Combining Farm and University/Industry Information
for Variable Rate Fertilizer Decisions

Fifth Annual Kansas Precision Agriculture Conference
Great Bend, Kansas, January 29-30, 2002

Terry L. Kastens, agricultural economist, Kansas State University
Kevin C. Dhuyvetter, agricultural economist, Kansas State University
John P. Schmidt, agronomist, Kansas State University

What do site-specific fertilizer managers need?

Farm managers and crop input providers desiring to variably apply fertilizers explicitly or implicitly recognize a number of important issues. First, they recognize that optimal fertilizer rates often depend on more than soil test fertility for the factor in question. For example, optimal fertilizer P might depend on how much fertilizer N is applied, or on soil pH, or on previous crop yields, or on whatever trait is classified in a management zone scheme. Second, optimal fertilizer rates can be quantified – variable rate application (VRA) software requires actual numerical values as inputs. Third, optimal fertilizer rates are continuous. After all, VRA providers routinely interpolate between the step-type fertilizer recommendation rates from soil testing laboratories.

Besides recognizing these important issues, VRA practitioners have wish lists that are expanding every day. For example, recognizing that crop yield is a function of many important factors, they would like to have a mathematical yield function that could be plugged into a spreadsheet so that they could ask "what-if" questions that could lead to more profitable crop production decisions.

Farmers wanting to use VRA and other precision agriculture technologies often feel strongly about two issues. First, they want to use data from their own farms to improve production management decisions. Second, they do not want to wait many years before using their farms' data. That is, they recognize that the increased profitability associated with new technologies is often short-lived because the gains are quickly bid into higher land values and rents.

Soil labs' fertilizer recommendations often fall short

Soil testing laboratories in universities and industry often provide fertilizer recommendation rates and underlying mathematical formulas, but virtually never provide the mathematical yield function underlying their recommendations. Even if underlying yield functions could be provided, those functions probably would include only one or two yield causing factors – because the underlying fertilizer trials were controlled experiments where only one factor was allowed to vary. Thus, for example, fertilizer P recommendations typically arise from yield response to only fertilizer P and soil test P.

A second problem associated with fertilizer recommendations from soil labs is that they usually are extremely short run – with an implicit time horizon of only one year. This could be a problem for P management, where evidence is mounting that higher soil test P (*STP*) leads to higher yields, and that *STP* can be changed over time via excess or deficient applications of fertilizer P (i.e., above or below crop removal P).

Why not just use farm data to determine recommended fertilizer rates?

If soil labs' fertilizer recommendations fall short of VRA practitioners' needs, why not simply use farm data to estimate a yield function that would generate recommended fertilizer rates conditional on yield causal factors of interest? Unfortunately, managers typically are reluctant to apply 0 or low fertilizer rates because they do not want to sacrifice the associated expected profitability. That means the range of farm fertilizer rate data may be insufficient to reliably estimate yield response to those measures. In short, using only farm data may result in a yield model that is the least reliable precisely where reliability is the most important, i.e., regarding recommended fertilizer rates.

Given the issues set forth, site-specific managers typically have three choices: 1) ignore non-fertility information and use university fertilizer recommendations as provided; 2) use a yield model estimated with only farm data, recognizing that some fertilizer recommendations may be unreliable if underlying fertilizer response data are poor; or 3) use farm non-fertility data to adjust university fertilizer recommendations in some ad hoc way.

Might university and farm data be combined?

If farm data on fertilizer response are poor, and if non-fertilizer non-fertility data are absent in university trials, might there be a way to formally combine information from both sources – to capture their comparative advantages? The objective of this research is to develop and evaluate a protocol for combining laboratory-based fertilizer recommendations with farm data to generate an integrated wheat yield model for site-specific fertilizer recommendations. More specifically, yield models are first generated that are consistent with fertilizer recommendations provided by several university and industry sources. Next, fertilizer and fertility response characteristics of these models are totally and partially imposed on a broader yield model estimated from farm data on a number of yield causing factors. An explicit consideration of the potential benefits of managing soil test P over time is a specific outcome of this research.

Yield models

Four key features underlie generally accepted descriptions of yield response functions: 1) diminishing returns to an input, 2) limiting factor response, 3) plateau yields, and 4) fertilizer and fertility are substitutes. One basic approach to yield model construction that accommodates the first three features is represented in the general multiplicative expression:

$$Y_i = \mathrm{A} * Z_i * N_i * P_i \ , \qquad [1]$$

where i indicates a particular location in time or space, A represents the yield plateau, N_i and P_i generally denote the expressions for N fertilizer and soil test N, and P fertilizer and soil test P, respectively, and Z_i denotes one or more other independent variables. Z_i, N_i, and P_i are assumed to be valued between 0 and 1. When everything is "perfect" for crop production (i.e., Z_i, N_i, and P_i each equal 1), the model predicts a yield equal to the plateau. A more explicit representation of the N and P terms, accommodating all four of the key features noted above, could be presented as Eq. 2:

$$Y_i = \mathrm{A} * Z_i * \left(1 - \mathrm{G}_1 * e^{-\mathrm{B}_1 * fertN_i - \mathrm{B}_2 * STN_i}\right) * \left(1 - \mathrm{G}_2 * e^{-\mathrm{B}_3 * fertP_i - \mathrm{B}_4 * STP_i}\right) + E_i \ . \qquad [2]$$

In Eq. 2, Y_i is wheat yield as bu per acre, $fertN_i$ is fertilizer N as lb N per acre , STN_i is soil test N as lb NO_3-N per acre, $fertP_i$ is fertilizer P as lb P_2O_5 per acre, STP_i is soil test P as ppm Bray 1P, *e* denotes the exponential function, *i* is an index for each observation in time or space, and E_i represents an error term so that the equation can hold true everywhere in observed data. Numerical constants (parameters) that must be estimated include A, B_1, B_2, B_3, B_4, G_1, and G_2. The following restrictions must be applied to correctly describe a yield response function: A>0, $0<G_1<=1$, $0<G_2<=1$, and B_1, B_2, B_3, $B_4>0$. Once these parameters are estimated, expected yield, Y_i, can be calculated for specific values of Z_i, $fertN_i$, STN_i, $fertP_i$, and STP_i, because the expected value of the error (E_i) is assumed to be 0.

