

Economically optimal nitrogen rate (EONR) can be predicted from simple two-plot experiments and a national prediction model calibrated with a field trial database

Piikki K.* and Stenberg B.

SLU, Swedish University of Agricultural Sciences, Department of Soil and Environment,
Precision agriculture and pedometrics, P.O. Box 234, SE-532 23 Skara

* Phone: +4651167222; E-mail: kristin.piikki@slu.se

Utkkast

Abstract

A Swedish field trial database was mined for information on economically optimal nitrogen rates (EONR). One hundred wheat (*Triticum aestivum* L.) trials and 47 barley (*Hordeum vulgare* L.) trials in Sweden were used to parameterize prediction models for EONR. Input data to the models were the yield in plots without any nitrogen (N) fertilization, intended to reflect N mineralization, and the yield in plots with a high N rate, intended to reflect the yield potential. An independent validation showed that the prediction models can be expected to predict EONR with a mean absolute error (MAE) <11 kg N / ha for wheat and a MAE < 10 kg N / ha for barley when applied for new sites and years. Simple experiments with two N levels were suggested as a tool to optimize the N fertilization, for example as a complement to the currently performed N fertilization trials to improve the spatial representation. There is also potential to use the model together with optical sensors for in-season EONR predictions at the time for fertilization. A large variation in EONR among sites and years confirmed that it is important to adapt the N fertilization rate to local and current conditions in order to minimize environmental risks and to improve the profit of N fertilization.

Keywords Nitrogen; economically optimal fertilization rate; barley; *Hordeum vulgare*; wheat, *Triticum aestivum*; empirical model; data mining

1. Introduction

Crops respond to nitrogen (N) application with increased yield up to a limit when other factors limit growth. The economically optimal N rate (EONR) is determined as the fertilization rate where the slope of the response curve equals the price ratio of the fertilizer N and the produced grain. Suboptimal fertilization means that the production potential is not exploited, while super optimal fertilization result in a poor profit because of high fertilization costs in relation to the income from the produced grain. EONR is also a threshold for increased environmental risks. Lord & Mitchell (1998) found that fertilization rates above EONR caused increased leaching and Delin & Stenberg (2010) confirmed the results in Swedish experiments. Both studies found that the leaching is not linearly related to the N rate. The nitrate leaching show a weak trend in relation to N for N rates < EONR and a strong accelerating trend in relation to the amount of applied N for N rates > EONR. Thus, it is crucial also from an environmental perspective not to exceed EONR. Tools for local predictions of EONR are necessary to enable optimal fertilization.

Raun et al. (2011) demonstrated that EONR is poorly correlated with yield and Scharf et al (2006) concluded that variation in EONR was to a large extent due to variations in soil N supply and N uptake efficiency. Scharf et al (2006) also calculated that variable rate application of N to corn (*Zea mays*) was much more profitable when the decision support was based on yield and soil N supply together (profit 38\$ ha⁻¹) compared to when it was based on yield alone (profit 2\$ ha⁻¹). This leads us to believe that soil N supply and crop yield potential would be good proxy variables for EONR prediction. These can be represented by 1) yield without N fertilization and 2) yield where N is not limiting to growth.

Today, recommendations for the N fertilization rate are based on a limited number of N fertilization trials with several N levels (usually 4-7) and usually with four replicates.

However, the spatial variation in EONR is expected to be large. The soil N supply by mineralization depends on soil type (Delin and Lindén, 2002) and the variation between and within fields can be $> 100 \text{ kg N ha}^{-1}$. Also crop yield, has been demonstrated to vary considerably, by several tonnes ha^{-1} , within Swedish fields (Algerbo et al., 2003). With the present experimentation to support the advisory services, EONR is determined with high precision at the sites where the experiments are conducted but are not always representative for other farms or fields. Simple experiments for local determination of EONR could potentially constitute a valuable supplement to the currently performed experiments. For this to be doable, the experiments must be simplified but what is lost in precision (how well EONR is predicted) should be more than gained in better representativeness, as the experiments are performed locally.

We propose a method to predict EONR for winter wheat (*Triticum aestivum* L) and barley (*Hordeum vulgare* L.) from N fertilization trials with only two N levels. One level (0 kg N / ha) will give an indication of the soil N supply and the other level (an N rate not limiting to growth) will give an indication of the yield potential. The method capitalizes on a field trial database and the idea is simple: EONR is calculated for each experiment in the database and empirical prediction models are parameterized for prediction of EONR from the yield at the two N fertilization levels.

An important note here is that the yield in the experimental plots is not known at the time for fertilization so if fertilization is to be based on current year experiments, the yield in the plots has to be estimated, e.g. by use of an optical sensor, and this would decrease the accuracy of the two-level N trials.

