Replicability of nitrogen recommendations from ramped calibration strips in winter wheat

Roberts, David C.; Brorsen, B. Wade; Taylor, Randal K.; Solie, John B. and Raun, William R.
Oklahoma State University

2011

Replicability of Nitrogen Recommendations

from Ramped Calibration Strips in Winter Wheat

David C. Roberts,[1] B. Wade Brorsen, Randal K. Taylor, John B. Solie, and William R. Raun

David C. Roberts, Ph.D.
Assistant Professor
Dept. of Agribusiness & Applied Economics
North Dakota State University
NDSU Dept. 7160
P.O. Box 6050
Fargo, ND 58108-6050
david.c.roberts@ndsu.edu
Tel: 1-701-231-9796
Fax: 1-701-231-7400

B. Wade Brorsen, Ph.D.
Regents Professor and Jean & Patsy Neustadt Chair
Dept. of Agricultural Economics
Oklahoma State University
414 Agricultural Hall
Stillwater, OK 74078

Randal K. Taylor, Ph.D.
Professor
Dept. of Biosystems and Agricultural Engineering
Oklahoma State University
111 Agricultural Hall
Stillwater, OK 74078

John B. Solie, Ph.D.
Regents Professor and Sarkey's Distinguished Professor
Dept. of Biosystems and Agricultural Engineering
Oklahoma State University
111 Agricultural Hall
Stillwater, OK 74078

William R. Raun, Ph.D.
Regents Professor and Nutrients for Life Professor
Dept. of Plant and Soil Sciences
Oklahoma State University
044 N. Agricultural Hall
Stillwater, OK 74078

[1] Corresponding author

Replicability of Nitrogen Recommendations

from Ramped Calibration Strips in Winter Wheat

Abstract Ramped calibration strips have been suggested as a way for grain producers to determine nitrogen needs more accurately. The strips use incrementally increasing levels of nitrogen and enable producers to conduct an experiment in each field to determine nitrogen needs. This study determines whether predictions from the program Ramp Analyzer 1.2 are replicable in Oklahoma hard red winter wheat (*Triticumaestivum*). Predictions are derived from 36 individual strips from on-farm experiments—two pairs of adjacent strips at each of nine winter wheat fields in Canadian County, OK. The two pairs of strips within each field were between 120 and 155 m apart. Each strip was analyzed three times during the 2006-2007 growing season. Nitrogen recommendations from Ramp Analyzer 1.2 are not correlated even for strips that were placed side by side, and recommendations from strips in the same field show no more homogeneity than randomly selected strips throughout the county. The results indicate that ramped calibration strips are unlikely to produce accurate nitrogen requirement predictions at any spatial scale, whether at the county level or for subsections of a single field. In contrast, a procedure that uses only measures from the plot with no nitrogen and the plot with the highest level of nitrogen applied does show replicability. Thus, improvements in the ramped calibration strip technology are needed if it is to become viable.

Key words Fertilizer · Nitrogen · Precision Agriculture · Ramped Calibration Strip · Winter Wheat

Introduction

Potential benefits to producers and the environment have motivated a flurry of research, as well as the development of precision agriculture technologies and techniques to improve nitrogen-use efficiency (NUE) (Cassman et al. 1998; Greenhalgh and Faeth 2001; López-Bellido and López-Bellido 2001; López-Bellido et al. 2004; Raun and Johnson 1999). Low NUE is due substantially to producers' response to the uncertainty of nitrogen (N) requirements across space and time (Babcock 1992; Tembo et al. 2008). Producers 'over-apply' N in most years and fields because they want to ensure that enough N is available if crop N requirements are higher than expected.

A ramped calibration strip (RCS) system has been designed to predict crop N requirements based on midseason measures of the normalized difference vegetation index (NDVI) from growing wheat (Arnall et al. 2008; Raun et al. 2008). The RCS system is based on the assumption that NDVI data at Feekes Growth Stage 5 (Large 1954) are directly related to wheat grain yield (Raun et al. 2008). The Oklahoma Cooperative Extension Service recommends that, shortly after planting, producers who use the RCS prediction technique apply incrementally increasing N rates on one or more strips in a field (Arnall et al. 2008). The RCS applicator applies 16 incremental N rates, increasing the N rates sequentially from one end of the strip to the other (Raun et al. 2008). Some N may be applied to the entire field before creating the RCS to prevent early season N stress, or the strip may be superimposed on the rate the farmer usually applies (Arnall et al. 2008; Raun et al. 2008). Producers can either use visual inspection or a hand-held Greenseeker sensor at Feekes Growth Stage 5 (Large 1954) to predict the N application rate at which yield will cease to respond to topdressed N. A large part of the impetus behind the development of the RCS method was producer receptiveness to other visual methods for determining midseason topdressed N requirements (Raun et al. 2008), such as the calibration stamp technology developed by Raun et al. (2005b). To date, no published research has tested the accuracy of RCS-based predictions of crop N requirements, although the RCS system is already in use by some winter wheat producers who have built their own RCS applicators (Raun et al. 2008).

If NDVI data recorded from a RCS can reduce uncertainty sufficiently about crop N requirements—including uncertainty caused by spatial variability and by prediction and measurement errors—the RCS might pay for itself by reducing average annual N expense and increasing average yield. The cost of using a RCS might be reduced further if the RCS predicts strong spatial correlation between N requirements for fields in a region. Ample research has shown that average crop yields (and yield potential) vary at regional and sub-regional scales. For example, based on the European Nomenclature of Territorial Units for Statistics (NTUS), Bakker et al. (2005) showed that average yield not only varies significantly between large territorial units but also between smaller subdivisions of these, and that the variation in yield is strongly correlated with soil and climate characteristics. Moen et al. (1994) developed a crop simulation model to predict regional yield potential based on field-level simulation with planting date, soil type and climate data as modeling inputs, pointing out that regional predictions of yield potential could assist producers in identifying reasonable yield goals and economically optimal N application rates. Similarly, Easterling et al. (1998) created a model to predict wheat yield in the US Great Plains using climate and wheat yield data from 1984 to 1992, and found that their model had the highest predictive power when climate data were disaggregated at a spatial scale of 84 km by 111 km. They also found that spatial disaggregation of soil data did not significantly improve model fit. Although the variation of mineral N availability has also been well-documented within individual fields (Blackmer et al. 1996; Scharf et al. 2005; Stenberg et al. 2005), and even at resolutions of less than one m^2 in Oklahoma (Solie et al. 1999), several studies have determined that it is not *always* profitable to address small-scale spatial variation (Batte2000; Biermacher et al. 2009; Boyer et al. 2010).Thus, a producer might need only one RCS per five fields (if they are close together, similar in soil type, and planted on the same date), rather than five or more strips for the same set of fields, as currently recommended (Arnall et al. 2008). Perhaps RCS data could even be collected at experimental stations and then disseminated at no cost to producers. Region-level information from RCS experiments might be especially valuable to producers who grow wheat for both grain and grazing—a common practice in

Oklahoma—because they might find establishing calibration strips to be prohibitively costly due to the need for fencing to exclude cattle. The objective of this paper is to determine the statistical relationships between N recommendations based on RCS predictions at different spatial scales (within and between fields) and at varying times during the same growing season, and to use this statistical analysis to draw inferences regarding the suitability of the RCS method for predicting crop N requirements at the single-field and regional levels.