Using Eq. 2, the expected profit (crop revenue less fertilizer costs) associated with point *i* can be described as in Eq. 3:

$$PROF_i = \mathrm{WPRICE} * Y_i - \mathrm{NPRICE} * fertN_i - \mathrm{PPRICE} * fertP_i \ , \qquad [3]$$

where WPRICE, NPRICE, PPRICE are expected wheat price as $/bu, fertilizer N price as $/lb, and fertilizer P price as $/lb P_2O_5, respectively, $PROF_i$ is expected profit at point *i* in $/acre, and Y_i is expected wheat yield.

If Eq. 2 is substituted for Y_i in Eq. 3, profit-maximizing levels of fertilizer N and fertilizer P (conditional on specific values of Z_i, STN_i, STP_i, and prices) for point *i* can be determined by taking the first derivative of $PROF_i$ with respect to $fertN_i$ and $fertP_i$, setting those derivatives equal to 0, and solving for $fertN_i$ and $fertP_i$ (Eq. 4 and 5):

$$fertN_i^* = \frac{\ln\left[\left(\dfrac{\mathrm{NPRICE} * \mathrm{WPRICE}^{-1}}{\mathrm{A} * Z_i * \left(1 - \mathrm{G}_2 * e^{-\mathrm{B}_3 * fertP_i^* - \mathrm{B}_4 * STP_i}\right) * \mathrm{G}_1 * \mathrm{B}_1}\right)\right] + \mathrm{B}_2 * STN_i}{-\mathrm{B}_1} \ , \qquad [4]$$

and

$$fertP_i^* = \frac{\ln\left[\left(\dfrac{\mathrm{PPRICE} * \mathrm{WPRICE}^{-1}}{\mathrm{A} * Z_i * \left(1 - \mathrm{G}_1 * e^{-\mathrm{B}_1 * fertN_i^* - \mathrm{B}_2 * STN_i}\right) * \mathrm{G}_2 * \mathrm{B}_3}\right)\right] + \mathrm{B}_4 * STP_i}{-\mathrm{B}_3} \ , \qquad [5]$$

where ln denotes the natural logarithm, and $fertN_i^*$ and $fertP_i^*$ denote the model-recommended optimal levels of fertilizer N and P, respectively, for point *i*.

The framework for developing a wheat yield model and associated fertilizer recommendations consistent with agronomic and economic theory is represented by Eq. 1 to 5. Model parameters (A, B's, and G's) could be estimated based on farm data using an algorithm that minimizes in-sample yield prediction error (usually the sum of squared errors, SSE) associated with Eq. 2. However, assuming that farm data on fertilizer may be unreliable, we estimate the parameters in a way that makes fertilizer N and fertilizer P recommendations from Eq. 4 and 5 similar to those provided directly by soil testing laboratories.

Four independent fertilizer recommendations

Because the modeling procedures developed in this research are ultimately tested on data from a Rawlins County, Kansas wheat farm, we considered four soil testing laboratories relevant to that area: Olsen's Agricultural Laboratory, McCook, NE; Kansas State University, Manhattan, KS; University of Nebraska, Lincoln, NE; and Colorado State University, Ft. Collins, CO. Fig. 1 and 2 show the four labs' fertilizer N and P recommendations, respectively. For the figures, yield goal (YG) was set at 10% above the 10-year (1991-2000) average wheat yield of 38 bu/acre reported by Kansas Agricultural Statistics for the northwest Kansas Crop Reporting District.

