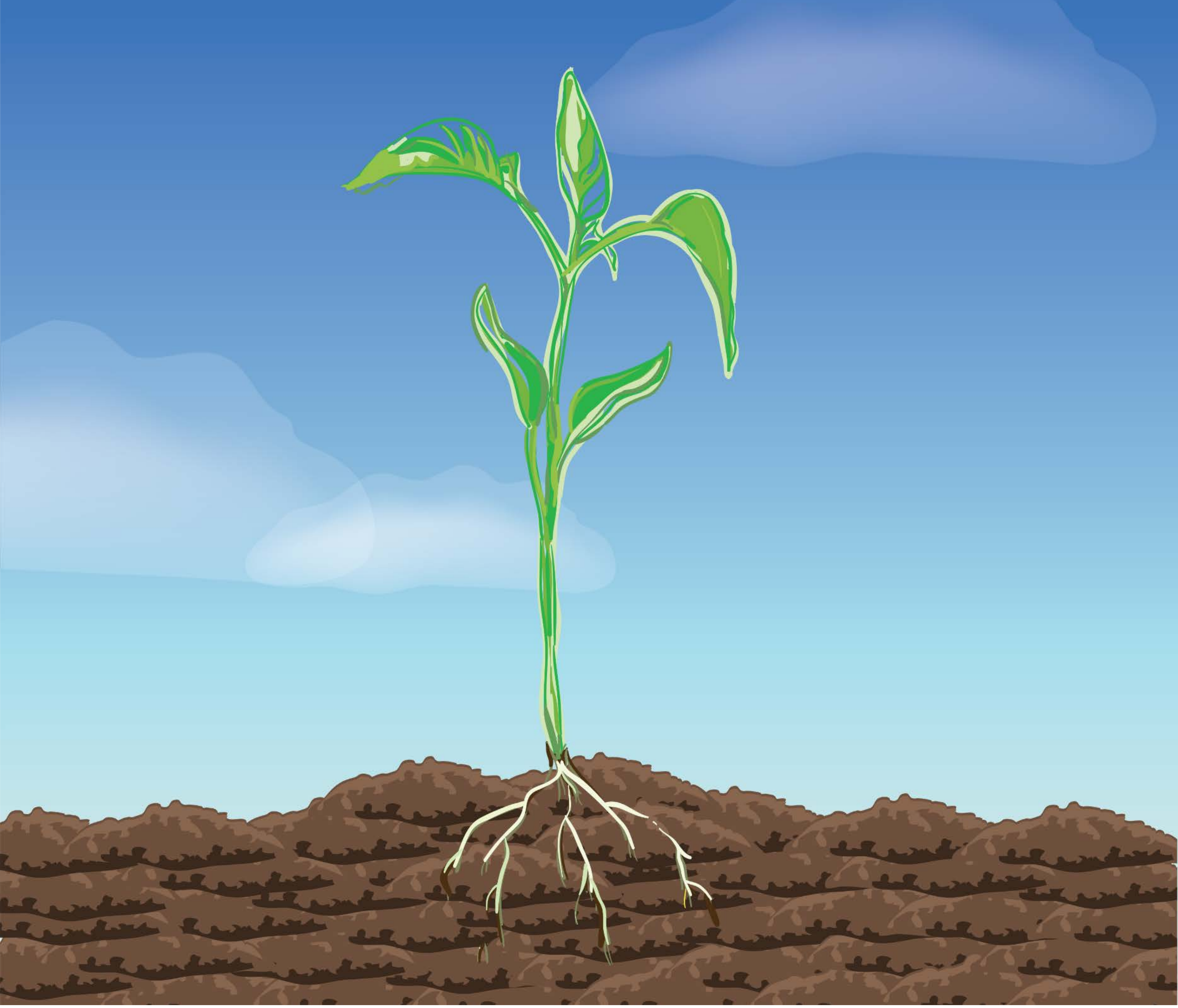
VFRC Report 2015/6

Preliminary evaluation of the feasibility of using geospatial information to refine soil fertility recommendations



Bas Kempen, Pieter Vereijken, Paul Keizer, María Ruipérez González, Prem Bindraban, John Wendt

VFRC

WAGENINGEN UR
For quality of life

ISRIC World Soil Information

IFDC

Preliminary evaluation of the feasibility of using geospatial information to refine soil fertility recommendations

Bas Kempen[1], Pieter Vereijken[2], Paul Keizer[2], María Ruipérez González[1], Prem Bindraban[3], John Wendt[4]

[1] ISRIC – World Soil Information, Wageningen, The Netherlands
Email: bas.kempen@wur.nl

[2] Biometris, Wageningen University and Research Centre, Wageningen, The Netherlands
Email: paul.keizer@wur.nl

[3] Virtual Fertilizer Research Center, Washington, D.C. 20005, USA
Email: pbindraban@vfrc.org

[4] IFDC East and Southern Africa Division, Nairobi, Kenya
Email: jwendt@ifdc.org

VFRC Report 2015/6

Washington, D.C. USA

Executive Director
Virtual Fertilizer Research Center
1901 Pennsylvania Avenue NW
Suite 800
Washington, D.C. 20006
USA
Email: contact@vfrc.org

Additional information on VFRC can be accessed through: http://www.vfrc.org.

Citation:
Kempen, B., Vereijken, P.F.G., Keizer, L.P.C., Ruipérez González, M., Bindraban, P.S., Wendt, J., 2015.
Preliminary evaluation of the feasibility of using geospatial information to refine soil fertility recommendations. VFRC Report 2015/6. Virtual Fertilizer Research Center, Washington, D.C. 67 pp.; 20 tables; 31 fig.; 30 ref.; 4 appendices.

Virtual Fertilizer Research Center

Contents

Abstract

IFDC carries out fertilizer field trials at hundreds of (georeferenced) locations throughout eastern and southern African countries. Yield data analyses from these trials typically focus on average yield responses and economic returns of specific fertilizer treatments with respect to control treatments. However, a deeper analysis of yield data might generate information about the spatial pattern of the response and provide insight in factors that cause differences in yield responses. These insights might be taken into account to make future fertilizer targeting more location or region-specific, as an alternative to blanket recommendations. The information may also be valuable for fertilizer companies and other actors in the chain at regional scale where to sell which fertilizer composition.

The objective of this exploratory study was to gain insight into factors that cause differences in fertilizer response, which might aid in tailoring fertilizer recommendations to local or regional conditions. This is done through (1) a linear mixed model analysis of fertilizer field trial data using spatial soil and climatic data as covariates and (2) an exploratory geostatistical analysis of soil and fertilizer field trial data. The latter included the development of digital soil maps for eleven soil fertility parameters with random forest modeling and kriging. Soil and crop yield data for Burundi were used for these purposes.

The results of the modeling exercises indicate that there is a general lack of correlation between soil type and climatic variables and yield. Based on these results, recommendations may not to be refined for Burundi, indicating that fertilizer formulations used in the omission trials may be valid throughout the study zone. The soil nutrient maps do not appear to be indicative of crop nutrient response, but are indicative of nutrient deficiencies that might be anticipated. We note, however, that eliminations of mistakes in the current datasets and used of more spatially and temporally synchronized covariate data would add value to the analysis. These findings may also prove different over more variable environments such as Rwanda, where soils are considerably more variable.

The added value of geospatial analysis for improving fertilizer recommendations is not immediately evident from these exploratory analyses but the results and the lessons learnt can be used to scope a more robust and thorough strategy for further geospatial data gathering and analyses of fertilizer response trials and for anticipating required nutrient composition of fertilizers.

Introduction

IFDC carries out fertilizer field trials at hundreds of (georeferenced) locations throughout eastern and southern African countries that have been analysed using basic statistical analyses. To advance our insights from these analyses, a consortium of research institutes, including IFDC, VFRC, ISRIC–World Soil Information and the department of Biometrics at Wageningen UR, has engaged in additional analyses.

The objective of the fertilizer trials is to develop better fertilizer recommendations for the respective crops than those currently available. The ultimate parameter of interest in considering what a 'better' fertilizer comprises is an improved return on investment compared to current recommendations. To address this objective two types of trials are carried out: omission trials and best bet trials. The purpose of the omission trials is to determine which nutrients in the fertilizer composition are economically justified. 'Best bet' trials are designed to evaluate current fertilizer recommendations versus improved fertilizer recommendations.

So far, yield data analyses from these trials have focused on average yield responses and economic returns of specific fertilizer treatments with respect to control treatments. However, a deeper inspection of yield data might yield information regarding a spatial response. A spatial analysis might, therefore, provide insight in other factors that cause differences in yield responses observed in the trial data. These insights might be taken into account to make future fertilizer targeting more location- or region-specific, as an alternative to blanket recommendations.