Materials and methods

The dataset used consists of observations from nine on-farm RCS trials conducted in Canadian County, Oklahoma. The data are from farmers' fields that were planted in the autumn of 2006. Two pairs of RCSs were established in each field using topdress urea-ammonium nitrate solution (UAN) shortly after plant emergence. Paired strips were created by making two adjacent passes over the field with the RCS applicator, therefore, the rates in the paired strips increased in opposite directions. The strips began with 242 kg ha^{-1}, and every 3 m the N was reduced in 16 equal steps. No N was applied on the seventeenth section. The maximum rate of 242 kg ha^{-1} was then applied on an eighteenth section to help in locating the end of the strip. The strips were 4 m wide. The two pairs of strips in each field ranged from 120 to 155 m apart. Figure 1 shows the basic RCS placement in the on-farm trials in this dataset.

The NDVI data from each strip were obtained with a hand-held Greenseeker optical sensor three times during the growing season. Readings were taken every 0.1 seconds from 0.6 m wide by 1 cm long sections as the operator walked the down the center of each strip, taking approximately 150 readings. The same operator made each reading, so operator variability was not a source of error. The operator was trained to avoid areas of the strip where bare soil was visible.

The program Ramp Analyzer 1.2 fitted a linear response-plateau function to the NDVI data to determine N requirements (Raun et al. 2008). For each RCS reading, N input requirements were predicted in two ways: 1) by direct use of the fitted linear response-plateau parameters to determine the application rate at which NDVI response to N ceased (hereafter called the RCS recommendation) and 2) by using a N fertilizer optimization algorithm (NFOA) developed by Raun et al. (2005a) (hereafter called the NFOA recommendation). Thus, three RCS and three NFOA recommendations were available for each RCS—one of each type at each of three dates.

A linear plateau model was defined mathematically as

$$y = \min(a + bN, Plateau), \tag{1}$$

where y is NDVI, N is N applied and *Plateau* is the NDVI plateau at which additional N had no effect on NDVI. Ramp Analyzer 1.2 used a heuristic estimation procedure for the parameters in Eq. 1. The program for this procedurewas written in Visual Basic in such a way that it can be solved on a personal digital assistant (PDA). The parameters were selected based on least squares regressions of NDVI against the number of readings taken at that point. These regressions were computed as

$$NDVI_i = \beta_0 + \beta_1 i + \varepsilon_i, \tag{2}$$

where the coefficients β_0 and β_1 from this equation are analogous to a and b, respectively, from Eq. 1. To select the intercept and slope parameters, parameters of Eq. 2 were estimated with the first few readings from the strip, and then further values were added one at a time. The intercept and slope parameters were selected from the regression with the largest R^2. The plateau level of NDVI was estimated by starting with the last few readings and then further values were added one at a time. The *Plateau* in Eq. 1 was the value of β_0 from the regression in which the slope parameter, β_1, in Eq. 2 was closest to zero. The predicted optimal value of N was the value of N corresponding to the point where the two functions crossed, *i.e.* $(Plateau - a) / b$.

The NFOA used only the NDVI values from the part of the strip with no N and those from the part of where the highest level was applied. The NFOA had a maximum predicted yield of 6048 kg ha^{-1} (90 bu ac^{-1})—the approximate biological maximum yield for rain-fed hard red winter wheat in Oklahoma (Raun et al. 2002)—even when the predicted intercept was above this level. Such censoring may mean that the NFOA predicted no N response even when the raw NDVI data clearly showed one. Table 1 lists the planting dates and sensing dates for each field. These data were used to determine how repeatable NFOA and RCS recommendations were over space and across sensing dates within fields as a measure of the amount of noise present in the predictions.

The important question of repeatability of RCS and NFOA recommendations across time (within a single growing season) and space was addressed. Poor repeatability of these recommendations at the same strip over time, or weak correlation between recommendations from two adjacent strips would indicate that the RCS or NFOA recommendations were too noisy to be useful in predicting N requirements at the single-field level. Such noise could stem from either measurement error or considerable spatial variability within the field.

Graphical analyses and correlation coefficients were used to determine the strength and significance of the relationships between both RCS and NFOA recommendations from: 1) strips in the same pair at the same sensing date, 2) different pairs (mean recommendation) in the same field at the same sensing date and 3) the same strip at the second and third sensing dates. The second and third sensing dates were chosen because the second date was usually closest to Feekes 5 (Large 1954)—the growth stage at which topdressed N is usually applied—and because the third sensing date (usually in March) was closest to harvest, and might therefore have been the most accurate. The correlation and plot of the relationship between RCS and NFOA recommendations at the same strip for the same sensing date were also calculated.

To provide further statistical confirmation of the graphical results, a Tobit model (Greene 2008) was used because the N recommendations are truncated at zero to avoid recommending a negative amount of N. The following no-intercept Tobit model was estimated:

$$r_{ijt} = \begin{cases} \sum_{j=1}^{J} \gamma_j D_j & \text{if } \; r_{ijt}^{*} = \sum_{j=1}^{J} \gamma_j D_j + \varepsilon_{ijt} > 0 \\ 0 & \text{if } \; r_{ijt}^{*} = \sum_{j=1}^{J} \gamma_j D_j + \varepsilon_{ijt} \leq 0, \end{cases} \tag{3}$$

where r_{ijt} is the recommended N application rate on strip *i* in pair *j* on sensing date *t*, γ_j is a fixed effect for pair *j*, D_j is an indicator variable equal to one for pair *j* and zero otherwise, r_{ijt}^{*} is a latent variable representing the level of N (including residual and applied N) the plants in strip *i* in pair *j* on sensing date *t* need to reach the predicted plateau yield, ε_{ijt} is a random error term distributed with mean zero and variance σ_{ε}^{2} and *J* is the number of strip pairs.

The first hypothesis tested was that N requirement predictions from the RCS did not vary between pairs within the same field, i.e. $\gamma_1 = \gamma_2, \gamma_3 = \gamma_4, \ldots, \gamma_{J-1} = \gamma_J$. Rejection of this hypothesis would indicate that predicted N requirements from the RCS varied consistently by pair within each field. Failure to reject the hypothesis would indicate that RCS recommendations from strips in one pair are no more homogeneous than randomly selected recommendations from strips in the field, because either there was little variation in actual N requirements among locations within a field or the RCS was not precise enough to detect this variation. If the RCS lacks predictive power because of noise, this means it cannot be used successfully to predict N requirements for nearby fields. Next, the model was restricted so that predicted N requirements could not vary by field, i.e. $\gamma_j = \gamma_y, \forall\ j, y,$ to determine whether the RCS detected significant variation in N requirements between fields. Failure to reject this restriction would indicate that RCS recommendations within one field are no more homogeneous than randomly selected recommendations from Canadian County as a whole. Equation 3 was then re-estimated using the NFOA predictions as the dependent variable (r_{ijt}) to determine whether the NFOA recommendations varied consistently within and between fields.