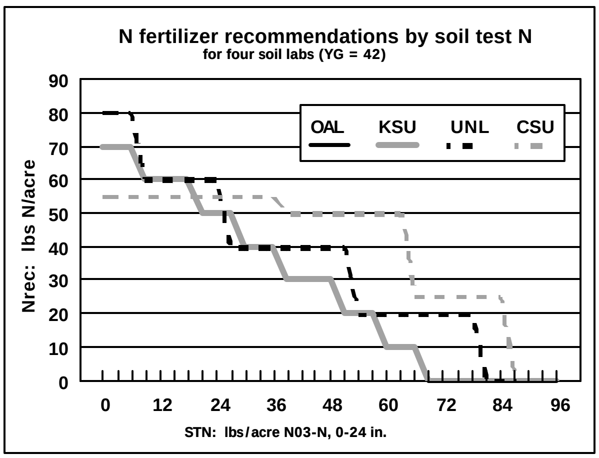

Figure 1

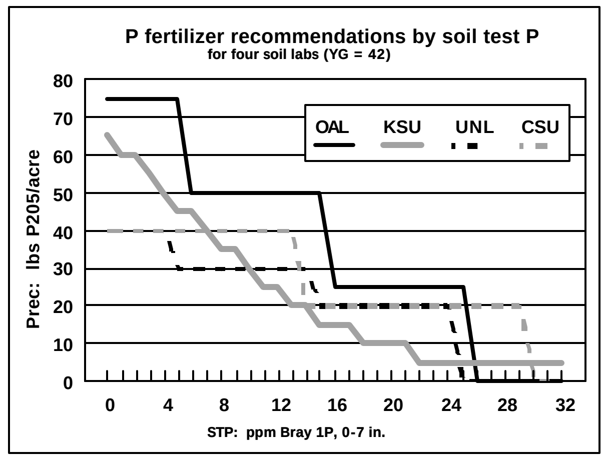

Figure 2

Generating a yield model from fertilizer recommendations

Based on the premise that lab-based fertilizer recommendations are appropriate for fields represented in the three state region of Kansas, Nebraska, and Colorado, simulated data were generated to represent actual field-level data. Yield data (*Y*) were simulated with a mean of 38 bu/acre and a CV (coefficient of variation, which is the standard deviation divided by the mean) of 0.25. *STN* data had a mean of 40 lb/acre and a CV of 0.50; *STP* data had a mean of 16 ppm and a CV of 0.50. Simulated data were used to generate two sets of fertilizer recommendations for each of N and P. First, the soil testing laboratories' formulas underlying Fig. 1 and 2 were used to provide lab fertilizer recommendations. Second, conditional on model parameter values and wheat and fertilizer prices ($3.22/bu for wheat, $0.19/lb for *fertN*, and $0.22 for *fertP*), Eq. 4 and 5 were used to generate model fertilizer recommendations. Parameter values were selected to minimize the SSE between lab and model recommendations. Parameter estimates and corresponding measures of accuracy for the model-generated N and P recommendations are provided in Table 1. Overall, the model-generated N and P recommendations are quite similar to the corresponding lab's fertilizer recommendations, with r^2's ranging between 0.74 and 0.97 for N and ranging between 0.66 and 0.93 for P.

Table 1. Laboratory-based yield model parameter estimates and fertilizer recommendation accuracy measures. Each lab-based yield model was estimated from the same simulated series of 10,000 data points.[a]

	Laboratory whose recommendations underlie the model			
	OAL	KSU	UNL	CSU
parameter value				
G_1	0.2925	0.2897	1.0000	0.2541
B_1	0.0131	0.0140	0.0510	0.0119
B_2	0.0139	0.0143	0.0366	0.0055
G_2	1.0000	1.0000	0.9999	0.9998
B_3	0.0469	0.0688	0.0964	0.0754
B_4	0.1003	0.1559	0.1085	0.0996
accuracy measure				
N: relative RMSE	0.1641	0.1757	0.5167	0.2877
N: r^2	0.9733	0.9710	0.7390	0.9182
P: relative RMSE	0.5292	0.2854	0.5073	0.5887
P: r^2	0.7246	0.9285	0.7443	0.6593

[a] RMSE is the square root of the 10,000-observation average difference between laboratory and lab-based model fertilizer recommendations; relative RMSE is the RMSE divided by the standard deviation of laboratory recommendations; r^2 is the squared linear correlation coefficient between the laboratory and the lab-based model fertilizer recommendation series.

As one example of the simulation exercise, Fig. 3 shows the N recommendations for the OAL-based model and for the underlying laboratory. Similarly, Fig. 4 shows the P recommendations for the KSU-based model and for the underlying laboratory.

It is important to note that the simulation and modeling process used here does not merely generalize fertilizer recommendations of laboratories – as it appears to do in Fig 3 and 4. Rather, a yield model (like that depicted in Eq. 2) is generated in the process, one that can be used to answer useful "what-if" questions to aid fertilizer management decisions. For example, Fig. 5 shows the expected yield associated with average *STN*, optimal *fertN*, and various *fertP* rates, all at four levels of *STP* and using the KSU-based model. Because the functional form of Eq. 2 captures the four fundamental features of yield models (diminishing returns, limiting factor, plateau, and fertilizer/fertility substitution) and also represents KSU's fertilizer recommendations quite well, the responses depicted in Fig. 5 likely represent yield responses inherent to KSU's P recommendations.