The objective of this exploratory study is to gain insight into factors that cause differences in fertilizer response, which might aid in tailoring fertilizer recommendations to local or regional conditions. This is done through (1) a linear mixed model analysis of fertilizer field trial data using spatial soil and climatic data as covariates and to (2) an exploratory geostatistical analysis of soil and fertilizer field trial data. These studies did not address the economic dimension but analyzed the trials for local fertilizer effects and tested whether these effects are correlated with environmental cofactors. In the first part of this report, yield data from Burundi for beans, potato and maize were used for this purpose. In addition to treatment effects (nutrient omissions, dolomite application), the effects of soil type and historic climatic variables on the observed yield responses were analyzed with linear mixed models. Finally, several recommendations are given to adapt the design of the field trials to better allow for local, geospatial analysis of the field trial data should this be required in the future. In the 'Results' section, general findings are presented. Details of the analyses are given in Appendices 1 to 4. The second part of the report presents the results of an exploratory geostatistical analysis of yield and soil data.

Part 1: Statistical analysis of yield data

This chapter presents the results of an exploratory geospatial analysis of the farm fertilizer yield data. Linear mixed models were used to explore if crop yields were correlated to environmental spatial data.

Materials and Methods

Data

Yield data came from Burundi omission trials (season 2014a, i.e. the first harvest season of 2014) for beans, potato and maize. These trials include a control treatment for each crop, and omissions of a nutrient in comparison to the control. For beans there were 6 treatments: the control treatment 'Tot' (containing N, P, K, S, Zn, B) and five nutrient omissions in each of which one nutrient of the control treatment was left out except P. N was not omitted but rather applied at a reduced rate. For maize, there were 6 treatments: the control treatment 'Tot' and the K, S, Zn, B and Cu omission treatments. For potato, there were 4 treatments: the control 'Tot' and the S, Zn and B omission treatments. At sites with pH<5, dolomite was included. The trials were performed on farms with each treatment allocated to a single plot at a farm. No replicates were feasible at the farms; rather, each farm was considered a replication. For beans and maize, 58 farms were selected, and for potato 44 farms. No randomization was used, but the country was rather well covered by the selected farms, especially for beans and maize. Also, there was no preferential allocation of the trial plots with respect to field gradients. In the allocation of treatments to the plots at a farm, no randomization was imposed, as it was assumed that gradients at each site effectively served to randomize treatments at each site. The spatial distribution of the trial sites is shown in Figure 1.

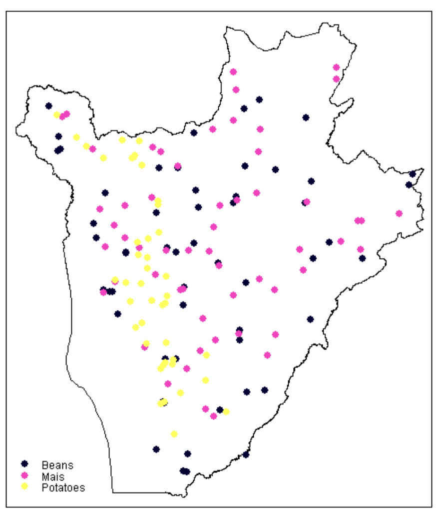


Figure 1. *Locations of the beans, maize, potato omission trial sites for the 2014 season.*

Apart from the yield data, information on the covariates soil type and altitude was obtained from external sources to investigate whether mean crop yield relates to these covariates.

Soil type information was derived from the Soil and Terrain Database of Central Africa (van Engelen et al., 2006). The yield dataset covers six soil types: Acrisols (AC), Cambisols (CM), Ferralsols (FR), Gleysols (GL), Leptosols (LP), and Nitisols (NT). Climate classes were derived from the USGS Isobioclimate map (Rivas-Martinez et al., 2011), which is one of the products of the USGS Africa Ecosystems Mapping database[1]. The yield dataset covers six bioclimate classes: Lower Thermotropical Humid (37), Upper Thermotropical Subhumid (46), Upper Thermotropical Humid (47), Lower Mesotropical Subhumid (56), Lower Mesotropical Humid (57), Upper Mesotropical Humid (67).

Rainfall and temperature data (long term averages) were obtained from the African Soil Information Service (AfSIS). Temperature (K) was derived from the BIO1ALT dataset; precipitation (mm) from the BIO12ALT dataset (Africa Soil Information Service [AfSIS], 2013). Both dataset contain 1950-2000 time series averages and have a spatial resolution of 1000 m. In addition, land surface day temperature was derived from the MYLSTNALT dataset, and land surface night temperature from the MYLSTDALT dataset (Africa Soil Information Service [AfSIS], 2014). Both datasets contain 2002-2014 8-day time series averages.

To investigate whether treatment effects correlate with the soil nutrient status of the farm, the concentration of 18 soil parameters P, K, Ca, Mg, Mn, S, Cu, B, Zn, Na, Fe, Al, pH, cation exchange capacity (CEC), organic matter (OM), electrical conductivity (EC), exchangeable acidity (Hp) and the Calcium-Magnesium ratio (CaMgRt) were measured at each farm.

Statistical analysis

Linear mixed models were used to test the effect of the fertilizer treatments and covariates on mean crop yield. Linear Mixed models are defined by a combination of a fixed part and a random part. The fixed part catches the response to the treatments and the random part describes the structure present in the experiment as variance components. The fixed part or explanatory components has the common statistical assumptions and parametrization for each level.

The random part makes it possible to capture the design of the experiment in various components and enforce their relationships in terms of dependencies. This is done by imposing structure, which reflects the experimental layout of the design, on the variance-covariance matrix of the components while estimating these. The variance components each assume a Normal distribution and independence, representing the physical structure of the experiment as model deviations.

The simplest linear mixed model in this case has ‘Treatment’ as a fixed factor and ‘Farm’ and ‘Treatment-within-Farm’ as random factors. Additional co-factors/covariates can be included in the model. These should be simultaneously recorded or show variation on the scale between farms within regions such as temperature, local rainfall and of course soil composition. The linear mixed models used for analysis here are presented below in a simplified notation.

[1] http://rmgsc.cr.usgs.gov/ecosystems/africa.shtml

Testing treatment effects on mean crop yield: treatment effects have been assessed using the mixed model

$$Yield_{ij} = c + farm_i + \text{treatment}_j + plot_{ij} \quad (1)$$

where c is the intercept. $Yield_{ij}$ is the yield at farm i, i = 1,2,...k, and treatment j, j = 1,2,...l, with k the number of farms in the trial and l the number of treatments. Treatment is modeled as a fixed effect. Crop yield data measured on plots at a particular farm will be correlated. Therefore, in the statistical analysis mixed models with a random *farm* effect and a random *plot* effect were used to describe this correlation in the yield data. Here the *plot* effect is the interaction farm*treatment and is the error term. Farm and plot effects were assumed to be independent normally distributed with mean 0 and variances σ^2_{farm} and σ^2_{plot} respectively. For reasons of clarity the indices have been omitted from the equations below.

Dolomite effects on mean crop yield: effect of dolomite (no: pH≥5, yes: pH<5) has been assessed using an extension of the model in Eq. 1:

$$\text{Yield} = c + \text{dolomite} + farm + \text{treatment} + \text{dolomite*treatment} + plot \quad (2)$$

where dolomite is modeled as a fixed effect. An interaction between dolomite and treatment is also included in the model since the effect of dolomite on mean crop yield can differ between treatments.

Cofactor/covariates effects on mean crop yield: importance of the cofactors soil type and climate class as well as of the calculated cofactors (with low, medium and high categories) for the regional covariates: altitude, rainfall, temperature, night temperature and day temperature have been assessed using the mixed model:

$$\text{Yield} = c + \text{cofactor} + farm + \text{treatment} + \text{cofactor*treatment} + plot \quad (3)$$

Soil nutrient status effect on mean crop yield: importance of soil nutrient status has been assessed by fitting to the scores of the first three components PC1, PC2, PC3 the mixed model:

$$\text{Yield} = c + \text{PC1} + \text{PC2} + \text{PC3} + farm + \text{treatment} + \text{PCx*treatment} + plot \quad (4)$$

The mixed models have been fitted to the data using the restricted maximum likelihood (REML) procedure in the statistical program GenStat (VSN International, 2013). Fixed effects in the models were assessed by approximated *F*-tests for the corresponding Wald statistic under the null hypothesis that the mean yields for the fixed effects are equal. In case the null hypothesis is rejected (p<0.05 for the *F*-statistic; i.e. the mean yields are not equal), Fisher's protected Least Significant Differences (LSD) tests were carried out to find out which of the means are significantly different from each other. These are like independent two-sample Student's *t*-tests with the difference that the variance computed from all the groups of the fixed effect (the pooled variance) is used instead of the variance computed from the two groups that are being compared. The test statistic of Fisher's LSD test is *t*-statistic, which is calculated by computing the pairwise difference of the means and dividing this difference with the standard error of difference (SED).