Results and discussion

Figure 2 shows scatter plots and the correlations of N recommendations from strips in the same pair at the same sensing date for the RCS (Fig. 2a) and NFOA (Fig. 2b). The correlation between RCS recommendations from adjacent strips in Fig. 2a is slightly negative, although not significant (p = 0.61). This result indicates that the RCS is a noisy predictor of N requirements. On the other hand, the correlation between NFOA recommendations from adjacent strips in Fig.

2b is 0.56, and is statistically significant ($p < 0.01$). Figure 3a shows the mean RCS recommendation from one pair of strips plotted against the mean RCS recommendation from the other pair of strips in the same field at the same sensing date. Figure 3b plots the NFOA recommendations in the same way. The mean RCS recommendations from pairs in the same field have a weak correlation (0.01) that is not statistically significant ($p = 0.98$). However, the mean NFOA recommendations from the different pairs are strongly correlated ($r = 0.74$, $p < 0.01$). Because RCS recommendations from *adjacent* strips are not correlated, lack of correlation between mean recommendations from strips in the same field probably indicates that variation in RCS recommendations was caused predominantly by noisy measurements rather than by actual variation in N requirements within the field. Although N requirements almost certainly varied throughout and amongst the trial fields (Lobell et al. 2005; Mamo et al. 2003; Scharf et al. 2005), the RCS recommendations did not match expected patterns of spatial variation in N requirements, e.g. strong positive correlation between N requirements at proximal locations, or even positive correlation between N requirements in different parts of the same field.

Figures 4a and 4b show scatter plots of recommendations at the same strip at the second sensing date (usually February) and the third sensing date (usually March) for the RCS and NFOA, respectively. For the RCS, the correlation is 0.10 and is not statistically significant ($p = 0.57$). The correlation for the NFOA recommendations is 0.56 and is significant at the 0.01 confidence level. The weak correlation between N recommendations from the same RCS in February and March is particularly disconcerting as it indicates that a RCS is likely to give widely disparate N recommendations on different dates. Thus, in this set of fields in 2006, the RCS was almost certainly not an adequate predictor of crop N requirements. One caveat is that, while our dataset includes nine fields, it includes only one year. A single county-wide weather event in 2006 could have caused the weak correlation between RCS recommendations at the different sensing dates. Even so, given the lack of expected spatial patterns in RCS N recommendations, lack of temporal correlation between recommendations from the same strip is not surprising.

One reason the NFOA recommendations show stronger spatial and temporal correlation may be the NFOA's propensity to predict optimal rates of zero kg ha^{-1}. The NFOA restricts the predicted plateau yield for each strip to be no greater than 6048 kg ha^{-1}. Thus, when the NFOA predicts a yield intercept >6048 kg ha^{-1}, the predicted plateau yield is still no greater than 6048 kg ha^{-1}, without regard to NDVI response to N.

Figure 5 shows a scatter plot of the NFOA recommendations against those of RCS from the same strip at the same sensing date. Note that the NFOA often recommends no N application, whereas the RCS recommends applying N in 36 out of 100 observations. Even when NDVI data indicate an N response, i.e. the average NDVI reading at one end of the strip is different from that at the other end, the NFOA still assumes no N response by assuming that the relationship between NDVI and yield is estimated without error.

The estimated parameters of Eq. 3 for the RCS recommendations are given in Table 2. The model with pair effects allows the mean predicted N requirement to be unique for each pair of adjacent strips, whereas the model with field effects is restricted such that pairs in the same field must have the same mean prediction, and the pooled model assumes the same mean N requirement for all strips in the dataset.

To determine whether the field itself affects the RCS N recommendation, the field effects model was tested against the pooled model using a likelihood ratio test. The null hypothesis that field does not affect RCS N recommendations cannot be rejected (LR = 11.66 < critical value = 13.36, 1 df, α = 0.01), suggesting that the RCS recommendation system does not detect variation in N requirements among fields, or that RCS N recommendations within a field are no more homogeneous than recommendations from randomly selected strips from Canadian county. Because variation in N requirements among fields is well documented (Lobell et al. 2005; Mamo et al. 2003; Washmon et al. 2002), this result probably indicates that the RCS system is not precise enough to detect such variability, and cannot make accurate N recommendations at the county-level. The likelihood ratio test to determine whether mean N recommendations vary among pairs of adjacent strips also fails to reject the null hypothesis of no difference (LR = 24.94 < critical value = 27.59, 17 df, α = 0.05). The inference is that recommendations from two adjacent strips in a pair selected at random are no more homogeneous than readings from two randomly selected strips from different pairs—perhaps on opposite sides of Canadian county. The fact that RCS predictions of N requirements do not show more homogeneity within pairs than within the field as a whole indicates that the predictions are imprecise. The lack of replicability of RCS N requirement predictions over space (especially between strips in the same pair) might be caused by considerable small-scale spatial variation in N requirements within and amongst the strips in a single field. However, the finding that N recommendations from the same strip show no repetition on different dates within the same growing season indicates that the RCS N requirement predictions are simply very noisy. It is clear from Raun et al. (2005a) that the relationship between midseason NDVI data and yield not only varies from field-to-field and year-to-year but also that the error terms in the estimated relationship are heteroskedastic, i.e. the variance of the error increases with NDVI. This means that large small-scale variation in NDVI, especially when NDVI values are large, does not necessarily indicate considerable small-scale variation in yield potential or crop N requirements. Yet, such variation in NDVI readings over the length of a RCS could result in systematic errors in the N response function and flawed N recommendations. This issue might be attenuated by using shorter RCSs, which would presumably exhibit less within-strip spatial variation in NDVI response to N.

Table 3 gives the mean N application rate recommended by the NFOA with and without fixed effects for the strip pair and field, as estimated by Eq. 3. The null hypothesis that NFOA N recommendations do not differ among fields is rejected (LR = 113.48 > critical value = 20.09, 8 df, $\alpha = 0.01$). Thus, it is possible that the NFOA detects actual variation in N requirements in different fields. Including pair effects slightly improves the model fit relative to field effects alone (LR = 19.06 > critical value = 16.92, 9 df, $\alpha = 0.05$), providing evidence of modest spatial variation within the fields. The clear replicability of the NFOA recommendations provides evidence that the lack of replicability of the RCS recommendations is not because of some anomaly in the sites.