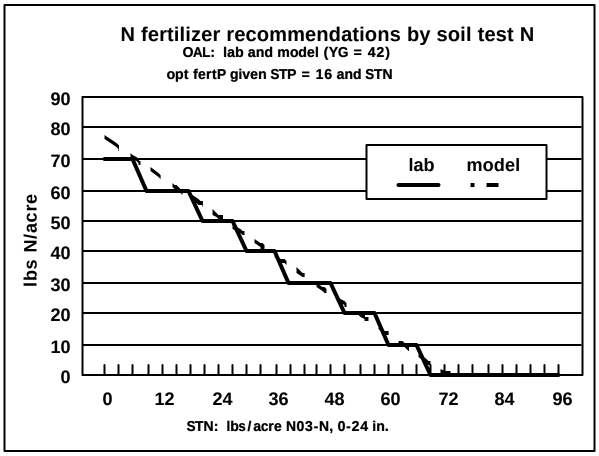

Figure 3

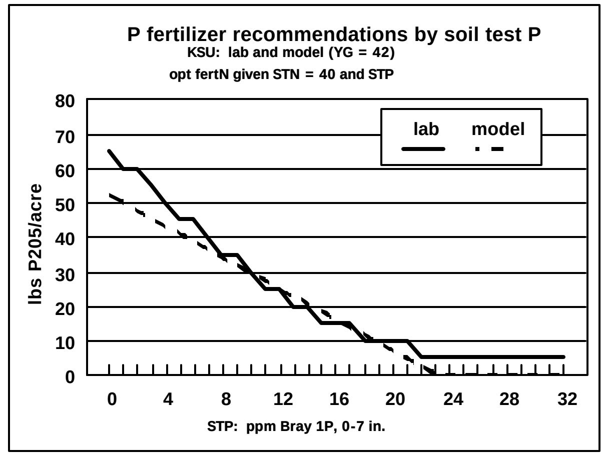

Figure 4

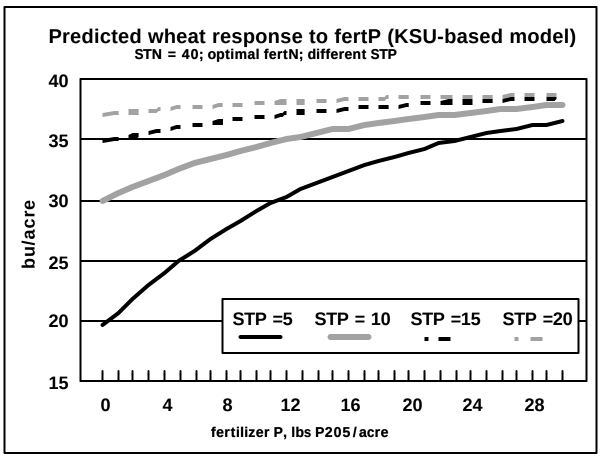

Figure 5

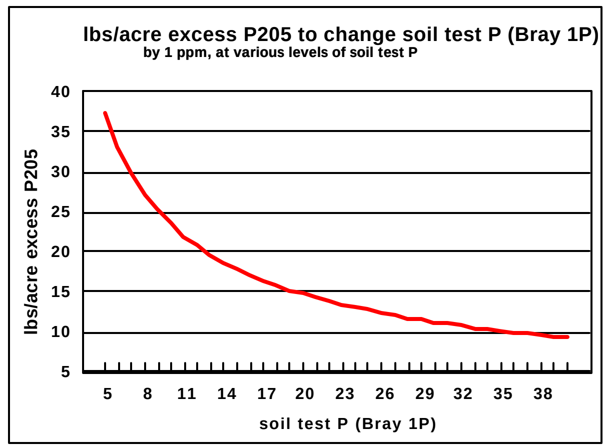

Figure 6

Long run steady-state soil test P level

It happens that, at low levels of *STP*, the four soil labs recommend more P than taken off by crops. Similarly, at high levels of *STP*, the labs recommend less P than taken off by crops. Thus, using a lab's P recommendations over time will result in *STP* approaching some long run steady-state value peculiar to that lab's recommendations. Also, in the long run, each lab's P recommendations ultimately will approach crop-removal P (assumed to be 0.60 lb P_2O_5/bu). How fast *STP* approaches the steady-state level will depend on an assumption of transformation rate, *TR*, where *TR* is the number of pounds of excess (above crop removal, but can be negative) *fertP* required to increase (or decrease when applying less than crop removal) *STP* by 1 ppm. Often a constant *TR* value (e.g., 18) has been used in the academic literature. However, the transformation rate for most soils probably is not constant. Thus, based on University of Missouri data, a transformation rate that depends on *STP* was estimated (shown as Fig. 6).

Using the transformation rate underlying Fig. 6, Fig. 7 depicts optimal *fertP* rates over time for the four lab-based models and Fig. 8 shows resultant *STP* levels over time. That long run steady-state *STP* values are different across the four labs is clearly shown in Fig. 8. The infinite-horizon steady-state *STP* values for the four labs are OAL = 21.7 ppm, KSU = 13.3 ppm, UNL = 16.0 ppm, and CSU = 20.0 ppm. Obviously, the different labs must, at least implicitly, target different *STP* values over time.

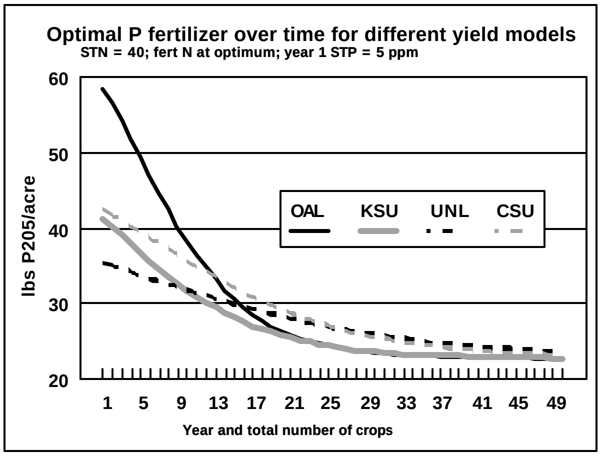

Figure 7

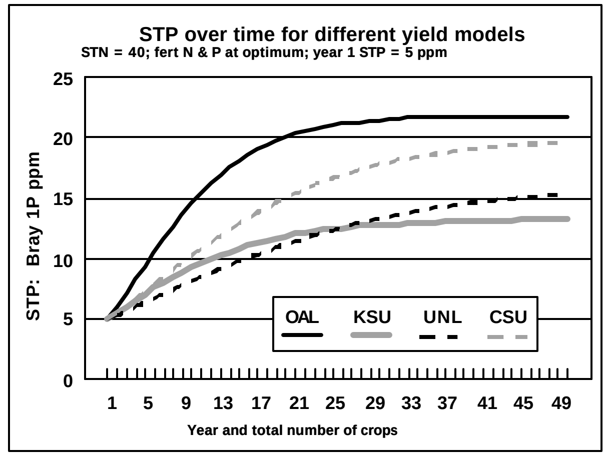

Figure 8

Profit and amortized profit

An important aspect of this simulation exercise, given a model that incorporates and adequately represents reliable N and P recommendations from soil testing laboratories, is the ability to evaluate annual profit (crop revenue less fertilizer expense in $/bu) for each year into the future. Each model, representing the optimal fertilizer recommendation from each laboratory, shows an increase in profitability with time (Fig. 9) as *STP* increases to the steady-state level.