The purpose of this study was threefold:

1. To examine the magnitude of the variation in EONR, soil N supply (yield without N fertilization) and yield potential (yield without N limitation) between years and between experimental sites.
2. To validate the proposed method to predict EONR from yield at two N fertilization rates and a national prediction model calibrated with a field trial database.
3. To estimate how much the accuracy of the predictions is deteriorated by errors in input data, i.e. to simulate prediction accuracy when using sensor based yield estimates at the time of fertilization instead of measured yield at the time of harvest.

2. Materials & methods

Field trial data

A subset of data from the Swedish field trial database (Field Research Unit, Swedish University of Agricultural Sciences) consisting of 100 wheat trials and 47 barley trials were used. The wheat trials were conducted 1999-2011 and the barley trials were conducted 2002-2011. The selection criteria listed below were applied. No restrictions were made on cultivar or N strategy (number of fertilization times –only the total amount of N was considered). For agronomic details of the individual experiments, search the field trial database of the Field Research Unit at the Swedish university of agricultural sciences (www.ffe.slu.se). Each experiment has a unique Adb-number that can be used in the query. Included Adb numbers are listed in Appendix 1.

Barley:

- Highest N rate at least 160 kg N / ha
- $0 \text{ kg N / ha} < \text{EONR} < 160 \text{ kg N / ha}$

- At least four N levels
- Positioned

Winter wheat:

- Highest N rate at least 180 kg N / ha
- $0 \text{ kg N / ha} < \text{EONR} < 300 \text{ kg N / ha}$
- At least four N levels
- Cereals as precrop
- Positioned

EONR calculation

A second grade polynomial (Equation 1) was fitted for each trial and EONR (Equation 2) was determined as the N rate where the derivative of equation 1 equalled the price ratio of the fertilizer N and the produced grain. This means that at EONR the cost for additionally applied N equals the economic value of the resulting increase in grain yield. At higher N rates any extra N added would not result in a yield increase that would cover the cost of that N and at lower N rates the yield potential would not be utilised. We used a prize ratio of 10 as a case study but models could be parameterized for other prize ratios as well.

$$\text{Grain yield} = i + j \times \text{N rate} + k \times (\text{N rate})^2 \quad \text{Equation 1}$$

$$\text{EONR} = \frac{\text{Price ratio} - j}{2k} \quad \text{Equation 2}$$

Study design

First, five different predictor sets were evaluated for EONR prediction (yield without N fertilization (Y_0) + yield at five different N rates supposed to be non-limiting to growth). Then, effects of error in input data (4 error levels) were simulated for the best predictor set for each crop. Empirical prediction models (multivariate adaptive regression splines; MARSplines) were calibrated for different predictor sets and error levels and validated by the mean absolute error (MAE) and the modelling efficiency (ME).

Evaluating predictor sets

The present study aimed to evaluate empirical prediction models of EONR based on yield. In two experimental plots reflecting soil N supply and yield potential. The soil N supply was supposed to be reflected by Y_0 but it needed to be decided which N rate that best reflected the yield potential. Therefore five different N rates were tested for each crop (Table 1). The yields at the different N rates were calculated for each experiment from the parameterized N response curves (Equation 1).

Error simulation

If empirical prediction models of EONR are to be used as decision support for fertilization the current year, the yield in the two experimental plots will not be known but have to be estimated e.g. from optical sensor measurements. This means that the accuracy of the EONR prediction will decrease. Therefore, for the best predictor set for each crop, four levels of errors were added to the validation data. The errors were randomly sampled from Gaussian distributions with means of zero and standard deviations that equalled 5%, 10%, 20% and 40% of the means of the predictors. Validations were made for errors added to either one or to both of the predictors in the model.

Model calibration

The idea of the present study was that the empirical prediction model will make use relationships between the soil N supply, the yield potential and the EONR that are general for Swedish conditions. Thus new N fertilization trials doesn't have to have enough N-levels to parameterize a new N response curve. Instead, only two essential characteristics of the site and the year (the soil N supply and the yield potential) has to be determined in simple trials with two N levels and EONR will be determined by utilizing empirical knowledge from the large number of previous experiments in the database, parameterized by the prediction models.

MARSplines models were chosen because they are flexible and can describe non-linearities in relationships among variables. Another benefit of MARSplines models is that the calibration procedure includes a pruning step that simplifies the model equation. This minimizes the risk of overfitting. Overfitting means that the model is fitted to non-general relationships that are present only in the calibration dataset. It will make the model explain more of the calibration data but it is still undesirable because the predictions will be less accurate when the model is applied on a new datasets. The model is said to be less robust.