Why are the RCS predictions so inaccurate when the RCS uses more observations than does the NFOA? One possible source of error is the heuristic estimation algorithm that was used. It may be possible to obtain more accurate estimates of the parameters of the linear plateau model with nonlinear least squares or nonlinear maximum likelihood. Another possible weakness occurs when an area being sensed has bare ground. These cases create errors with large negative skewness, and the estimation procedure used by Ramp Analyzer 1.2 might be overly sensitive to these observations. Estimating a stochastic frontier function as in Coelli (1996) would provide one way to remove the effects of these outliers. Another problem can occur when the response to N is almost flat. In this case, even a small change in measurement can change the recommendation from a very high level of N to a very low level.

Conclusions

First and foremost, the results indicate that the use of Ramp Analyzer 1.2 and RCS is too noisy a method to be useful to predict accurately and consistently optimal N application levels at any spatial scale, i.e. at the regional, single-field or sub-field levels. Thus, the RCS method needs substantial modification to improve its predictive power. Potential ways to improve the RCS approach include using shorter strips and more statistically advanced estimation procedures. Incorporating RCS data into a Bayesian statistical framework could improve the value of the RCS method by enabling it to be used in conjunction with other readily available data, such as historical yields and weather data. Research that validates the predictive power of the RCS technology is lacking, although many producers find the visual information acquired from the strips appealing. What is clear from the results of this analysis is that further research and development are necessary to improve and verify the accuracy of RCS N recommendations.

Acknowledgements

This research was partially funded by the Target Research Initiative Program of the Oklahoma Agricultural Experiment Station. The authors thank the editor of *Precision Agriculture*, as well as the two anonymous peer reviewers for their helpful insights and suggestions.

References

Arnall, D. B., Edwards, J. T., & Godsey, C. B. (2008). *Reference strip series: Applying your nitrogen-rich and ramped calibration strips*. Stillwater, OK, USA: Oklahoma Cooperative Extension Service Fact Sheet CR-2255.

Babcock, B. A. (1992). The effects of uncertainty on optimal nitrogen applications. *Review of Agricultural Economics*, *14*, 271-280.

Bakker, M. M., Govers, G., Ewert, F., Rounsevell, M., & Jones, R. (2005).Variability in regional wheat yields as a function of climate, soil and economic variables: Assessing the risk of confounding. *Agriculture, Ecosystems and Environment*, *110*, 195-209.

Batte, M. T. (2000). Factors influencing the profitability of precision farming systems. *Journal of Soil and Water Conservation*, *55*, 12-18.

Biermacher, J., Epplin, F. M., Brorsen, B. W., Solie, J. B., & Raun, W. R. (2009). Economic feasibility of site-specific optical sensing for managing nitrogen fertilizer for growing wheat. *Precision Agriculture*, *10*, 213-230.

Blackmer, T. M., Schepers, J. S., Varvel, G. E., & Meyer, G. E. (1996). Analysis of aerial photography for nitrogen stress within corn fields. *Agronomy Journal*, *88*, 729-733.

Boyer, C. N., Brorsen, B. W., Solie, J. B., & Raun, W. R. (2010).Profitability of variable rate nitrogen application in wheat production. *Precision Agriculture*, http://dx.doi.org/10.1007/s11119-010-9190-5.

Cassman, K. G., Peng, S., Olk, D. C., Ladha, J. K., Reichardt, W., Doberman, A., & Singh, U. (1998). Opportunities for increased nitrogen-use efficiency from improved resource management in irrigated rice systems. *Field Crops Research*, *56*, 7-39.

Coelli, T. (1996). A guide to FRONTIER version 4.1: A computer program for frontier production and cost function estimation. CEPA Working Paper 96/07. Armidale, Australia: University of New England.

Easterling, W. E., Weiss, A., Hays, C. J., & Mearns, L. O. (1998). Spatial scales of climate information for simulating wheat and maize productivity: The case of the US Great Plains. *Agricultural and Forest Meteorology*, *90*, 51-63.

Greene, W. H. (2008). *Econometric Analysis*. Sixth Edition. Upper Saddle River, NJ: Prentice-Hall.

Greenhalgh, S., & Faeth, P. (2001). A potential integrated water quality strategy for the Mississippi River Basin and the Gulf of Mexico. *The Scientific World*, *1*, 976-983.

Large, E. C. (1954). Growth stages in cereals: Illustration of the Feekes Scale. *Plant Pathology*, *3*, 128-129.

Lobell, D. B., Ortiz-Monasterio, J. I., Asner, G. P., Naylor, R. L., & Falcon, W. P. (2005). Combining field surveys, remote sensing, and regression trees to understand yield variations in an irrigated wheat landscape. *Agronomy Journal*, *97*, 241-249.

López-Bellido, R. J., & López-Bellido, L. (2001). Efficiency of nitrogen in wheat under Mediterranean conditions: Effect of tillage, crop rotation and N fertilization. *Field Crops Research*, *71*, 31-46.

López-Bellido, R. J., Shepherd, C. E., & Barraclough, P. B. (2004). Predicting post-anthesis N requirements of bread wheat with a Minolta SPAD Meter. *European Journal of Agronomy*, *20*, 313-320.

Mamo, M., Malzer, G. L., Mulla, D. J., Huggins, D. R., & Strock, J. (2003). Spatial and temporal variation in economically optimum nitrogen rate for corn. *Agronomy Journal*, *95*, 958-964.

Moen, T. N., Kaiser, H. M., & Riha, S. J. (1994). Regional yield estimation using a crop simulation model: Concepts, methods and validation. *Agricultural Systems*, *46*, 79-92.

Raun, W. R., & Johnson, G. V. (1999). Improving nitrogen use efficiency for cereal production. *Agronomy Journal*, *91*, 357-363.

Raun, W. R., Solie, J. B., Johnson, G. V., Stone, M. L., Mullen, R. W., Freeman, K. W., Thomason, W. E., & Lukina, E. V. (2002). Improving nitrogen use efficiency in cereal grain production with optical sensing and variable rate application. *Agronomy Journal*, *94*, 815-820.

Raun, W. R., Solie, J. B., Stone, M. L., Martin, K. L., Freeman, K. W., Mullen, R. W., Zhang, H., Schepers, J. S., & Johnson, G. V. (2005a). Optical sensor-based algorithm for crop nitrogen fertilization. *Communications in Soil Science and Plant Analysis*, *36*, 2759-2781.

Raun, W. R., Solie, J. B., Stone, M. L., Zavodny, D. L., Martin, K. L., & Freeman, K. W. (2005b). Automated calibration stamp technology for improved in-season nitrogen fertilization. *Agronomy Journal*, *97*, 338-342.

Raun, W. R, Solie, J. B., Taylor, R. K., Arnall, D. B., Mack, C. J., & Edmonds, D. E. (2008). Ramp calibration strip technology for determining midseason nitrogen rates in corn and wheat. *Agronomy Journal*, *100*, 1088-1093.