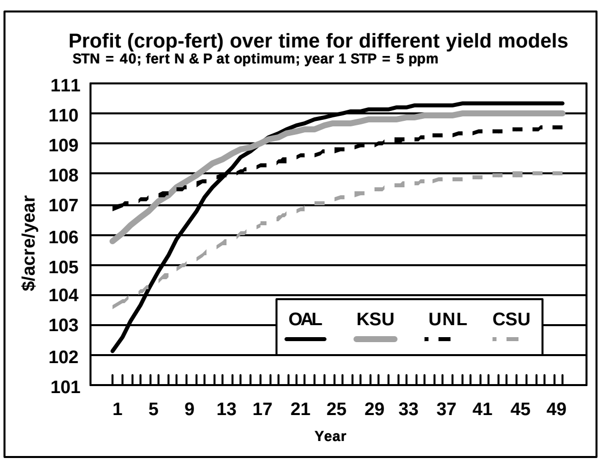

Figure 9

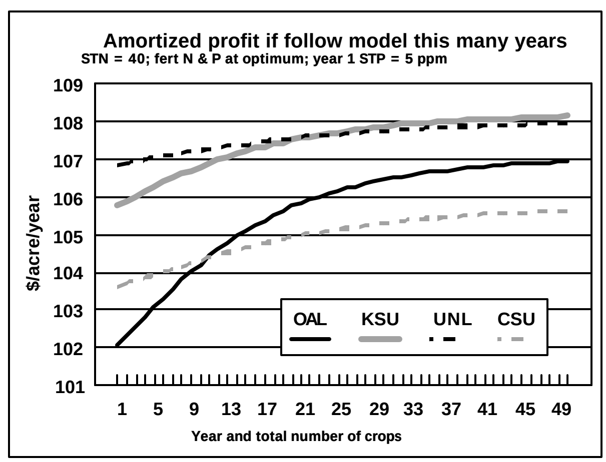

Figure 10

Initially, profits from the OAL-based model were substantially lower than profits from the other lab-based models (Fig. 9). In later years (>10), OAL-based model profits were similar to KSU- and UNL-based model profits, and higher than profits from the CSU-based model. However, over a producer's time horizon, early profits are more valuable than later profits because they can earn interest. To account for this, Fig. 10 shows the annually amortized profits associated with each lab-based model, had the producer followed the model over the number of years shown on the x-axis (amortization factor based on a 9% interest rate and a 35% marginal income tax rate).

It should be noted that the differences in long-run profits among the lab-based models shown in Fig. 10 mainly are due to different assumptions regarding *STP*. For example, the KSU-based model provided the greatest profits because it had assumed that optimal yields can be obtained with the lowest *STP* levels (Fig. 8). Whether optimal yields actually can be obtained at lower levels of *STP* cannot be determined here. After all, though each of the four laboratories' fertilizer recommendations were assumed equally plausible for generating mean wheat yields historically observed in northwest Kansas, it could be that producers following OAL recommendations have higher than average wheat yields and KSU-following producers have lower than average wheat yields. Ultimately, the producer must decide on a recommendation (or some combination of one or all of them) to "believe." From that decision, the producer should consider using the corresponding yield model reported here as a foundation from which to develop a farm-specific, or site-specific, yield model.

Because the fertilizer P recommendations from the four laboratories considered here are *conditional on STP*, they essentially are 1-year horizon recommendations; they do not explicitly consider the potential benefit associated with increasing *STP* with time (or allowing *STP* to decrease when it is sufficiently high for obtaining optimum yield). If producers expect to control land more than 1 year, then these fertilizer P recommendations should be considered minimum levels (conditional on *STP*) and not optimal. In particular, Fig. 5 suggests that higher *STP* leads to higher yields and Fig. 8 suggests that *STP* can be modified over time. Taken together, these figures suggest that *STP* might be treated as a managed crop production factor over time.

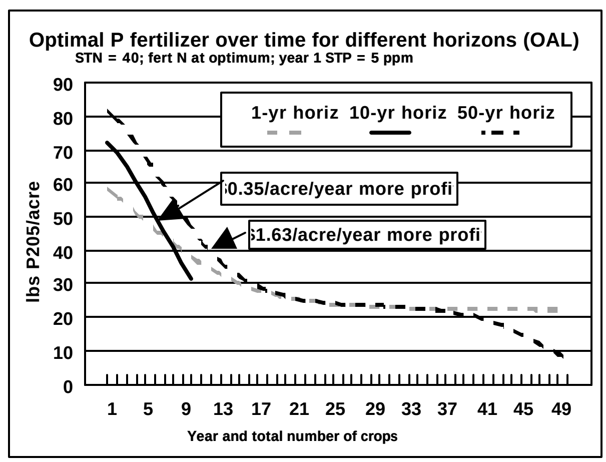

Figure 11

A benefit of the simulation process developed here is that resultant yield models can be used to explicitly answer the question, How much fertilizer P should be applied each year in order to maximize profits over some finite time horizon? As one example, Fig. 11 shows the P recommendation for the OAL-based model over 50 years, assuming the 1-year horizon implicit in the lab's recommendations (same as the OAL line of Fig. 7). It also shows rates for 10 years when *fertP* rates were selected to maximize annually amortized 10-year profits (the 10-year horizon line), and rates for a 50-year horizon. Relative to the 1-year horizon, maximizing profits over a 10-year horizon resulted in fertilizer P rates that were 23% higher in year 1, 16% lower in year 10, and 10% higher overall. Despite applying more fertilizer, annual profits were $0.35/acre higher when a 10-year horizon was explicitly considered rather than simply following the 1-year horizon for 10 years. Explicitly considering time horizon was even more profitable for longer horizons. For example, annual profits for the 50-year horizon rates were $1.63/acre higher than for rates based on a 1-year horizon each year for 50 years.