In essence, a MARSplines model is the sum of a number of so-called basis functions. The basis functions are simple univariate linear regressions described by a slope and an intercept but they are only defined above or below a threshold value, a so-called knot, of the predictor variables. MARSplines models can be allowed to include interactions among basis functions but in the present study simple additive models without interactions were parameterized (also because of the risk for over fitting). More information about MARSplines models and the parameterization procedure are found in Hastie et al. (2009) and Milborrow (2013). In the

present study, the MARSplines models were parameterized using the statistical software R (R Development Core Team, 2012), package Earth (Milborrow, 2013)

Method validation

For a prediction model to be of any value, it has to make reliable predictions also when applied for new sites and years not present in the calibration dataset. Therefore a cross-validation strategy was designed as follows: Thirteen (wheat) or ten (barley) different model calibrations were made. Data from one year were withheld from each calibration. Then the model was applied for the withheld year. It was also checked that there were no trials in the calibration dataset that were located close to (< 1 km) any of the trials in the validation dataset. If so, that trial was omitted from the calibration dataset. This procedure means that one prediction is made for each trial but not all predictions are made by the same model (thirteen different models for wheat and ten different models for barley). Thus, it is not a specific model that is evaluated. It rather evaluates the prediction performance of MARSplines models on the present dataset. After the cross-validation, however, final models were calibrated using all data and these can be expected to yield predictions of the accuracy indicated by the cross-validation when applied for new sites (in Sweden) and years.

The error magnitude was quantified by the MAE (Equation 3) and the model performance in relation to using the mean EONR value of the calibration data was quantified by the ME (Equation 4). An ME = 0 indicate that using the model is not better than using the mean of the calibration data while an ME = 1 would be obtained if the model predicted all EONR values correctly.

$$MAE = \frac{\sum |p-m|}{n} \quad \text{Equation 3}$$

$$ME = 1 - \frac{\sum (m-p)^2}{\sum (m-\bar{m})^2} \quad \text{Equation 4}$$

Results

Quantifying variation

It can be inferred from Figure 1 and 2 that there was a considerable variation in EONR, both between years (yearly averages are indicated by the line markers) and between trial sites the same year (each trial is marked by a hollow circle). The range of EONR for wheat was 245 kg N / ha and the range of EONR in barley was 112 kg N / ha.

There was likewise a considerable variation both between sites and between years for the yield without N fertilization, the yield at 300 kg N / ha (wheat) and the yield at 180 kg N / ha (barley). The ranges in yield without N fertilization were 4623 kg / ha (wheat) and 6086 kg / ha (barley) and the ranges at the high fertilization rates were 9920 kg / ha (wheat) and 8945 kg / ha (barley).

Finding the best predictors

Validation measures for the five different predictor sets of each crop are presented in Table 2. For wheat, predictor set 5 (yield without fertilization together with yield in plots fertilized with 300 kg N / ha) had the lowest MAE and the highest ME and was judged to be the best. For barley, predictor set 4 (yield without fertilization and yield in plots fertilized with 180 kg N / ha) was best on the same grounds of judgement. For the best predictor set for each crop, predicted values of EONR are plotted against EONR determined from the N response curves in Figure 3.

Simulating effects of erroneous input data

The results from the simulation of effects of errors in input data are presented in Figure 4. The errors increase rapidly with errors in input data and the models were more sensitive to errors in yield potential (Y_{300} and Y_{180}) compared to errors in Y_0 .

Final models for EONR prediction

The final MARSplines models for EONR prediction are given in Equation 5 (wheat) and Equation 6 (barley).

$$\begin{aligned}
 \text{EONR} = & 135.160907 - 0.055857 \max(0, Y_0 - 3102.38) \\
 & + 0.041956 \times \max(0, Y_0 - 3659.85) \\
 & + 0.029571 \times \max(0, Y_{300} - 5612.13) \\
 & - 0.015401 \times \max(0, Y_{300} - 7597.71) \\
 & + 0.015352 \times \max(0, Y_{300} - 9761.98)
 \end{aligned}
 \tag{Equation 5}$$

$$\begin{aligned}
 \text{EONR} = & 66.547604 - 0.016588 \times \max(0, Y_0 - 3364.44) \\
 & 0.018973 \times \max(0, 3364.44 - Y_0) \\
 & + 0.026882 \times \max(0, Y_{180} - 3988) \\
 & - 0.015294 \times \max(0, Y_{180} - 6361)
 \end{aligned}
 \tag{Equation 6}$$

As the models are only bivariate, it was possible to visualize the Equations as surfaces (Figure 5). The surfaces were curvilinear and somewhat similar for both crops.

3. Discussion

The large magnitude of variation in EONR indicates that recommendations for N fertilization based on average values would be rather erroneous for a certain site (farm/field/management zone) and year. It also shows that experimentation on only a few sites to support the advisory services would be very sensitive to where the trials are located. The presently described method to perform simple experiments with two N-levels could be a valuable complement to the experimentation with more N-levels, in order to better cover the spatial variation in crop N response on a national scale and to provide better advice for farmers based on local conditions.