Scharf, P. C., Kitchen, N. R., Sudduth, K. A., Davis, J. G., Hubbard, V. C., & Lory, J. A. (2005). Field-scale variability in optimal nitrogen fertilizer rate for corn. *Agronomy Journal*, *97*, 452-461.

Solie, J. B., Raun, W. R., & Stone, M. L. (1999). Submeter spatial variability of selected soil and bermudagrass production variables. *Soil Science Society of America Journal*, *63*, 1724-1733.

Stenberg, B., Jonsson, A., & Börjesson, T. (2005). Use of near infrared reflectance spectroscopy to predict nitrogen uptake by winter wheat within fields with high variability in organic matter. *Plant and Soil*, *269*, 251-258.

Tembo, G., Brorsen, B. W., Epplin, F. M., & Tostão, E. (2008). Crop input response functions with stochastic plateaus. *American Journal of Agricultural Economics*, *90*, 424-437.

Washmon, C. N., Solie, J. B., Raun, W. R., & Itenfisu, D. D. (2002). Within field variability in wheat grain yields over nine years in Oklahoma. *Journal of Plant Nutrition*, *25*, 2655-2662.

Table 1. Planting date and sensing dates for each field

Field	Planting date	Sensing dates
1	11/6/2007	01/31/2008
2	10/10/2007	02/01/2008
		02/19/2008
		03/11/2008
3	10/14/2007	01/31/2008
		02/19/2008
		03/11/2008
4	10/12/2007	01/31/2008
		02/20/2008
		03/11/2008
5	10/5/2007	01/31/2008
		02/19/2008
		03/11/2008
6	10/9/2007	01/23/2008
		01/31/2008
		02/19/2008
7	10/12/2007	02/04/2008
		02/19/2008
		03/11/2008
8	10/12/2007	02/04/2008
		02/20/2008
		03/11/2008
9	10/10/2007	01/31/2008
		02/19/2008
		03/11/2008

Table 2. Mean ramped strip recommendation, with and without fixed effects for strip pair and field

Parameter	Definition	Model		
		Pair effects	Field effects[a]	Pooled[b]
γ_1	Fixed effect for pair 1	10.08 (21.18)	19.04 (16.07)	35.19***[c] (3.42)[d]
γ_2	Fixed effect for pair 2	28.00 (21.18)	19.04 (16.07)	35.19*** (3.42)
γ_3	Fixed effect for pair 3	65.15*** (12.23)	48.91*** (9.28)	35.19*** (3.42)
γ_4	Fixed effect for pair 4	32.67 (12.23) ***	48.91*** (9.28)	35.19*** (3.42)
γ_5	Fixed effect for pair 5	59.36*** (12.23)	49.75*** (9.28)	35.19*** (3.42)
γ_6	Fixed effect for pair 6	40.13*** (12.23)	49.75*** (9.28)	35.19*** (3.42)
γ_7	Fixed effect for pair 7	35.47*** (12.23)	23.07** (9.37)	35.19*** (3.42)
γ_8	Fixed effect for pair 8	10.40 (12.54)	23.07** (9.37)	35.19*** (3.42)
γ_9	Fixed effect for pair 9	26.88** (12.23)	31.08*** (9.28)	35.19*** (3.42)
γ_{10}	Fixed effect for pair 10	35.28*** (12.23)	31.08*** (9.28)	35.19*** (3.42)
γ_{11}	Fixed effect for pair 11	35.47*** (12.23)	34.91*** (9.28)	35.19*** (3.42)
γ_{12}	Fixed effect for pair 12	34.35*** (12.23)	34.91*** (9.28)	35.19*** (3.42)
γ_{13}	Fixed effect for pair 13	16.07 (12.51)	18.61** (9.49)	35.19*** (3.42)
γ_{14}	Fixed effect for pair 14	21.72* (12.48)	18.61*** (9.49)	35.19*** (3.42)
γ_{15}	Fixed effect for pair 15	24.64** (12.23)	33.13*** (9.28)	35.19*** (3.42)
γ_{16}	Fixed effect for pair 16	41.63*** (12.23)	33.13*** (9.28)	35.19*** (3.42)

Table 2. Mean ramped strip recommendation, with and without fixed effects for strip pair and field

γ_{17}	Fixed effect for pair 17	68.48***	47.25***	35.19***
		(12.33)	(9.34)	(3.42)
γ_{18}	Fixed effect for pair 18	26.69**	47.25***	35.19***
		(12.23)	(9.34)	(3.42)
σ_{ε}^{2}	Variance of error	29.95***	32.15***	34.05***
		(2.18)	(2.34)	(2.48)
Log Likelihood		-466.79	-473.43	-479.26

Note: Units are kg ha^{-1}.

[a] This model is restricted such that $\gamma_1 = \gamma_2, \gamma_3 = \gamma_4, \gamma_5 = \gamma_6, \ldots, \gamma_{17} = \gamma_{18}$.

[b] This model is restricted such that $\gamma_1 = \gamma_2 = \gamma_3 =, \ldots, = \gamma_{18}$.

[c] One, two or three asterisks (*) indicate statistical significance at the 0.10, 0.05 or 0.01 confidence level, respectively.

[d] Numbers in parentheses are standard errors.

Table 3. Mean nitrogen fertilizer optimization algorithm recommendation, with and without fixed effects for strip pair and field

		Model		
Parameter	Definition	Pair effects	Field effects[a]	Pooled[b]
γ_1	Fixed effect for pair 1	7.28 (19.12)	10.64 (15.47)	16.11***[c] (6.24)[d]
γ_2	Fixed effect for pair 2	14.00 (19.12)	10.64 (15.47)	16.11*** (6.24)
γ_3	Fixed effect for pair 3	-156.84 (0.00)	-46.00*** (17.01)	16.11*** (6.24)
γ_4	Fixed effect for pair 4	-30.53* (16.83)	-46.00*** (17.01)	16.11*** (6.24)
γ_5	Fixed effect for pair 5	63.65*** (11.04)	63.00*** (8.93)	16.11*** (6.24)
γ_6	Fixed effect for pair 6	62.35*** (11.04)	63.00*** (8.93)	16.11*** (6.24)
γ_7	Fixed effect for pair 7	-156.84 (0.00)	-186.66 (0.00)	16.11*** (6.24)
γ_8	Fixed effect for pair 8	-156.84 (0.00)	-186.66 (0.00)	16.11*** (6.24)
γ_9	Fixed effect for pair 9	107.71*** (11.04)	81.48*** (8.93)	16.11*** (6.24)
γ_{10}	Fixed effect for pair 10	55.25*** (11.04)	81.48*** (8.93)	16.11*** (6.24)
γ_{11}	Fixed effect for pair 11	43.12*** (11.04)	26.81*** (9.20)	16.11*** (6.24)
γ_{12}	Fixed effect for pair 12	9.86 (12.00)	26.81*** (9.20)	16.11*** (6.24)
γ_{13}	Fixed effect for pair 13	26.74** (11.23)	39.94*** (9.06)	16.11*** (6.24)
γ_{14}	Fixed effect for pair 14	53.94*** (11.15)	39.94*** (9.06)	16.11*** (6.24)
γ_{15}	Fixed effect for pair 15	-29.28* (16.63)	-45.06*** (16.81)	16.11*** (6.24)
γ_{16}	Fixed effect for pair 16	-156.84 (0.00)	-45.06*** (16.81)	16.11*** (6.24)