Validating the laboratory-based yield modeling approach

For the model-generating framework developed in this research to have merit, it should result in fertilizer N and P recommendations that are, in some important way, an improvement over those generated directly from a yield model estimated from only farm data. To that end, the – and P-yield-response information from the lab-based models are integrated into a yield model estimated using data from the example Rawlins County, Kansas wheat farm (124 field-year observations, 1994-2001).

As considered here, the validation process begins by using historical data from the example farm to estimate a multi-factor yield model of the general form shown in Eq. 2, but involving more yield-causing factors. Model parameters are selected that minimize the sum of squared in-sample yield-prediction errors (SSE), subject to certain parametric constraints such as those described immediately following Eq. 2. The yield model estimated was

$$\begin{aligned} Yield = {} & \mathrm{A}\left(1-\mathrm{B}_0\, frost\right)\left(1-\mathrm{G}_1 e^{-\mathrm{B}_1 fertN-\mathrm{B}_2 STN}\right)\left(1-\mathrm{G}_2 e^{-\mathrm{B}_3 fertP-\mathrm{B}_4 STP}\right)\left(1-\mathrm{G}_3 e^{-\mathrm{B}_5 om}\right) \\ & \left(\mathrm{B}_6+\mathrm{B}_7 sand+0.25\mathrm{B}_7^2\left[\mathrm{B}_6-\mathrm{A}\right]^{-1} sand^2\right)\mathrm{A}^{-1} \\ & \left(\mathrm{B}_8+\mathrm{B}_9 clay+0.25\mathrm{B}_9^2\left[\mathrm{B}_8-\mathrm{A}\right]^{-1} clay^2\right)\mathrm{A}^{-1} \\ & \left(\mathrm{B}_{10}+\mathrm{B}_{11} pH+0.25\mathrm{B}_{11}^2\left[\mathrm{B}_{10}-\mathrm{A}\right]^{-1} pH^2\right)\mathrm{A}^{-1}+E\ , \end{aligned} \qquad [6]$$

where *Yield* is wheat yield (bu/acre); *frost* is a binary variable equal to 1 if a substantial late spring frost had occurred, else 0; *fertN* is fertilizer N (lb N/acre); *STN* is soil test N (lb NO_3-N /acre); *fertP* is fertilizer P (lb P_2O_5/acre); *STP* is soil test P (ppm); *om* is soil organic matter content (%); *sand* and *clay* are each soil texture variables (%), *pH* is soil pH; and *E* is an error term. In the model, A is the positive plateau parameter, parameters B_0, G_1, G_2, and G_3 are constrained between 0 and 1, and parameters B_1-B_6, B_8, and B_{10} are constrained to be positive. The Eq. 6 model estimated from only farm data is referred to as the *Farm-Only* model.

Information from the lab-based models is incorporated into the farm information by estimating a second yield model of the form shown in Eq. 6. However, in this model, during the SSE-minimizing process, N- and P-related parameters (the G_1, B_1, B_2, G_2, B_3, and B_4 values) are constrained to exactly equal those of the lab-based models depicted as Eq. 2 (parameter values shown in Table 1). This forces the yield response to *fertN*, *STN*, *fertP*, and *STP* to equal (in a proportion-of-plateau-yield sense) that of the lab-based models – farm data on N and P essentially are ignored in model estimation. These models, one for each laboratory, are referred to as *Lab-Constrained* models.

In the quest to improve yield models based on farm data, it may be inappropriate to totally ignore farm data on N and P during the model estimation process. For example, a farm manager may be gathering ever better farm information on – and P-response over time by conducting on-farm experiments with substantial fertilizer rate variability. Thus, another approach to integrating information from lab-based models into the farm information is considered here. This approach takes the soil testing laboratory's N and P fertilizer recommendations for the farm as what the recommendations are "expected" to be. Then, farm information on N, P, and other factors is used to estimate a yield model conditional on the prior fertilizer recommendation information. The formal process used here is maximizing cross entropy (MCE), and the four resulting models, one for each laboratory, are referred to as *Lab-MCE* models.

Within an estimation sample, increasing the constraints on model parameters will always reduce the yield prediction accuracy relative to less constrained optimization. Here, that means that the *Farm-Only* model will always predict yield more accurately in-sample than either the *Lab-Constrained* or the *Lab-MCE* models described above. The hope is that imposing greater parametric constraints during model estimation will lead to a model that performs better (predicts yield more accurately) out-of-sample. Here, finding that one of the models that integrated farm and lab information (i.e., *Lab-Constrained* or *Lab-MCE*) predicted yield out-of-sample no less accurately than the *Farm-Only* model means that a manager would: 1) have at least as much confidence in his integrated yield model as he might have in a model estimated using only his farm's information, 2) especially have confidence in the model-generated yield relationships with managed variables such as N and P (because they are consistent with the selected laboratory's recommendations), and 3) have a model that allows optimal fertilizer rates to vary with other measured variables of interest, allowing VRA N and P to proceed using other variables besides only *STN* and *STP*.

To test the out-of-sample prediction accuracy of the *Farm-Only*, *Lab-Constrained*, and *Lab-MCE* models discussed above, a jackknife approach was used. In this case, the jackknife approach predicts all yield observations in one year, using a model estimated with only information from the other years. After each yield observation for the example farm was predicted, the entire predicted series was compared to the actual yield series using three measures of prediction accuracy: 1) RMSE; 2) r^2; and 3) MAPE (mean absolute percent error), which is the average absolute percent error, where an absolute percent error is the absolute value of an error, divided by the actual value, times 100.