The SED is a measure precision with which the pairwise difference is estimated from the sample data. If the pairwise difference between means is roughly twice as large as the SED, then the difference is significant. The precision for estimated treatment effects depends on the variation (the pooled variance) between plots σ^2_{plot} only; the SED for pairwise differences between predicted mean yields for the treatment is:

$$SED = \sqrt{\frac{2\sigma^2_{plot}}{n_{farms}}} \tag{5}$$

where n_{farms} is the number of farms. For the precision of pairwise differences between predicted mean crop yields for the classes of the cofactors the size of the variance for variation in yield between farms σ^2_{farm} and of the variation between plots σ^2_{plot} is of importance. The relationship between the SED for comparing the predicted mean yields for class *I* and class *J* and the components of variance is:

$$SED = \sqrt{\left(\frac{1}{n_{farms\ for\ class\ I}} + \frac{1}{n_{farms\ for\ class\ J}}\right)\left(\sigma^2_{farm} + \frac{\sigma^2_{plot}}{n_{treatments}}\right)} \tag{6}$$

where $n_{treatments}$ is the number of treatments.

The SED is used to calculate the LSD as *t*(0.025)*SED, where *t*(0.025) is approximately equal to 2. The LSD is used to calculate the 95% confidence interval of the estimated pairwise differences as the pairwise mean difference ± LSD.

The estimates of the variance components σ^2_{farm} and σ^2_{plot} and the significance of the fixed effects in the models are provided in Appendix 1. In case of a significant *F*-test, two-way tables of predicted mean crop yields and profile plots of predicted mean yields were constructed, as well as diagrams with the test results from comparing pairwise differences between predicted mean yields. These are presented in Appendix 2. Because of the unbalanced design of the field trials, σ^2_{plot} might not be homogeneous when comparing mean yields for co-factor levels. To avoid needless complexity, the median SED was used for pairwise comparison of mean yields. There are, however, modeling methods that can deal with unbalanced experimental designs and non-homogenous error variances (see ‘Question 6’ in the next section).

Results

Question 1. Does omission of a nutrient from the control treatment have a negative effect on treatment yield?

Treatment effects have been assessed using the mixed model with fixed term for treatment effects and random farm and plot effects. Hereafter this model will be called the minimal model. The predicted mean yields with LSD and the estimated variance components σ^2_{farm} and σ^2_{plot} are presented in Table 1. The 95% confidence interval of the estimated pairwise differences can be calculated from the LSDs. For example, the confidence interval of the difference between B omission and the control yields {-0.686 - -0.414}. Since 0 is not included in this interval we can conclude that B omission results in a significantly lower yield.

Table 1. *Predicted mean crop yields (Mt ha^{-1}) with SED and estimated variance components. For the omission treatments, the percentage yield reduction compared to the control is shown between parentheses. For the variance components, the standard error of the components is shown between parentheses.*

Crop	Tot	oN[1]	oK	oS	oZn	oB	oCu	LSD	σ^2_{farm}	σ^2_{plot}
Beans	2.55	2.05(20)	2.10(18)	1.81(29)	2.23(13)	2.00(22)	-	0.136	0.15(0.03)	0.14(0.01)
Maize	4.41	-	3.99(10)	4.27(3)	3.63(18)	3.25(26)	3.15(29)	0.150	0.31(0.06)	0.15(0.01)
Potato	21.0	-	-	20.6(2)	18.4(12)	18.3(13)	-	0.713	2.99(0.80)	2.86(0.36)

[1] "o" stands for "omission". In case of beans, N was not omitted but only reduced.

For beans, omission of S from the control gave the strongest yield reduction: 29%. For maize, the strongest yield reductions occurred when B (26%) and Cu (29%) were omitted. For potato, omission of Zn (12%) and B (13%) gave the strongest reduction in yield. Testing pairwise differences shows significant negative effects on mean predicted crop yield for all nutrient omission treatments except for maize and potato, when S is omitted. The non-significant effect omission of S for potato might be explained by the fact that the potato trial sites (Figure 1) are all located in an area where soils are relatively rich in S (Figure 12).

Question 2. Is soil type and climate class of importance for crop yield and do they influence the size of the negative effect on crop yield when a nutrient is omitted?

To answer this question the minimal model has been extended with a fixed term for effect of soil type and the interaction term between soil type and treatment. The same has been done in assessing an effect of climate class.

This analysis allows statements about marginal effects of soil type and climate class on crop yield, i.e. soil type effects ignoring climate class effects and climate class effects ignoring soil type effects. Since the study is observational, however, marginal effects are not causal due to possible confounding with effects of other factors. In potato, for example, soil type and climate class are associated as is illustrated by the climate class by soil type frequency distribution of counts of farms for potato (Table 2).

Table 2. *Frequency distribution of farms counts for potato for climate class and soil type.*

Farm counts	soil_AC‡	soil_CM	soil_FR	soil_NT	Total
Potato-clim_37†	5	6	1		12
Potato-clim_46	3	1			4
Potato-clim_47	3		2		5
Potato-clim_56		1		1	2
Potato-clim_57		1	10	1	12
Potato-clim_67			9		9
Total	11	9	22	2	44

‡ Soil types: Acrisols (AC), Cambisols (CM), Ferralsols (FR), Nitisols (NT)

† Climate Classes: Lower Thermotropical Humid (37), Upper Thermotropical Subhumid (46), Upper Thermotropical Humid (47), Lower Mesotropical Subhumid (56), Lower Mesotropical Humid (57), Upper Mesotropical Humid (67)

Table 2 shows a number of empty cells with no farms at all and widely varying farm counts in the remaining cells, even for the total marginal counts. Effects of soil type and climate class are confounded and a proper interpretation of marginal effects of climate class and soil types requires additional information from outside the study. In future trials, one should try to avoid confounding of important co-factors by taking a balanced set up of the farms for these factors. 'Balanced' means that all cells (treatment combinations) have the same number of experimental plots (here farms). One should investigate whether this is possible in practice. For example, potato is not grown in all areas of Burundi (see Figure 1), so not all soil-climate class combinations will be practically feasible. As a solution, soil and climate classes could be combined, in a sensible way, to avoid empty cells.

For beans and maize, the mixed model analysis of the yield data did not show significant marginal effects for soil type and climate class nor significant interaction between these factors and treatment. For potato, however, significant interactions of soil type and climate class with treatment were found ($P<0.05$). For the omission treatments oS, oZn and oB, the percentage yield reduction as compared to the control are shown in the following Tables. Profile plots and two-way tables for the significant effects are shown in Appendix 2.

Table 3. *Absolute and relative (between parentheses) potato yield (Mt ha^{-1}) reduction compared to the control for soil type. A significant difference is labeled with an asterisk.*

Soil type	AC	CM	FR	NT
Tot	21.5	21.6	20.4	23.3
oS	0.37 (1.7)	0.35 (1.6)	0.46 (2.3)	0.14 (0.6)
oZn	2.06 (9.6)	4.00 (18.5)*	2.04 (10.0)*	6.55 (28.1)*
oB	2.99 (13.9)*	3.72 (17.2)*	2.00 (9.8)*	4.12 (17.7)*

Table 4. Absolute and relative (between parentheses) potato yield (Mt ha^{-1}) reduction compared to the control for climate classes. A significant difference is labeled with an asterisk.

Climate class	37	46	47	56	57	67
Tot	22.0	20.4	21.2	20.7	19.7	21.8
oS	0.26 (1.2)	0.24 (1.2)	0.51 (2.4)	0.12 (0.6)	0.24 (1.2)	0.87 (4.0)
oZn	3.81 (17.3)*	0.92 (4.5)	3.52 (16.6)*	4.89 (23.6)*	2.68 (13.6)*	0.83 (3.8)
oB	3.67 (16.7)*	1.55 (7.6)	3.54 (16.7)*	2.92 (14.1)	1.50 (7.6)	2.94 (13.5)*

Tests of pairwise differences between predicted mean yields show (Appendix 2):

- a small non-significant reduction in mean potato yield when nutrient S is omitted;
- a substantial significant reduction in mean potato yield when nutrient Zn or B is omitted, except for soil type AC where the negative effect on yield is not significant when Zn is omitted;
- a significant substantial negative effect for climate classes 37, 47, 56 and 57; when Zn is omitted;
- a significant substantial negative effect for climate classes 37, 47 and 67 when B is omitted.

Question 3. Does regional information on altitude, precipitation, and temperature influence mean crop yield, and are the negative effects on mean yield of omission of a nutrient influenced by these covariates?

When the relationship between yield and the covariate is linear, an efficient approach would be to have a mixed model with different linear trend of the covariate for the treatments. However, we chose to use a more robust analysis by constructing (categorical) cofactors from the (continuous) covariates. Here cofactors have been used with three categories: low, medium and high for values below the 35 percentile, between 35 and 65 percentile and values above the 65 percentile. The calculated 35 and 65 percentiles are presented in Appendix 3.

For beans and maize, influence of the cofactors altitude, long-term average rainfall and temperature on mean yield seems to be less important, since no significant interactions with treatment were found. For potato, however, significant interactions with treatment have been found for long-term average rainfall, temperature and night temperature. For the omission treatments oS, oZn and oB, the percentage yield reduction as compared to the control are shown in the Tables 5, 6 and 7. Profile plots and two-way tables for the significant effects are shown in Appendix 2.

Table 5. Absolute and relative (between parentheses) potato yield (Mt ha^{-1}) reduction compared to the control for rainfall classes. A significant difference is labeled with an asterisk.