Table 3. Mean nitrogen fertilizer optimization algorithm recommendation, with and without fixed effects for strip pair and field

γ_{17}	Fixed effect for pair 17	45.36***	40.42***	16.11***
		(11.17)	(8.99)	(6.24)
γ_{18}	Fixed effect for pair 18	36.03***	40.42***	16.11***
		(11.04)	(8.99)	(6.24)
σ_{ε}^{2}	Variance of error	27.04***	30.93***	55.42***
		(2.53)	(2.89)	(5.56)
Log Likelihood		-292.12	-301.65	-358.39

Note: Units are kg ha^{-1}.

[a] This model is restricted such that $\gamma_1 = \gamma_2, \gamma_3 = \gamma_4, \gamma_5 = \gamma_6, \ldots, \gamma_{17} = \gamma_{18}$.

[b] This model is restricted such that $\gamma_1 = \gamma_2 = \gamma_3 =, \ldots, = \gamma_{18}$.

[c] One, two, or three asterisks (*) indicate statistical significance at the 0.10, 0.05 or 0.01 confidence level, respectively.

[d] Numbers in parentheses are standard errors.

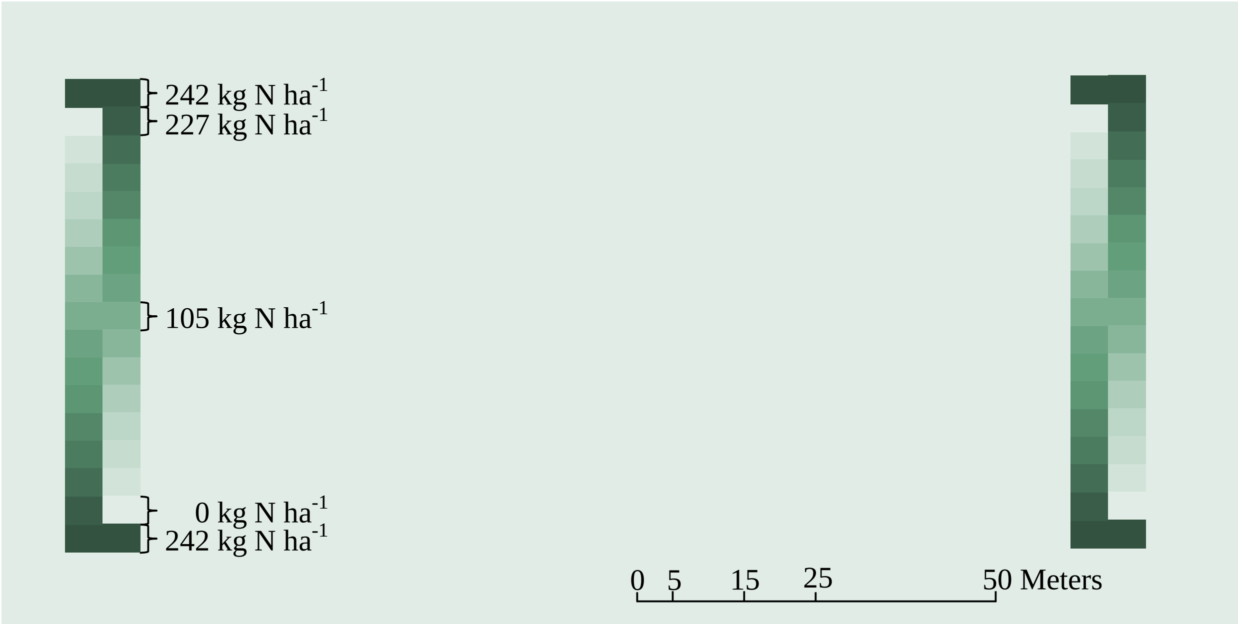

Fig. 1 Diagram of basic placement of two pairs of ramped calibration strips in an on-farm trial.

