, D.E. and Millen, J.A. 1998. Site-specific modeling of corn yield in the SE Coastal Plain. In: Proceedings of the 1st International Conference on Geospatial Information in Agriculture & Forestry, Lake Buena Vista, FL,USA June 1-3. pp. I-214-221. ERIM International, Ann Arbor, MI, USA.

Sadler, E.J., Gerwig, B.K., Evans, D.E., Busscher, W.J. and Bauer, P.J. 2000. Site-specific modeling of corn yield in the SE Coastal Plain. Agricultural Systems 64 189-207.

Sadler, E.J., Barnes, E.M., Batchelor, W.D., Paz, J. and Irmak, A. 2002. Addressing spatial variability in crop model applications. In: Agricultural System Models in Field Research and Technology Transfer, eds. L. R. Ahuja, L. Ma, and T.A. Howell. Lewis Publishers, Inc., Boca Raton, FL, USA. Chapter 12, pp. 253-264.

SAS Institute. 2006. SAS User's Guide, V9.0. SAS Institute, Cary, NC, USA.

Wang, F., Fraisse, C.W., Kitchen, N.R. and Sudduth, K.A. 2003. Site-specific evaluation of the CROPGRO-soybean model on Missouri claypan soils. Agricultural Systems 76 (3) 985-1005.