Rainfall class (mm)	Low:1340	Medium:1443	High: 1567
Tot	19.6	21.9	21.7
oS	0.20 (1.0)	0.72 (3.3)	0.31 (1.4)
oZn	2.49 (12.7)*	2.98 (13.6)*	2.50 (11.5)*
oB	1.35 (6.9)*	2.72 (12.4)*	3.99 (18.4)*

Table 6. *Absolute and relative (between parentheses) potato yield (Mt ha^{-1}) reduction compared to the control for temperature classes. A significant difference is labeled with an asterisk.*

Temperature class (ºC)	Low: 16	Medium:16.75	High: 18.91
Tot	21.0	19.8	22.2
oS	0.72 (3.4)	0.18 (0.9)	0.29 (1.3)
oZn	1.81 (8.6)*	2.91 (14.7)*	3.24 (14.6)*
oB	2.63 (12.5)*	1.70 (8.6)*	3.69 (16.6)*

Table 7. *Absolute and relative (between parentheses) potato yield (Mt ha^{-1}) reduction compared to the control for night temperature classes. A significant difference is labeled with an asterisk.*

Night temperature class (ºC)	Low: 9.62	Medium: 11.15	High: 13.49
Tot	21.6	21.0	20.3
oS	0.54 (2.5)	0.29 (1.4)	0.28 (1.4)
oZn	1.40 (6.5)	3.95 (18.8)*	2.94 (14.5)*
oB	2.92 (13.5)*	2.58 (12.3)*	2.50 (12.3)*

Pairwise comparisons of predicted mean potato yields reveal:

- for all three temperature categories a significant negative effect on mean potato yield when Zn or B are omitted; when S is omitted a negative non-significant effect is found;
- for all three night temperature categories a non-significant negative effect on mean potato yield when S is omitted;
- for all three night temperature categories a significant negative effect on mean yield when B is omitted;
- a significant negative effect when Zn is omitted for the high and medium night temperature categories.
- for all three rainfall categories a significant negative effect on mean potato yield when Zn or B are omitted; when S is omitted a negative not significant effect is found.

For maize and beans, soil type and climatic factors have no significant effect on individual nutrient response, so no regional modification of formulae would be suggested.

For potato, no significant S response is observed, suggesting that S could be omitted from the formulation. Concerning Zn and B, while certain soil classes resulted in non-significant yield losses due to their omission, the percentage yield reduction would suggest that Zn and B should still be included in the formulations. For example, the lowest percentage yield reduction (for Zn omission at low night time temperature class) is 6.5%, which amounts to a yield loss of 1.4 Mt/ha of potato, valued at $280 per ha, assuming a farm-gate price of $0.20/Mt. We can roughly estimate the cost of the 3 kg/ha of Zn applied at $12, which still offers, on average, a substantial return on investment.

Thus it can be concluded that in spite of some significant differences in response for potato, there is little reason to modify fertilizer formulations for any of the 3 crops to suit specific soil classes or climatic variables.

Question 4. Does the use of dolomite at farms with soil pH less than 5 have an effect on mean yield and on the size of the negative effect on yield from omission of a nutrient?

Due to the design an effect of dolomite is confounded with a pH effect, as it was originally assumed that crops on soils at pH 5 or below would respond to dolomite. This assumption is likely not justified, as it is clear that beans and maize which received dolomite not only responded to dolomite, but also out-yielded beans and maize grown on soils of pH>5 which did not receive dolomite. This suggests that soils with pH>5 would also have responded to dolomite. It was also found that the dolomite used contained substantial quantities of Zn and B, which may be the reason that beans (which have moderate Zn and low B requirements) did not respond to Zn and B in the presence of dolomite (see Table 8). Maize, by contrast, still responded to Zn and B in the presence of dolomite. In subsequent experiments (not reported here) maize responded to dolomite, irrespective of pH class (<5; 5-5.5, and >5.5), suggesting that at least for maize, the dolomite response may be due to supply of Mg and/or Ca rather than to the contribution of dolomite to neutralizing soil acidity. Above pH 5.5, exchangeable acidity is insignificant.

The combined dolomite/pH effect has been explored by fitting a mixed model with fixed terms for effects of dolomite, treatment and the interaction between dolomite and treatment. For potato, the use of dolomite seems to be less important since no significant interaction neither dolomite effect has been found. For beans and maize however, a significant interaction with treatment has been found. For beans, testing pairwise differences shows a significant negative effect on mean bean yield for all omission treatments, except for omission of Zn or B when dolomite was supplied, possibly due to the small quantities of Zn and B in the dolomite. Absolute yields of the dolomite trials where significantly higher than for the non-dolomite trials for all omission treatments (Appendix 2 and Table 8), except for the omission of K. For maize, all nutrient omission treatments, except for S, show a negative effect for both dolomite treatments. Percentage yield reductions for the omission treatments are shown in the tables below. Profile plots and two-way tables for the significant effects are shown in Appendix 2.

Table 8. *Absolute and relative (between parentheses) bean yield (Mt ha^{-1}) reduction compared to the control for dolomite application. A significant difference is labeled with an asterisk.*

Beans	No Dolomite, pH>5	Dolomite, pH<5
Tot	2.37	3.01
oN	0.45 (19)	0.63 (21)
oK	0.33 (14)	0.75 (25)
oS	0.69 (29)	0.90 (30)
oZn	0.38 (16)	0.15 (5)
oB	0.71 (30)	0.12 (4)

Table 9. *Absolute and relative (between parentheses) maize yield (Mt ha^{-1}) reduction compared to the control for dolomite application. A significant difference is labeled with an asterisk.*

Maize	No Dolomite, pH>5	Dolomite, pH<5
Tot	4.10	5.21
oK	0.45 (11)	0.36 (7)
oS	0.12 (3)	0.21 (4)
oZn	0.70 (0.17)	0.99 (19)
oB	1.03 (25)	1.51 (29)
oC	1.15 (28)	1.51 (29)

Question 5. Does the soil nutrient state have an effect on the negative effect on yield of omitting a nutrient?

The data from the seventeen soil nutrient parameters Ca, Mg, Mn, S, Cu, B, Zn, Na, Fe, CEC, Al, OM, EC, HP and CaMgRt were available. These parameters, however, are confounded because of dependencies between the parameters, which makes it difficult to assess the effect of one nutrient on the treatments independent of the others. Independent testing of seventeen parameters in a full factorial (orthogonal) design would require a 17-dimensional matrix that contains all possible nutrient concentrations (in levels). Such design is practically impossible not only because of its size (high-dimensional) but also because such matrix would be largely empty (sparse). Nutrient levels are correlated and many combinations will not occur. (Table 2 gives an example of a 2-dimensional matrix for climate and soil classes that already contain empty cells). To reduce independence between parameters and high dimensionality, a principal component was performed to see whether a substantial part of the variation in the data could be described by only a few orthogonal (i.e. independent) variables.

The component analysis has been performed on the standardized parameter data by subtracting the mean value and dividing by its standard deviation. Component analysis shows that the first dimension accounts for at least 30 percent of the variation and the first three dimensions together for 60% of the variation in the soil nutrient parameters. Next, a mixed model analysis with effect of the scores for components PC1, PC2 and PC3 and their interactions with treatment have been performed. Correlations have been calculated to explore pairwise relationships, including the relationships between the soil nutrient level and the omission of that level (Figures 28–30, Appendix 4).

For potato, no significant effects were found for the three components. For beans and maize, only PC1 showed a significant interaction. PC1 can be interpreted as the contrast between acid and non-acid soil conditions. The significant effect is likely to be caused by dolomite application at acidic sites that boosted yields for these sites. Effects of PC2 and PC3 seem not to be important and a final analysis was performed with PC1 effect and interaction of PC1 with treatment. Diagrams with the calculated correlations and the loadings (weights) of the soil parameters in the first component and the percentage of the variation explained in the data are presented in Table 22 in Appendix 4.

The correlation diagrams (Figures 28–30, Appendix 4) show that pairwise correlation between yield response and the soil nutrient parameters is low for beans and potato. These include correlations between soil nutrient level and yield response for the omission of that nutrient. High correlations are found between yield response for maize when K is omitted with pH, Ca, CEC and HP. Percentage yield reduction for the omission treatments are shown in Tables 10 and 11.