The setup of the method validation, where the calibration and the validation data are independent (the trials were performed different years and were never closer than 1 km), allows us to draw the conclusion that two-plot experiments performed at new sites and years

can be expected to yield EONR predictions with a MAE < 11 kg N / ha for wheat and a MAE < 10 kg N / ha for barley. However, the presently calibrated models should not be regarded as valid for all times. They should rather be regarded as ‘living models’ that are to be continuously updated using the latest full N fertilization trials performed in the country, in order to be up-to-date regarding crop breeding and agricultural practices.

A drawback of the method –as with all N fertilization trials- is that the EONR determination is based on grain yield and the yield will not be known until harvest. In order to minimize the environmental risk of nitrate leaching and to improve the profit, it is necessary to adjust the N fertilization rate both to *local* and to *current* conditions. This calls for the development of accurate yield prediction methods and there are already promising sensor-based methods to predict the yield in trial plots. For example, Overgaard et al. (2013) found that up to 94% of the yield variation could be predicted from hyper-spectral near infrared reflectance measurements in an independent validation of Norwegian field trials. This was a considerable improvement compared to yield predictions based on reflectance-based vegetation indices. The authors stressed that it was important to include several sites and years in the model calibration. Solie et al. (2012) fitted general yield prediction models to the NDVI of 390 N fertilization trials in wheat and found that yield could be predicted with coefficients of determination (r^2) between 0.1 and 0.8, depending on growth stage.

4. Conclusions

The following conclusions were drawn:

- There is a considerable variation in EONR both between sites and between years. The EONR for wheat differed at most 245 kg N / ha between the wheat experiments and at most 112 kg / ha between the barley experiments. Also the soil N supply (measured

as yield without N fertilization) and the yield potential (yield without N limitation) varied considerably between experimental sites and years.

- Provided that accurate input data are available, EONR can be predicted from simple two-plot experiments and a national prediction model calibrated with a field trial database with a MAE = 11 kg N / ha for wheat and a MAE = 10 kg N / ha for barley.
- The prediction errors increased rapidly with increased error levels in the input data. Simulation showed that EONR predictions were more sensitive to errors in yield without N fertilization compared with errors in yield without N limitation.
- The evaluated method can be used to extend the current experimentation for better geographical coverage. Used together with optical sensors, there is also potential to use the two-plot experiments for EONR prediction at the time for fertilization.

5. References

- Algerbo, P.A., Mattson, L., and Thylén, L. (2003). "Skörderelaterad kvävegödsling-teknik, metodik och erfarenheter." JTI-institutet för jordbruks- och miljöteknik.
- Delin, S., and Lindén, B. (2002). Relations between net nitrogen mineralization and soil characteristics within an arable field. *Acta Agriculturae Scandinavica Section B-Soil and Plant Science* **52**, 8.
- Delin, S., and Stenberg, M. (2010). Effect of nitrogen fertilization on nitrate leaching in relation to grain yield response in Sweden. In "5th International Nitrogen Conference 2010" (M. S. Sachdev, ed.), pp. 235, New Delhi, India.
- Hastie, T., Tibsharini, R., and Friedman, J. (2009). "The elements of statistical learning. Data Mining, Inference, and prediction," Springer Science + Business Media.