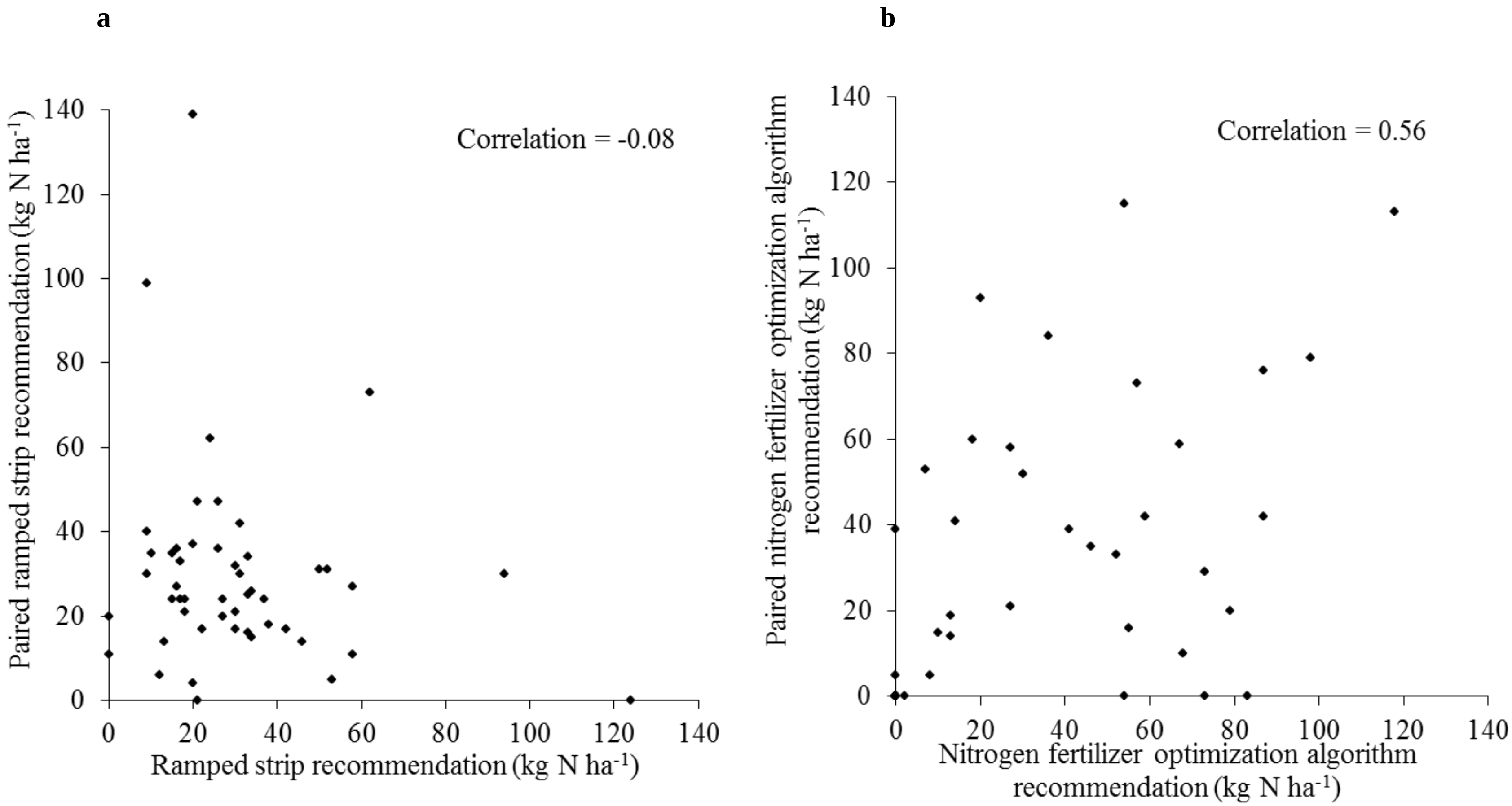

Fig. 2 Ramped calibration strip recommendation for one strip vs. that from the other strip in the same pair at the same sensing date (**a**) and nitrogen fertilizer optimization algorithm recommendation for one strip vs. that from the other strip in the same pair at the same sensing date (**b**) for all nine experiments.

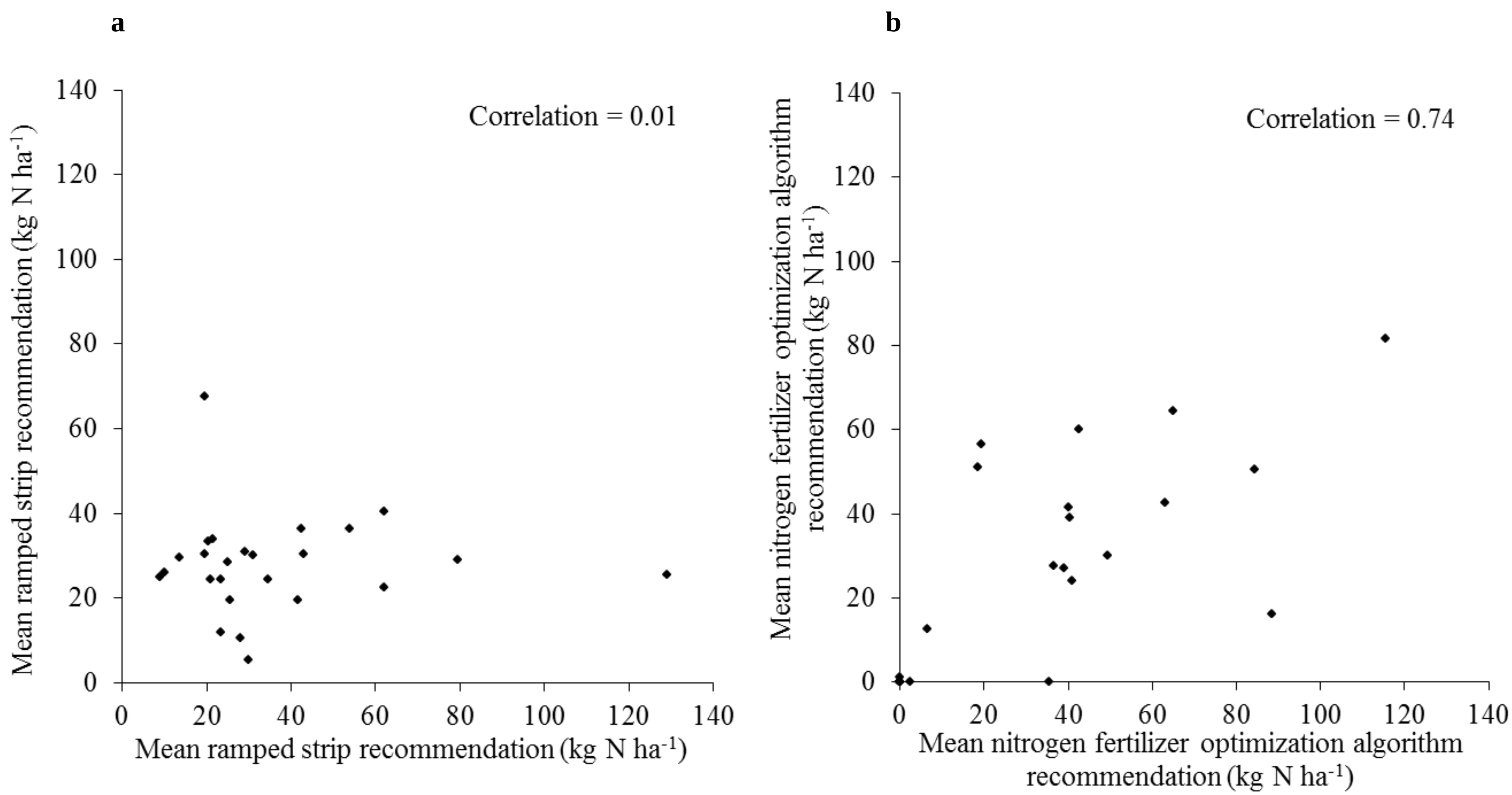

Fig. 3 Mean ramped calibration strip recommendation from one pair of strips vs. that from the other pair in the same field at the same sensing date (**a**) and mean nitrogen fertilizer optimization algorithm recommendation from one pair of strips vs. that from the other pair in the same field at the same sensing date (**b**) for all nine experiments.