Table 10. *Absolute and relative (between parentheses) bean yield (Mt ha^{-1}) reduction compared to the control for the first principal component. A significant difference is labeled with an asterisk.*

PC1 classes	**Low: -2.28**	**Medium: -0.407**	**High: 2.646**
Tot	2.77	2.50	2.36
oN	0.58 (21)*	0.53 (21)*	0.45 (17)*
oK	0.66 (24)*	0.30 (12)*	0.40 (15)*
oS	0.80 (29)*	0.76 (31)*	0.74 (28)*
oZn	0.28 (10)*	0.43 (17)*	0.29 (11)*
oB	0.36 (13)*	0.70 (28)*	0.66 (25)*

Table 11. *Absolute and relative (between parentheses) maize yield (Mt ha^{-1}) reduction compared to the control for the first principal component. A significant difference is labeled with an asterisk.*

PC1 classes	**Low: -2.506**	**Medium: -0.464**	**High: 2.924**
Tot	4.99	4.15	4.06
oK	0.30 (6)*	0.37 (9)*	0.57 (14)*
oS	0.25 (5)*	0.12 (3)	0.04 (1)
oZn	1.00 (20)*	0.75 (18)*	0.57 (14)*
oB	1.45 (29)*	1.08 (26)*	0.93 (23)*
oCu	1.50 (30)*	1.20 (29)*	1.10 (27)*

Pairwise testing of differences in predicted mean yields for beans and maize reveals:

- a significant negative effect on mean yield for beans for all nutrition omission treatments for the low, medium and high category of the first principal component;
- a significant reduction in mean yield for maize for all nutrition omission treatments except for omission of S in the medium and high category of the first component.

Although a PC analysis can be informative, it is of limited use for targeting purposes because the results are difficult to associate with a particular piece of land as a result of transforming physical data to artificial constructs. The results of the PC analysis are likely to be confounded by dolomite application that was only applied to sites with pH<5, and might have caused the significant effect of PC1 (which distinguishes acidic from non-acidic soils) for beans and maize.

For each of the crops, pairwise comparisons showed that nutrient response does not correlate well with the soil nutrient level.

Question 6. What statements do the data from the nutrient omission trial at farms allow?

The objective of the Burundi fertilizer field trials is to develop generalized fertilizer formulae that are substantially more profitable in terms of economic returns than current recommendations. The trials were designed to address this specific objective. This study had a different objective: analyzing the trials for local, between farm, fertilizer effects and testing whether these effects are correlated with environmental cofactors. Such analysis could potentially help to better tailor fertilizer applications to local or regional conditions. For example, a specific nutrient might show on average a positive significant effect on yield response. However, the magnitude of this positive effect can differ between regions in a country because of different local (environmental) conditions. If this is the case, the effects of local conditions on yield response could be taking into account to refine fertilizer recommendations. A study with this objective would have merited a different trial design. Below, recommendations are given to adapt the current design to allow for local analyses of the trial data, in case this becomes relevant for future studies.

- Since single replicate trials at farms have been performed, differences in mean yield between treatments between farms cannot be assessed. By taking replicates within farms one can assess differences between treatments in mean yield between farms. Yet, it is understandable that replications are not feasible under farm conditions, for which appropriate statistical methods should be used or developed; for example, an analysis of variance (ANOVA) generalized linear mixed model with the farm-to-farm random spatial variation (the error) modeled with an autoregressive model of order 1.
- Although marginal effects of soil type, climate class and the other co-factors have been calculated, these effects are not causal because of confounding with other factors as illustrated for soil type and climate class. In future studies using farms, one should aim for a balanced set up around the more important co-factors that ensures that the levels of the co-factors have the same number of experimental plots. Already, an improvement would be if the marginal totals (which are the row and column totals in Table 2) were equal. But even then the situation will be not optimal because these more important factors might still be confounded with other factors varying in an observational study. A balanced design will also help avoid having non-homogeneous error variances.

Yet, the design of the omission trials aimed to suit practical farm reality and to assess the impact of single nutrients against the background of adequate amounts of other nutrients. These conditions may not allow experimental designs to conform conventional statistical concepts. Optimization of representation of co-factors in the on-farm trial population and optimized number of on-farm trials may reduce variability, while complementary statistical methodologies appropriate for such conditions may need to be devised for effective analyses.

Part 2: Geostatistical analysis of soil and yield data

This chapter presents the results of an exploratory geostatistical analysis of the crop yield and soil data. Maps were generated of yield responses to fertilizer treatments and soil fertility parameters through spatial interpolation (kriging) of data measured at the field trial and soil sampling sites. In a subsequent analysis, the soil nutrient maps were correlated to yield responses to assess whether these could explain some of the observed spatial variation. Soil nutrient maps are made for Rwanda and Burundi, while yield data from on-farm trials in Burundi were sufficiently dense for bean and maize yields to allow production of their yield maps.

Materials and Methods

Spatial interpolation methods

To create a map, we need to predict target variable *y* at unvisited locations $\mathbf{s}_0$. The target variable can for example be a soil nutrient or crop yield. Most spatial interpolation methods assume that prediction $\hat{y}(\mathbf{s}_0)$ is a weighted linear combination of *n* observations $y(\mathbf{s}_i)$, for $i = 1...n$, in the vicinity of $\mathbf{s}_0$:

$$\hat{y}(\mathbf{s}_0) = \sum_{i=1}^{n} \lambda_i \, y(\mathbf{s}_i) \tag{7}$$

where λ_i is the weight assigned to observation point $\mathbf{s}_i$. See Knotters et al. (2010) for a comprehensive overview on spatial interpolation methods or (Isaaks and Srivastava, 1989) or (Webster and Oliver, 2007) for a practical guide. The prediction locations $\mathbf{s}_0$ are typically the nodes of a regular grid. Here a grid with a spatial resolution of 250 m was used. Hence, the Burundi soil maps are made up from over 420,000 pixels (grid cells), and the Rwanda maps over 360,000 pixels.

The main difference between the various spatial interpolation methods is in the way the weights are determined. In nearest neighbor interpolation, for example, a weight of one is assigned to the observation point $\mathbf{s}_i$ nearest to prediction point $\mathbf{s}_0$, all other observation points get weight zero. A disadvantage of this method is that not all available information is used. Local sample mean interpolation is an improvement in this respect. It takes more information into account by assigning equal weights to all observation points in the vicinity of $\mathbf{s}_0$. However, it seems reasonable to assign larger weights to nearby observations, and smaller weights to observations some distance away from the prediction point. This is what inverse distance (ID) interpolation methods do. Each weight is inversely proportional to the distance between the prediction point and the observation point. A further improvement is to relate the weights to a model of spatial continuity that is computed from the available data. The strength of the spatial continuity, or spatial dependence, is quantified with the semivariogram. If spatial dependence is weak, only observations close to the prediction location receive weight. If spatial dependence is strong, then also points further away receive some weight, though these are typically smaller than the weights of observations closer to the prediction location.

Kriging interpolation

This brings us to kriging interpolation methods. Each kriging weight is related to the distance between prediction and observation point. Contrary to ID interpolation, kriging weights are determined from the data. In addition, kriging weights down-weigh observations that are spatially clustered. In this way, predictions are less affected by redundant information. Another important feature of kriging is that it also provides a local quality measure for each prediction. A graphical example of kriging is given in Figure 2.

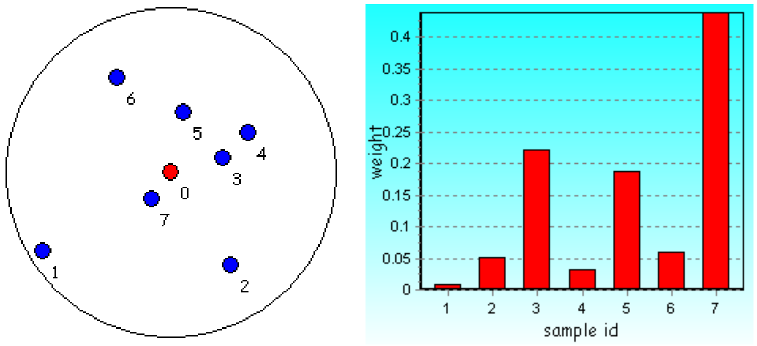


Figure 2. *Kriging example.*

The plot on the left shows the locations of seven hypothetical observation sites (points 1-7) and one prediction site (point 0). The seven observation points are used to predict at the prediction site. The plot on the right shows the kriging weights that are associated to the seven observation points given a specific semivariogram. For this point configuration the inverse distance (ID) weights (computed as 1/d2) would be 0.022 (1), 0.047 (2), 0.172 (3), 0.076 (4), 0.146 (5), 0.051 (6), 0.485 (7). Note that the kriging weights for points 3 and 5 are larger than the ID weights and that the kriging weight of point 4 is much lower than the ID weight. Since this point 4 is very close to point 3, kriging down-weighs the influence of this point on the prediction point.

The kriging weights are obtained by solving a linear system of equations, also known as the kriging system. The kriging weights result in unbiased predictions. This means that the systematic error is zero: there is no systematic over- or under-prediction (i.e. bias) of the property of interest. Furthermore, kriging ensures that the uncertainty associated to the predicted value is as small as possible. This means that among all linear predictors (examples of these are provided above), the kriging predictor is the most optimal one.