Although details are not shown here, estimated *Farm-Only*, *Lab-Constrained*, and *Lab-MCE* models generally had jackknife r^2's in the 0.05 to 0.10 range. Given these values, none of the models explain much out-of-sample wheat yield variability across fields and years – probably because random weather is still a large causal factor. However, this does not preclude using the model to improve fertilizer decisions, which are almost always made without knowing an upcoming growing season's weather. In general, the *Lab-Constrained* models do not predict out-of-sample yields statistically less accurately than the *Farm-Only* model. This supports the idea that yield-prediction accuracy may not have to be substantially sacrificed when lab-based models are used to constrain N and P parameters in estimating a farm-level yield model. More importantly, this farm's manager probably would have as much confidence in the *Lab-Constrained* models as he would in the *Farm-Only* model that did not incorporate lab-based information. Thus, the manager might view the *Lab-Constrained* models as suitable for use in VRA of N or P. Because the N and P fertilizer recommendations and related expected yield response are believable for the manager, this would be an added benefit to implementing a *Lab-Constrained* model.

Typically, the *Lab-MCE* models stand out as being the best models for predicting out-of-sample wheat yields. Three out of four (*OAL-*, *KSU-*, and *CSU-MCE*) are statistically more accurate than the *Farm-Only* model at the 0.95 confidence level. Further, all four *Lab-MCE* models are statistically more accurate than their *Constrained* counterparts. These findings are especially encouraging. They suggests that maximizing cross entropy, a systematic and theoretically well-supported modeling approach, may be a plausible way to incorporate ever better farm-level N and P information over time. As with the *Lab-Constrained* models, assuming that predicted N and P fertilizer recommendations are believable for the farm manager, he might consider the *Lab-MCE* models appropriate for VRA of N or P.

Using an integrated model to make VRA decisions

In general, using one of the *Lab-Constrained* models for making N or P VRA decisions for the example northwest Kansas farm would result in N and P responses very similar to those of the lab-based models discussed earlier in this paper (e.g., like those shown in Fig. 5 regarding P). Thus, *STN* would be especially important for determining optimal *fertN* rates and *STP* for *fertP* rates. Generally, using one of the *Lab-MCE* models for making N or P VRA decisions for the example farm would result in N and P responses much closer to those of the *Lab-Constrained* models than those of the *Farm-Only* model. As two examples, Fig. 12 and 13 show yield response to *fertN* and *fertP*, respectively, for the *Farm-Only*, *OAL-Constrained*, and *OAL-MCE* models (holding all other variable values at their means). Notice that the *Farm-Only* model shows nearly zero response to *fertN* and *fertP* – likely a tribute to the poor fertilizer response farm data. Notice also that, by design, the *OAL-Constrained* model gives more believable fertilizer response results than the *Farm-Only* model (i.e., closer to those of the laboratory).

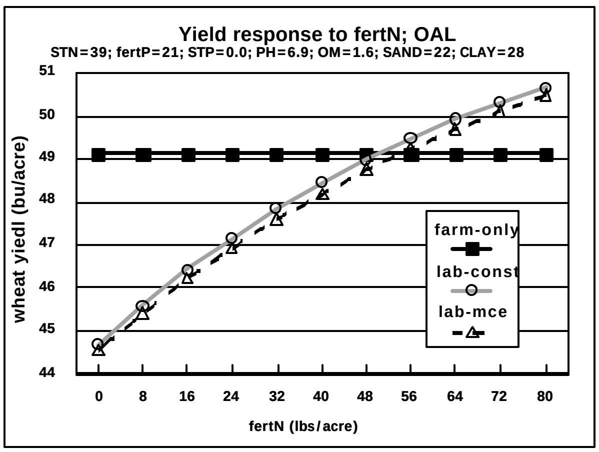

Figure 12

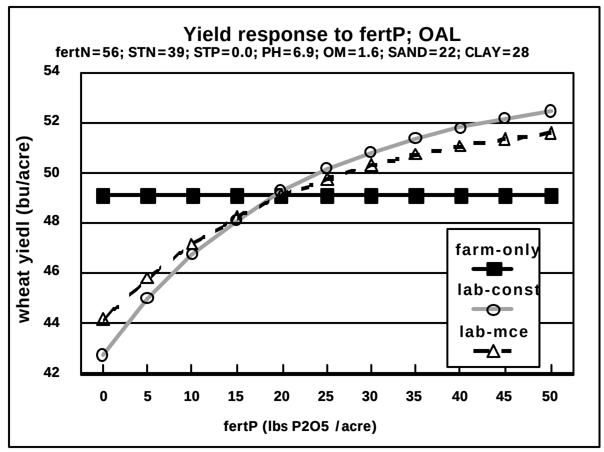

Figure 13

As one example of non-fertility yield response, Fig. 14 shows response to *pH* for the *Farm-Only*, *CSU-Constrained*, and *CSU-MCE* models. Unlike in Figs. 12 and 13, the *CSU-MCE* line is much closer to the *Farm-Only* line than to the *CSU-Constrained* line. Notice also that, in this western Kansas data set, higher *pH* leads to lower yields.