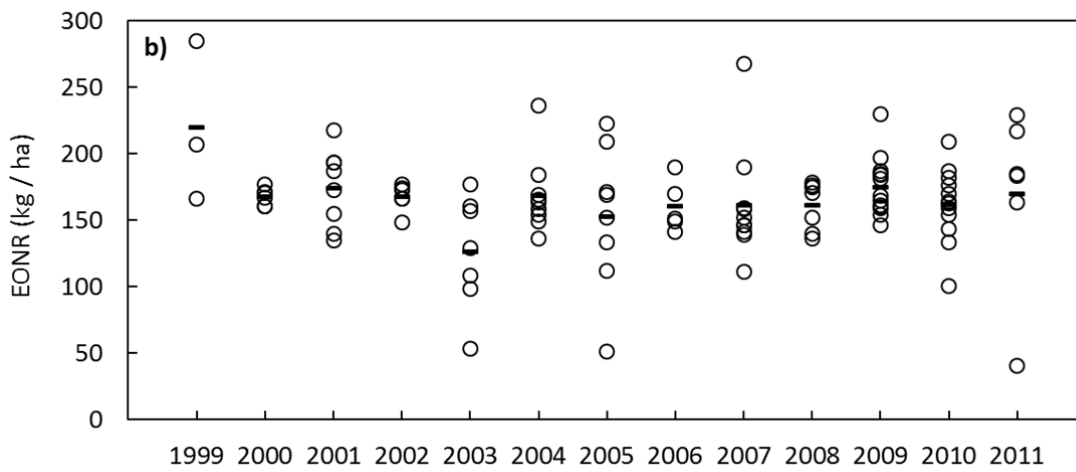
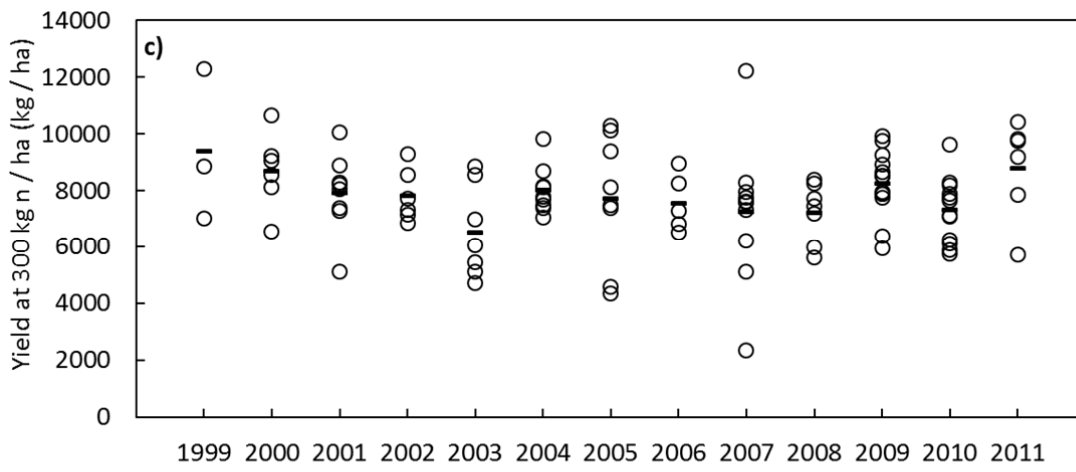
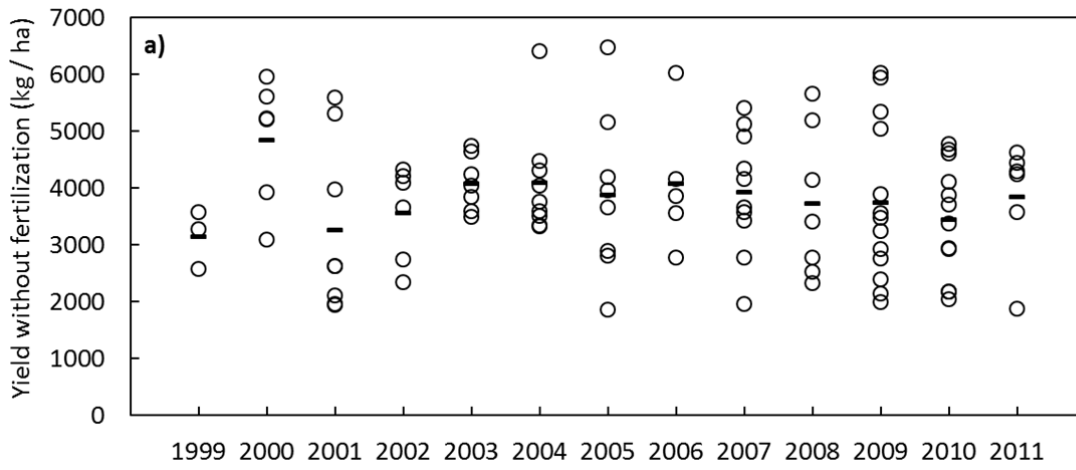
- Lord, E. I., and Mitchell, R. D. J. (1998). Effect of nitrogen inputs to cereals on nitrate leaching from sandy soils. *Soil Use and Management* **14**, 78-83.
- Milborrow, S. (2013). "Notes on the Earth package." Available at: <http://cran.r-project.org/web/packages/earth/vignettes/earth-notes.pdf>. Verified 16 december 2012.
- Overgaard, S. I., Isaksson, T., and Korsaaeth, A. (2013). Prediction of wheat yield and protein using remote sensors on plots-Part I: Assessing near infrared model robustness for year and site variations. *Journal of near Infrared Spectroscopy* **21**, 117-131.
- Raun, W. R., Solie, J. B., and Stone, M. L. (2011). Independence of yield potential and crop nitrogen response. *Precision Agriculture* **12**, 508-518.
- Scharf, P. C., Kitchen, N. R., Sudduth, K. A., and Davis, J. G. (2006). Spatially variable corn yield is a weak predictor of optimal nitrogen rate. *Soil Science Society of America Journal* **70**, 2154-2160.
- Solie, J. B., Monroe, A. D., Raun, W. R., and Stone, M. L. (2012). Generalized Algorithm for Variable-Rate Nitrogen Application in Cereal Grains. *Agronomy Journal* **104**, 378-387.
- R Development Core Team (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.