The simplest form of kriging in which only the observations are used is called ‘ordinary kriging’. Soil property maps of Burundi reported by Nduwimana et al. (2013) were generated from the Burundi soil sample dataset using ordinary kriging. Ordinary kriging can produce accurate maps, but it is based on the observed data only. We know, however, that soil properties and their spatial distribution are controlled by environmental and anthropogenic factors such as parent material, rainfall, land use, elevation, cropping system, agronomic management, etc. One can, therefore, expect that predictions will improve when one exploits the relationship between soil and environmental conditions in a prediction model. This is confirmed by the scientific literature (Hengl et al., 2004). Environmental conditions can be represented with MODIS satellite imagery, land cover or geological maps, and terrain property maps that can be computed from digital elevation models. But also legacy soil maps are often useful covariates. A commonly used digital soil mapping technique that includes environmental covariates in the prediction model is called ‘regression-kriging’ (RK) (Hengl et al., 2007; Hengl et al., 2004; Odeh et al., 1995).

In RK, soil properties are correlated to spatially exhaustive environmental data (e.g. MODIS imagery, soil or land cover maps, DEM) in a linear regression model. The model residuals, the difference between the measured value and the value predicted by the linear regression model at the measurement site, are interpolated by kriging. The RK model has the following generic form:

$$Y(\mathbf{s}) = \mathbf{d}(\mathbf{s})^{\mathrm{T}}\beta + \epsilon(\mathbf{s}) \tag{8}$$

where $\mathbf{d}(\mathbf{s})$ is a vector of environmental variables at spatial location $\mathbf{s}$, β is a vector of regression coefficients, and ϵ is the model residual. The residual is assumed to be spatially dependent, i.e. residuals at locations close together tend to be more similar than residuals locations that are further apart. Furthermore, we assume that the residual is a normally distributed random variable which has an expected value of zero everywhere, a variance equal to σ^2, and correlation function $\rho(h)$. The correlation function defines the strength of the correlation between two sampling locations as a function of the distance between the sampling locations. The degree of spatial dependence of the residuals is described by the semivariogram. The semivariogram is a continuous function that models the strength of spatial dependence between sampling locations, expressed by the semivariance, as a function of the distance between locations (Isaaks and Srivastava, 1989; Webster and Oliver, 2007). The semivariogram is used in the kriging system to compute the kriging weights λ_i (Eq. 7) that are associated to the sampling locations.

Implementing regression-kriging is fairly straightforward. It follows the following framework:

1. Extract the covariate information at the observation locations through a GIS overlay operation.
2. Calibrate between the target property and covariate data by fitting a linear regression model.
3. Apply the linear regression model to the prediction locations (typically the nodes of a regular grid) to predict the target property.
4. Predict the target property at each observation location by applying the linear regression model.
5. Compute the residual values at the observation locations by subtracting the predicted value from the observed value.
6. Model the spatial correlation structure of the regression residuals with a semivariogram.
7. Predict the regression residuals at prediction locations with ordinary kriging.
8. Sum the predicted property value and kriged residual value to obtain the final prediction.

Tree-based methods and random forests

Regression-kriging assumes a linear relationship between the target property and covariates. In case of soil properties, relationships between target property and covariates are often not linear. For this reason, non-linear models in general and tree models in particular are becoming increasingly popular tools in digital soil mapping to model the relationship between soil property and covariates (Grimm et al., 2008; Häring et al., 2012; Nauman and Thompson, 2014). Besides their ability to model non-linear relationships, tree models have several additional advantages to linear regression models. Most importantly, tree models do not pose any parametric assumptions on the data, i.e. the data does not need to follow a normal distribution, and they can more easily accommodate complex non-linear interactions between covariates (Smith et al., 2013; Strobl et al., 2009). However, their supposed superiority over linear regression models for digital soil mapping has not been proven yet in the scientific literature.

Tree-based methods are part of the large family of machine learning approaches. Machine learning employs computer science and statistics to uncover patterns and relationships in datasets. A tree model partitions the training dataset, for example a soil sampling dataset, recursively into increasingly homogeneous subsets (Strobl et al., 2009) by means of binary splitting (Figure 3). The covariates are used as partitioning variables, and each split is chosen in such a way that it maximizes the reduction of an accuracy measure, such as the sum of squares in case of a continuous target property or the Shannon entropy, Gini index or misclassification error in case of a categorical target property (Hastie et al., 2008). At the terminal nodes of the tree (the leaves), a simple model is fitted to the observations that are contained in the nodes. The model can be a constant, for instance the mean of the values. There are two types of trees: regression trees and classification trees. Regression trees are used to model continuous data (e.g. soil nutrients). Classification trees are used for categorical data (e.g. soil class).

Figure 3. *Example of a tree model. The numbers indicate the predictions at the leaves of the tree.*

One of the most popular tree models in digital soil mapping are random forest (RF) models. A random forest consists of an ensemble (a forest) of classification or regression trees, depending on the type of data to be modeled (Breiman, 2001; Strobl et al., 2009). Random forests combine bootstrap aggregation (bagging) and random covariate selection for partitioning, to grow a forest of trees. This means that for each tree in the forest a bootstrap sample is drawn from the calibration dataset randomly - a bootstrap sample is a subset of the full dataset drawn with replacement. Typically, the bootstrap sample contains two-thirds of the dataset. A random-forest tree is grown to the bootstrap sample. This is done as follows (Hastie et al., 2008):

1. Select *m* covariates at random from the full set of *p* covariates. *m* is often taken as the square root of *p* or *p*/3.

2. Select the best covariate/split-point among the *m*. This is the covariate/split-point that results in the largest reduction in error.

3. Split the node into two daughter nodes.

These steps are repeated recursively until a minimum node size is reached. The minimum node size is the minimum number of observations that must be included in a terminal node. These observations are used for the terminal node prediction (see above). Bootstrapping and growing trees is repeated many times, a random forest typically contains 500 to 1,000 trees. Each tree is used to make a prediction at the data points or at new points (for instance the nodes of a prediction grid). The predictions of the individual trees are then aggregated in a single prediction: the random-forests predictor. For regression trees, this is the average of the predictions of the individual trees. For classification trees, the prediction is based on the majority vote of the individual tree predictions (Breiman, 2001; Hastie et al., 2008).

The samples not included in the bootstrap sample are the called the 'out-of-bag' (OOB) samples. The OOB samples are used to assess the prediction accuracy of the tree by computing the OOB error. This is done as follows. Once a random-forest tree is grown to the bootstrap sample, it is used to predict the target variable for all OOB samples of that tree. This is done for every tree in the forest. Because bootstrapping the full dataset is done randomly, each sample in will be in the OOB set approximately one-third times the number of trees in the forest. The OOB predictor for a sample is the average of the individual OOB predictions for that specific sample. The OOB error is calculated by subtracting the observed value from the OOB predictor. From the OOB error typical validation can be calculated such as the mean error (a measure of bias) and the root mean squared error (a measure of accuracy) (Brus et al., 2011).

Random forests are not easy to interpret. There is no such thing as an average tree that can be visualized for interpretation for example to assess the relevance of the covariates. Despite this drawback, there are measures, the so-called 'variance importance measures' that can be used to assess the relevance of each covariate over all trees in the forest (Strobl et al., 2009). The most advance variance importance measure is the 'permutation accuracy importance'. Strobl et al. (2009) provide the following rationale for this measure. By randomly permutating the values of a covariate its originial association with the target variable is broken. Hence, if a permuted covariate, together with the remaining unpermuted covariates, is used to predict the target variable, then this will result in a decrease of prediction accuracy. Covariates that have the strongest association with the target variable will show the largest decrease in prediction accuracy. Thus, a reasonable measure for variable importance is the prediction accuracy before and after permuting a variable, averaged over all trees. Usually, the OOB error is used to compute the permutation importance.

Random forest modeling can be extended with a kriging step, like in regression-kriging following the same framework. The random forest model is used to predict the target property at the observation sites. From the predicted and observed value the residual is computed. A variogram is fitted to the residual and used for spatial interpolation of the residual values to the prediction locations. The random forest model is applied to the covariate data layers to obtain predictions at the prediction locations. Finally, the random forest predictions and the predicted residuals are summed.

Uncertainty assessment

Kriging takes a stochastic approach to quantifying and mapping spatial variation. It assumes that the sample data are not deterministic values but the outcomes of a stochastic (random) process. This means that the property of interest is assumed to be a random variable. A random variable is a variable which values are randomly generated according to some probabilistic mechanism. (Like the outcome of throwing a six-sided die. The outcome has six possible values, each having equal probability: 1/6). Thus, the sample data are regarded as realizations from an underlying probability distribution (Isaaks and Srivastava, 1989; Webster and Oliver, 2007). In kriging, this distribution is assumed to be a normal (Gaussian) distribution. The normal distribution has two parameters: the mean (or expected value) and variance.

The stochastic approach to spatial variation makes it possible to quantify uncertainty that is associated to the predicted value. This is a great advantage compared to the other spatial interpolation methods such as ID interpolation. The uncertainty of the kriging prediction is quantified by the kriging variance (also referred to as the prediction error variance). The target variable, here a soil nutrition parameter, at a prediction location is assumed to follow a normal distribution. The expected value of the distribution is the predicted value; the variance of the distribution is the kriging variance. This means that kriging does not provide a single estimate of the property of interest at each prediction location, but a probability distribution of values of which the predicted value is the most likely one.