Of special interest for VRA of fertilizer is how much other variables are expected to impact the choice of optimal fertilizer rates. Depending on the model used and the non-fertility variable examined, a number of substantial impacts emerged. As just one example, holding other variables (including *STN* and *STP*) at their means, varying *pH* across its approximate data range while optimally selecting *fertN* and *fertP* resulted in optimal *fertN* rates varying from 60.9 to 45.5 lb N/acre and optimal *fertP* rates varying from 19.0 to 15.9 lb P_2O_5/acre with the *KSU-MCE* model (Fig. 15). In this data set, because *pH* has a negative impact on

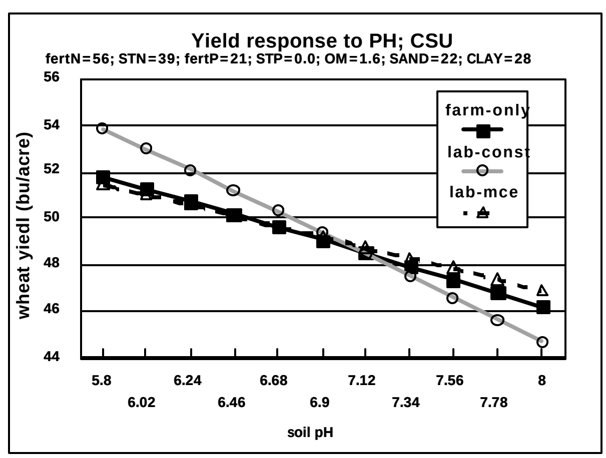

Figure 14

yield, it causes optimally selected *fertN* and *fertP* rates to fall with increased *pH* levels. Clearly, such differences in optimal fertilizer rates would be impossible to assess without farm information on *pH* and without a mathematical model that quantified *pH*'s impact on wheat yield. Further, assuming *pH* changes little over time, the annual cost of site-specific *pH* information could be low enough to make VRA of N and P based on *pH* profitable.

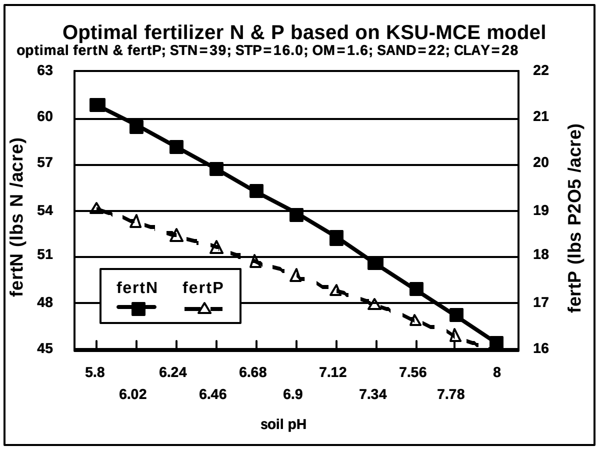

Figure 15

For the example farm, VRA N and P fertilizer rates would depend principally on measures of *STN* and *STP*, respectively, but also on other information. Gains to VRA might be expected from non-N non-P information if it were sufficiently inexpensive. Regardless, the *Lab-Constrained* and *Lab-MCE* modeling approaches would provide this farm's manager with reasonable ways to integrate other information into his fertilizer management program – methods that would not allow his fertilizer rates to stray "too far" from tried and true lab-based recommendations. Additionally, these approaches could easily accommodate new and better farm information as it became available over time.

Summary

Mathematical formulas used to calculate N and P fertilizer rate recommendations from measures of soil test N and P are readily available from university and industry soil testing laboratories. However, as precision farming technologies increase the availability and lower the cost of non soil test information, farm managers will expect recommendations that take advantage of this additional information. For example, for a variable rate application, farm managers want a prescription for N fertilizer rates that depends on measures of soil properties besides soil test N (e.g., soil test P, soil organic matter content, soil texture). Similarly, a prescription for P fertilizer that depends on properties other than soil test P will be necessary to take advantage of new yield-affecting information.

Even when fertilizer rates are optimal, farm managers want to know the economic impact of intrinsic soil properties – to help with land ownership/rental decisions. Although not always requested directly, managers ultimately want a mathematical yield model that shows how crop yield is expected to change conditional on a number of managed and intrinsic causal factors. Moreover, managers likely will want their own farms' data to play a key role in yield model generation. Perhaps most importantly, because the comparative advantage garnered from early adoption of technology dissipates with time, progressive managers will want to quickly adopt information-based crop production decisions. Despite their impatience, managers recognize limitations in their own data, because they rarely use low- or 0-rate fertilizer applications or may not have many years of observations. Thus, they are unlikely to completely ignore laboratory-based fertilizer recommendations in favor of models developed from only their own data.

Using wheat production in northwest Kansas, this research first developed a protocol for generating yield models from fertilizer recommendations from soil testing laboratories. Second, a protocol was developed for combining the N- and P-yield-response from laboratory-based models with farm-level data to generate an integrated yield model that includes numerous causal factors, yet which makes acceptable fertilizer recommendations. Third, to accommodate those farms using on-farm research to gather ever better fertilizer response information with time, this research considers maximizing cross entropy as one theoretically appropriate way of using fertilizer recommendations from soil testing laboratories to modify the N- and P-yield response that would be implied from a model using only farm data.

The analytical procedures discussed and developed here were tested on historical wheat production data from one farm in northwest Kansas. In general, the tests revealed that using these procedures would have provided a yield model that was no less, and possibly more, accurate in terms of predicting wheat yields out-of-sample conditional upon measures of various causal factors. Further, because of particular features of the analytical method, the yield model could be constrained to make only sensible fertilizer recommendations – those consistent with a manager's intuition gained from experience and an understanding of published fertilizer research.

Besides helping farm managers make better crop production decisions, this research should help public and private providers of fertilizer recommendations improve those recommendations. The fertilizer-decision issues discussed here often impact both farm managers and university researchers. Also, fertilizer recommendation providers may want to consider other issues noted here, such as the long-run implication of current fertilizer rates, or how soil test P might become a managed variable along with fertilizer P. Regardless, the demand for improved analytical methods will likely continue to increase as farm managers gather more and more farm-level information.