Table 1. Predictor sets. Y = grain yield. Subscripts denote nitrogen (N) fertilization rate (kg / ha). EONR = economically optimal N rate

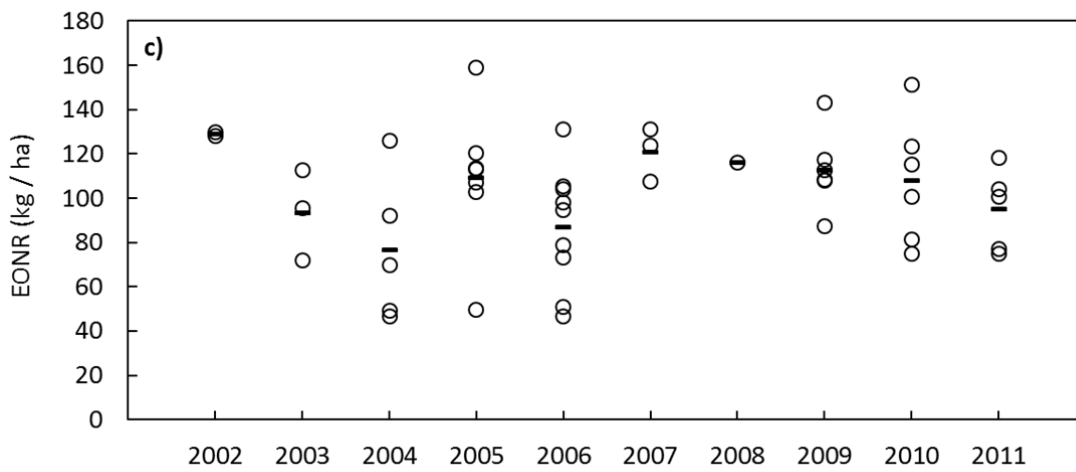
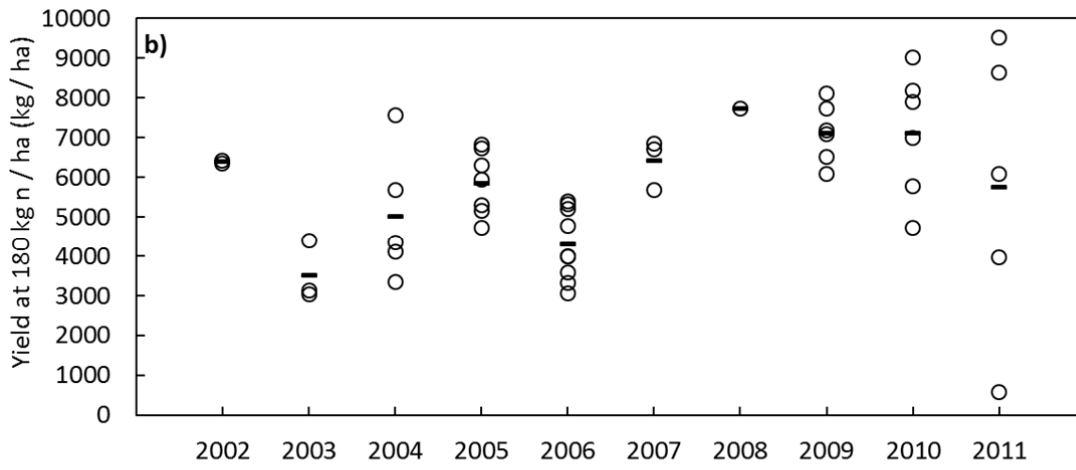
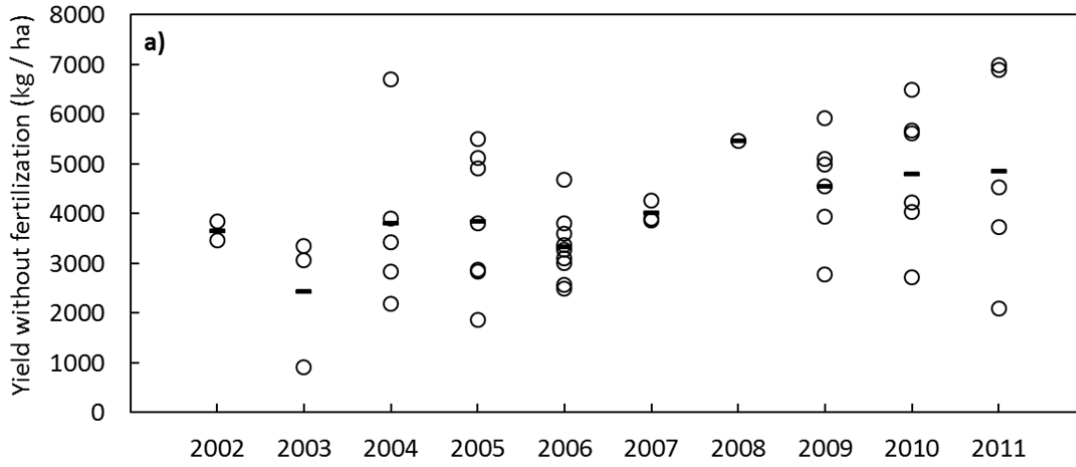
Predictor set	Predictors wheat	Predictors barley
1	$Y_0 + Y_{\text{EONR}}$	$Y_0 + Y_{\text{EONR}}$
2	$Y_0 + Y_{140}$	$Y_0 + Y_{240}$
3	$Y_0 + Y_{180}$	$Y_0 + Y_{260}$
4	$Y_0 + Y_{180}$	$Y_0 + Y_{280}$
5	$Y_0 + Y_{200}$	$Y_0 + Y_{300}$