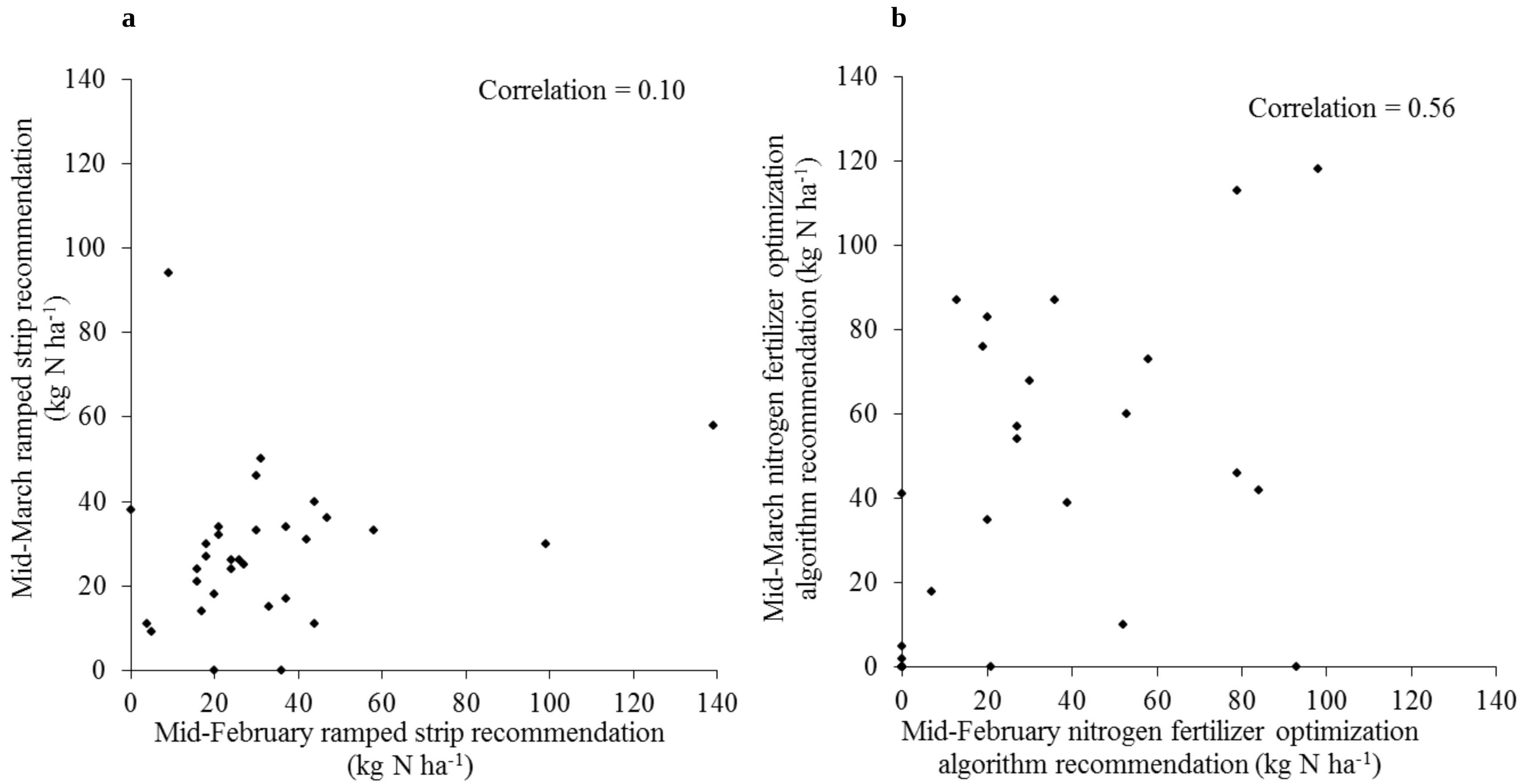

Fig. 4 Ramped calibration strip recommendation in mid-March vs. that from the same strip in mid-February (**a**) and nitrogen fertilizer optimization algorithm recommendation in mid-March vs. that from the same strip in mid-February (**b**) for all nine experiments.

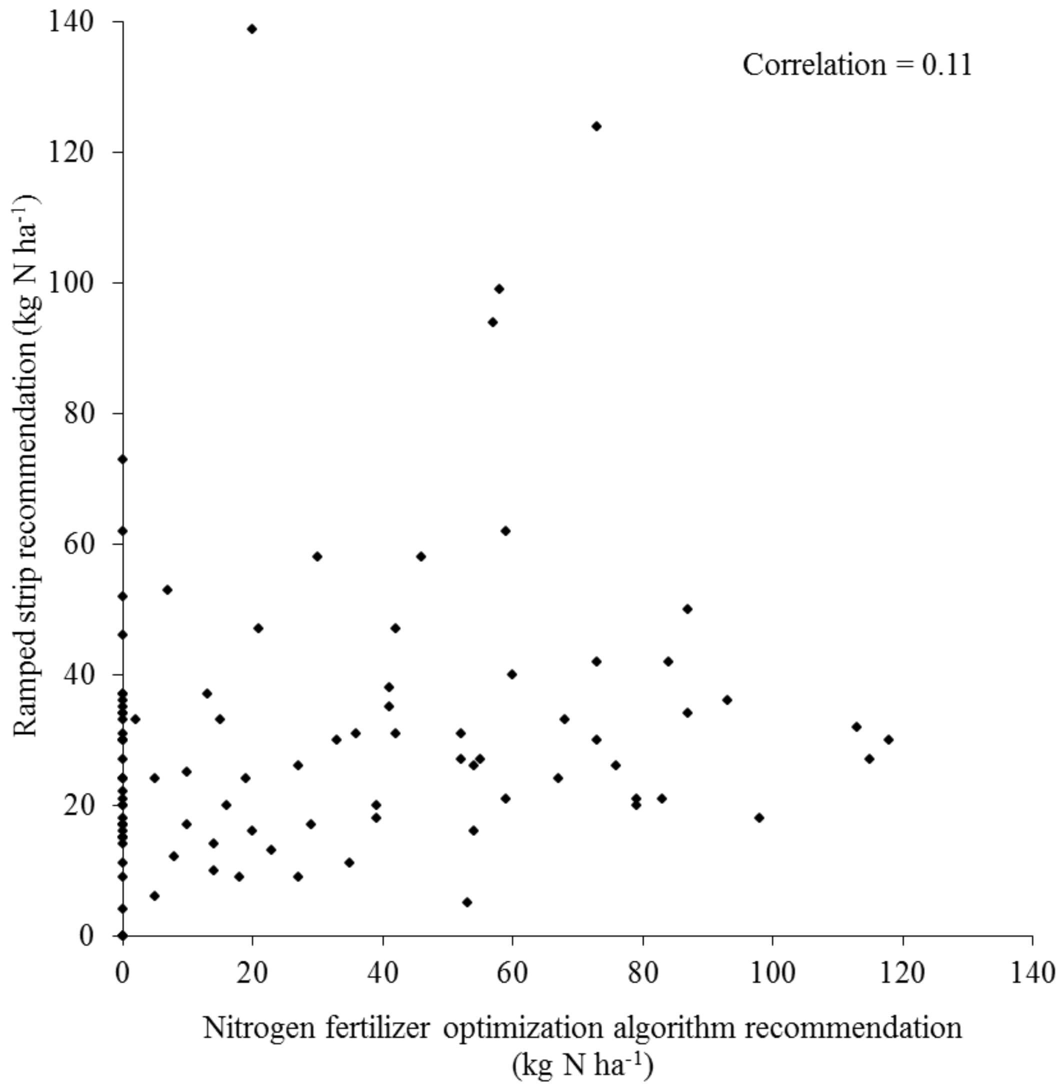

Fig. 5 Ramped strip recommendation vs. nitrogen fertilizer optimization algorithm recommendation from the same strip at the same sensing date for all nine experiments.