Kriging provides quantified information about the uncertainty associated to the kriging predictions. This information can be used for several purposes. The most common purpose is to derive the prediction interval (PI). The PI is an estimate of the interval in which future observations will fall with a certain probability. Here the 90% PI was used that is calculated from the kriging predictions following classical statistical theory:

$$\begin{aligned} lower &= \hat{y}\,(\mathbf{s}_0) - 1.645 * \sqrt{var\big(\hat{y}\,(\mathbf{s}_0)\big)} \\ upper &= \hat{y}\,(\mathbf{s}_0) + 1.645 * \sqrt{var\big(\hat{y}\,(\mathbf{s}_0)\big)} \end{aligned} \tag{8}$$

where $\hat{y}\,(\mathbf{s}_0)$ is the predicted value at prediction location $\mathbf{s}_0$, and $\mathrm{var}\big(\hat{y}\,(\mathbf{s}_0)\big)$ is the kriging variance at location $\mathbf{s}_0$.

Quantified information of the prediction uncertainty can also be used to estimate the probability that a specific threshold or critical level will be exceeded, for instance a soil nutrient level for a specific crop. This is done through geostatistical Monte Carlo simulation. During simulation, prediction locations are visited in random order. At each location, a value is drawn from the kriging distribution, while taking into account values drawn at other prediction locations through the spatial correlation structure. When all prediction locations are visited, then this gives a map of a possible realization of reality.

Burundi and Rwanda soil and covariate data

We used IFDC soil datasets of Rwanda and Burundi, which are to our best knowledge the most comprehensive countrywide datasets with soil nutrient data. The Rwanda dataset contains 999 georeferenced observations, the Burundi dataset 1,007. Both datasets have good spatial coverage, though the Burundi dataset is somewhat more clustered (Figure 4). The soil samples were collected and analyzed by Crop Nutrition Laboratory Services. The following soil properties were measured: pH, P (ppm), K (ppm), Ca (ppm), Mg (ppm), Mn (ppm), S (ppm), Cu (ppm), B (ppm), Zn (ppm), Na (ppm), Fe (ppm), CEC (meq $100g^{-1}$), Al (ppm), Si (ppm), organic matter (OM) (%), electrical conductivity (EC) (µS cm^{-1}), H (%), exchangeable acidity (meq $100g^{-1}$), Acid saturation (%). Summary statistics are presented in Table 12. This table shows that concentrations of all primary, secondary and micronutrients except for S are somewhat larger for the soils in Rwanda than for Burundi.

A total of 134 covariate layers were acquired from six sources: the Africa Soil Information Service (AfSIS)[2], ISRIC WorldGrids[3], USGS Africa Ecosystems Mapping database[4], soil and terrain database of Central Africa (SOTERCAF) (van Engelen et al., 2006), soil and terrain database for northeastern Africa (SOTERNEA) (FAO, 1997), Multipurpose Africover Database for the Environmental Resources (MADE)[5]. Maps of terrain (elevation, specific catchment area, topographic wetness index), climate (precipitation, day, night and long-term average temperature), soil moisture (monthly and long-term averages), land cover, vegetation indices (enhanced, normalized difference, leaf are), fraction of photosynthetically active radiation, primary production (gross, net, yearly and long-term averages), albedo (black and white sky, near-infrared, shortwave, visible) and spectral reflectance (red, blue, near-infrared, mid-infrared for January, April, August, long-term average) were obtained from AfSIS. Additionally, from the day and night temperate layers, a temperature difference layer was computed. From the spectral reflectance images, reflectance ratios were computed.

[2]http://africasoils.net/services/data/remote-sensing/

[3]http://www.worldgrids.org

[4]http://rmgsc.cr.usgs.gov/ecosystems/africa.shtml

[5] http://www.glcn.org/activities/africover_en.jsp

Table 12. *Summary statistics of soil fertility properties.*

Burundi						
	pH	OM (%)	P (ppm)	K (ppm)	Ca (ppm)	Mg (ppm)
Min	3.8	0	0.2	19	43	11
Q1	5.0	3.7	2.6	73	264	59
Median	5.3	4.5	4.2	115	496	118
Mean	5.9	4.6	7.5	150	752	172
Q3	5.8	5.2	6.9	191	995	218
Max	8.1	9.4	137.0	708	11200	2520

Burundi					
	S (ppm)	Zn (ppm)	B (ppm)	Cu (ppm)	ECEC (meq $100g^{-1}$)
Min	3.2	0.28	0.20	0.20	0.69
Q1	8.6	0.83	0.07	0.90	4.31
Median	11.8	1.19	0.12	1.44	5.59
Mean	13.4	1.78	0.15	1.90	6.99
Q3	16.0	1.94	0.19	2.33	7.92
Max	164.0	26.60	1.49	14.10	79.28

Rwanda						
	pH	OM (%)	P (ppm)	K (ppm)	Ca (ppm)	Mg (ppm)
Min	3.2	0.4	0.2	14	51	17
Q1	4.9	3.4	5.4	83	398	94
Median	5.6	4.1	9.3	153	792	173
Mean	5.6	4.3	29.6	226	1192	234
Q3	6.2	5.0	22.1	267	1480	310
Max	8.5	8.4	950.0	2960	9800	2050

Rwanda					
	S (ppm)	Zn (ppm)	B (ppm)	Cu (ppm)	ECEC (meq $100g^{-1}$)
Min	0.5	0.31	0.02	0.20	2.05
Q1	5.6	1.06	0.12	1.05	5.26
Median	10.4	1.83	0.22	1.87	7.13
Mean	13.1	3.65	0.35	2.11	9.72
Q3	16.7	3.75	0.44	2.69	11.13
Max	157.0	115.00	0.64	17.60	65.27

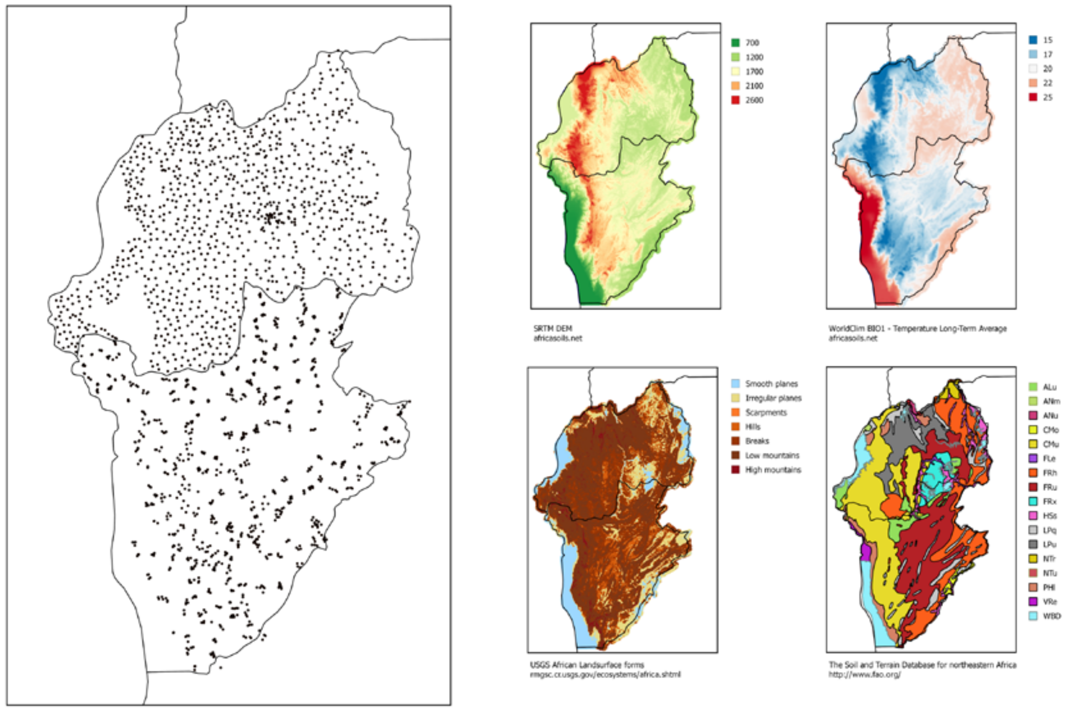


Figure 4. *Distribution of the soil sampling sites across Rwanda and Burundi (left panel) and four examples of covariate data layers that were used to fit the random forest models (right panel).*

The AfSIS digital elevation model was used to derive e-SOTER6 terrain attributes (elevation, relief intensity and slope classes) and units (Dobos et al., 2005; Dobos et al., 2010). From WorldGrids, maps of land cover (MERIS, GLC2000, ESA GlobCover), landform (SCALA), slope percentage, long-term averages of MODIS time series data (enhanced vegetation index, day and night temperature) were obtained. From the USGS Africa Ecosystem Mapping database, maps of isobioclimates, ecosystems, landsurface forms, surficial lithology, topographic position were obtained. From MADE maps of habitats and lithology were obtained. From SOTERNEA, a map of the dominant soil class was derived and from SOTERCAF maps of dominant soil class, landform and lithology. All datasets, soil and covariate, were brought to a common geographical projection before modeling (UTM Zone 36 south). Figure 4 shows four examples of covariate maps used for digital soil mapping.