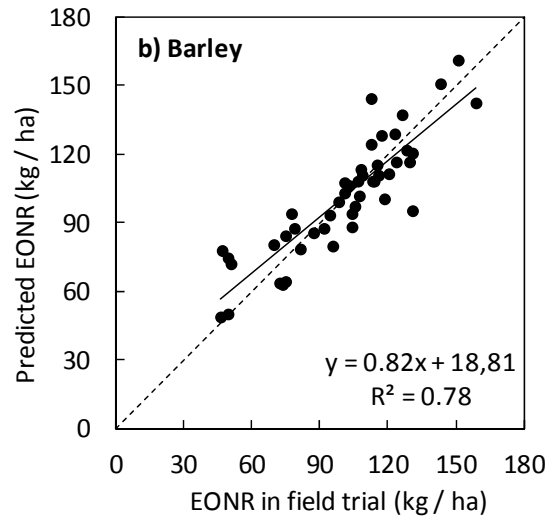
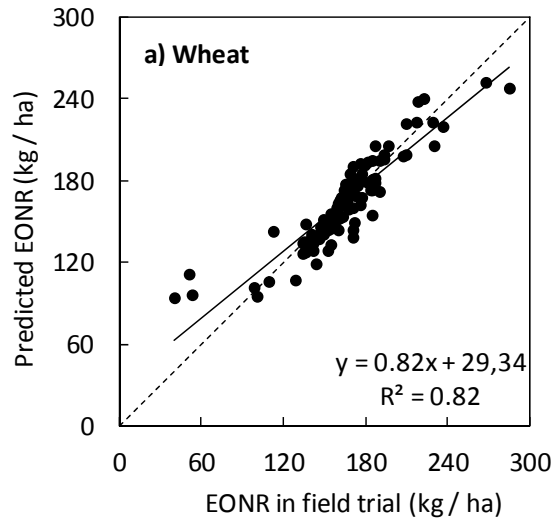
Table 2. Validation of predictions of economically optimal nitrogen rate. Predictor sets are presented in Table 1. MAE = mean absolute error, ME = modelling efficiency. Values within parenthesis are calculated without one outlying prediction.

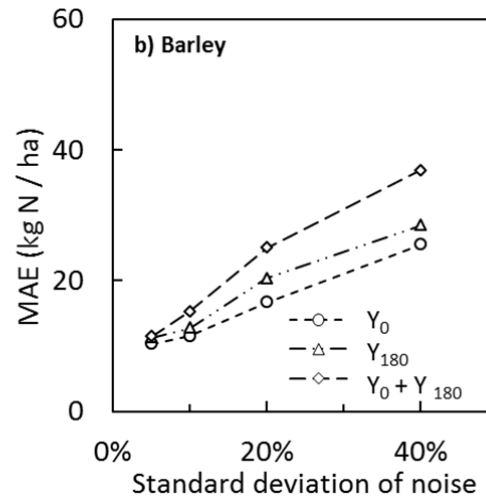
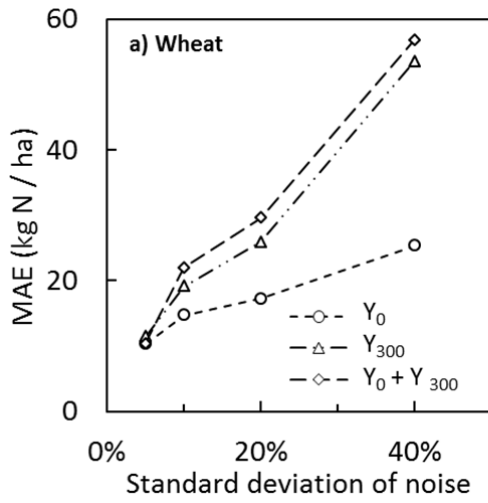
Predictor set	MAE	ME	Predictor set	MAE	ME
<i>Winter wheat</i>			<i>Barley</i>		
Set 1	15.5 (15.5)	0.84 (0.84)	Set 1	13.5	0.79
Set 2	15.4 (15.4)	0.82 (0.82)	Set 2	13.0	0.82
Set 3	13.0 (13.0)	0.85 (0.86)	Set 3	11.3	0.86
Set 4	11.5 (11.5)	0.85 (0.87)	Set 4	9.9	0.88
Set 5	11.0 (11.0)	0.86 (0.90)	Set 5	13.6	0.77



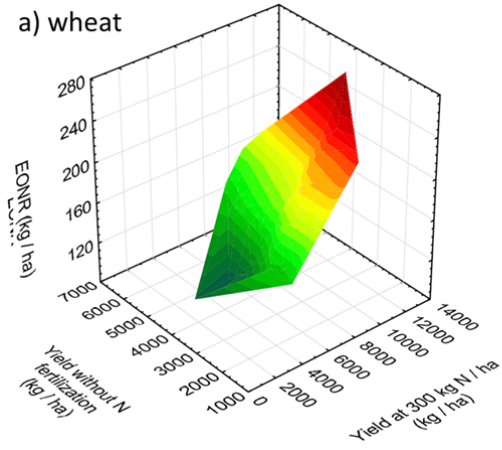
Utkast



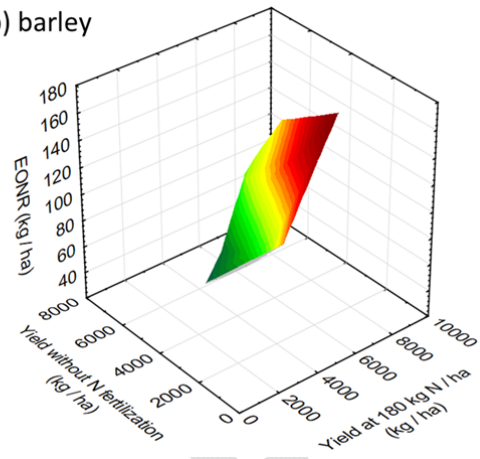




a) wheat



b) barley



Utkast

Figure captions

Figure 1. Trialwise (hollow circles) and yearly mean (line markers) values of a) yield without fertilization, b) yield at 300 kg n / ha and c) economically optimal nitrogen rate (EONR) in wheat.

Figure 1. Trialwise (hollow circles) and yearly mean (line markers) values of a) yield without fertilization, b) yield at 180 kg n / ha and c) economically optimal nitrogen rate (EONR) in barley.

Figure 3. Predicted values of economically optimal nitrogen rate (EONR) plotted against EONR determined in field trials for a) wheat and b) barley. The predictions were made by cross validation of models using predictor set 5 for wheat and predictor set 4 for barley.

Figure 4. Mean absolute error (MAE) in relation to the magnitude of added errors. The added errors were randomly sampled from a Gaussian distribution with mean = 0 and a standard deviation of 5%, 10%, 20% or 40% of the average of the predictor variables. The errors were added to the yield in plots without fertilization (Y_0), the yield in plots where N were not suspected to be limiting to growth (Y_{300} for wheat and Y_{180} for barley) or to both predictor variables.

Figure 5. Final multivariate adaptive regression splines models calibrated by all data.

EONR = economically optimal nitrogen rate.

Utkast

Appendix 1.

List of included field trials. More information on the individual experiments can be obtained from the the Field trial database run by the Field research Unit at the Swedish university of agricultural sciences (www.ffe.slu.se). Each experiment has a unique Adb-number that can be used in the search function of the database.

Adb numbers of wheat experiments:

039884; 039887; 039888; 003A022; 03A023; 03A024; 03A025; 03A026; 03A027; 03B100;
03B104; 03B105; 03B110; 03B111; 03B112; 03B113; 03B115; 03C035; 03C046; 03C047;
03C048; 03C049; 03C050; 03D105; 03D107; 03D126; 03D127; 03D128; 03D130; 03D163;
003E075; 03E078; 03E079; 03E080; 03E0103; 03E0104; 03E0105; 03E0106; 03E0107;
03F008; 03F009; 03F010; 03F011; 03F012; 03F092; 03F094; 03F095; 03G010; 03G014;
03G023; 03G024; 03G027; 03H015; 03H017; 03H018; 03H019; 03H091; 03H092; 03H094;
03H096; 03H097; 03H102; 03K016; 03K020; 03K082; 03K088; 03K089; 03K091; 03K092;
03L020; 03L022; 03L024; 03L109; 03L110; 03L111; 03L112; 03L113; 03L114; 03L116;
03L118; 03L119; 03L120; 03M081; 03M083; 03M084; 03M088; 03M091; 03M092;
03M093; 03M094; 03M095; 03M097; 03M098; 03M099; 03N089; 03N091; 03N092;
03N098; 03N099; 03N101

Adb numbers of barley experiments:

03C040; 03C042; 03D091; 03D092; 03D096; 03E069; 03E071; 03E081; 03E083; 03E085;
03F082; 03F084; 03F085; 03F086; 03F088; 03F089; 03F110; 03G015; 03G016; 03G018;
03G019; 03G021; 03G094; 03G095; 03G096; 03G097; 03H087; 03H088; 03H089; 03K108;
03L126; 03L127; 03L128; 03L129; 03L130; 03L131; 03M116; 03M117; 03M118; 03M119;
03M120; 03M121; 03N082; 03N084; 03N086; 03N087; 03N088.