Modeling

Random forest models were used to map the following soil fertility properties: pH, P, K, Ca, Mg, S, Cu, B, Zn, OM and effective CEC. Effective CEC was computed as Ca + Mg + Na + K + exchangeable acidity (all in meq $100g^{-1}$; for this purpose Ca, Mg, Na and K measurements were converted from ppm). Separate RF models were fitted for the Burundi and Rwanda data. Semivariogram models were fitted to the model residuals. If the semivariograms gave evidence of spatial correlation, then the residuals were kriged to the nodes of the prediction grids and added to the random forest

---

[6] www.esoter.net

predictions. All properties except pH and OM were transformed to natural logarithms to normalize skewed data distributions. The predictions were back-transformed to the original scale after the kriging step.

Random forest models were fitted with the randomForest package (Liaw, 2014) for the statistical software R (R Development Core Team, 2015). Each forest was an ensemble of 1,000 trees. Fifteen covariates were randomly selected as candidates for each split, and the minimum size of the terminal nodes was set to ten. Kriging was performed with the gstat package (Pebesma, 2004) for R. Model accuracy was assessed through 10-fold cross-validation (Brus et al., 2009). This was done with the caret package (Kuhn, 2015) for R in case only a random forest model was fitted, and with the gstat package in case the random forest predictions were complemented with a kriging step.

Cross-validation

In *n*-fold cross-validation, the dataset is split into *n* sets. A model is calibrated with the observations contained in *n-1* set, and used to predict at the validation set that is left out of the calibration. This is repeated *n* times, in which each time a different validation set is used. In this way, a cross-validation prediction is made at each observation site. The observed value is subtracted from this prediction giving the prediction error. From the prediction error, several accuracy measures can be calculated: the mean error (ME; indication of prediction bias), mean absolute error (MAE) and the root mean square error (RMSE) and root median squared error (RMedSE). The latter might give a better indication of the 'average' error in case of a skewed squared error distribution caused by presence of one or more extremely large errors. These inflate the estimate of the mean squared error. The RMedSE gives the midpoint of the error distribution: 50% of the errors are smaller and 50% are larger.

Spatial analysis of crop yield data

Like the soil data, the measured yield responses to fertilizer applications are also georeferenced. This means that we can assess if there is spatial structure in the response data, and if there is, create response maps through kriging interpolation. We have explored this using beans yield data from season 2014a and maize yield data from season 2015a. For this purpose, data from the comparison (or 'best bet') trials were used. The omission trial data cannot be combined with best-bet trial data because there is not a no-fertilizer control. Reliable estimation of the semivariogram requires approximately 100 data points (Webster and Oliver, 2007). For maize, yield responses (control vs. best-bet treatment) were available for 119 comparison trial sites and for beans for 330 trial sites. Figure 5 shows the distribution of the trial sites across Burundi. The comparison trials compare the responses of three treatments: the control (no fertilizer), current practice (NPK), best bet practice (NPK + soil micronutrients). For the maize season 2015a the three treatments have been applied with and without dolomite, giving six treatments in total. Figures 6 and 7 depict the observed yield responses for the fertilizer treatments. Figures 8 and 9 show the differences in yield response between the treatments. Semivariograms were fitted to the yield response data for each treatment, and for the differences between treatments. The absolute yield responses, as well as yield differences, were interpolated to the nodes of a grid with 250 m spatial resolution with ordinary kriging. Note that no environmental covariates were used for interpolation. There was no distinction made between the different maize and bean varieties that were used in the trials.

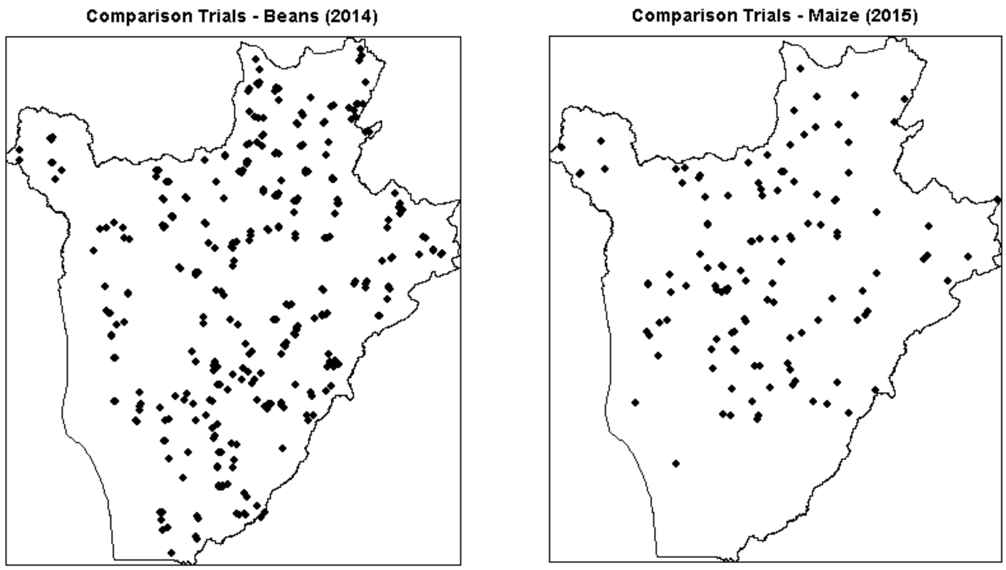


Figure 5. *Locations of the fertilizer comparison trials.*

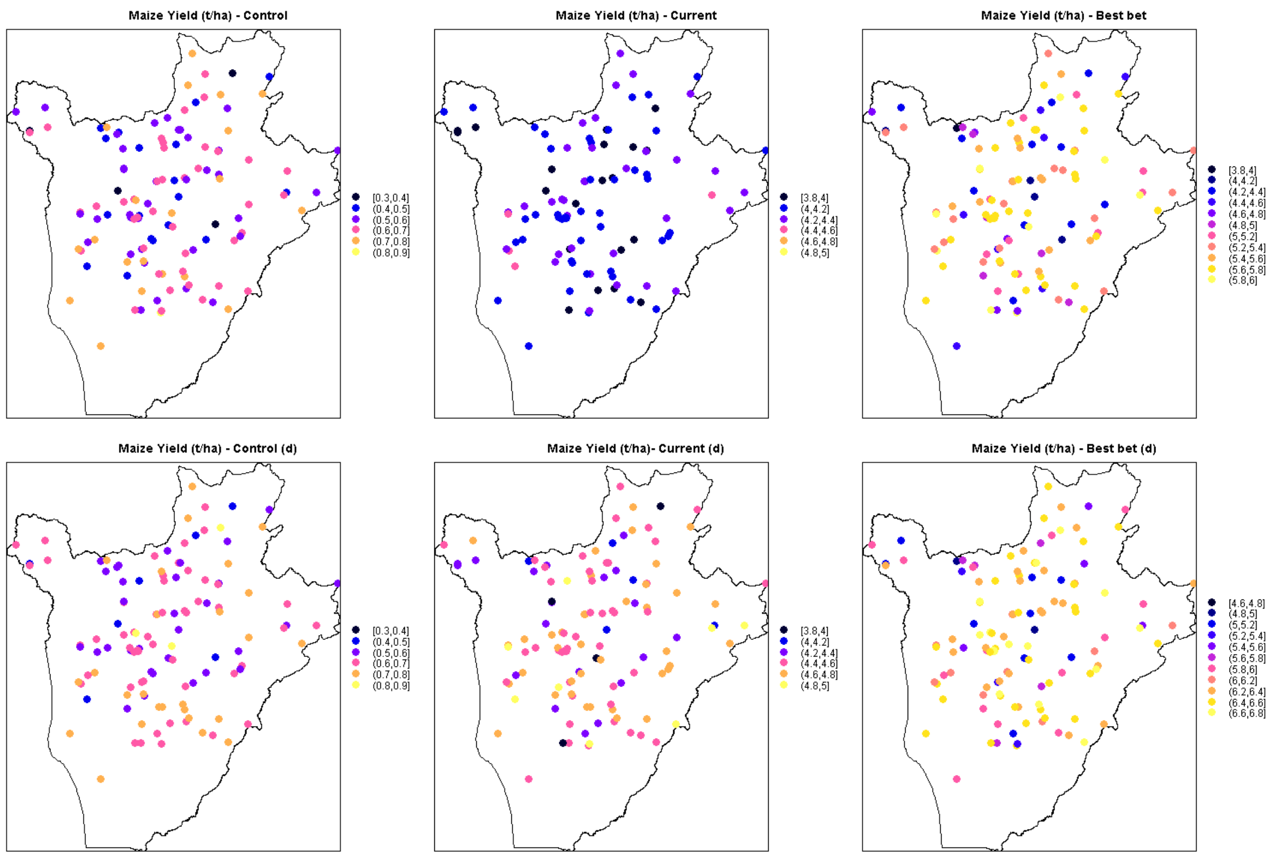


Figure 6. *Yield response to the fertilizer treatments for maize.*

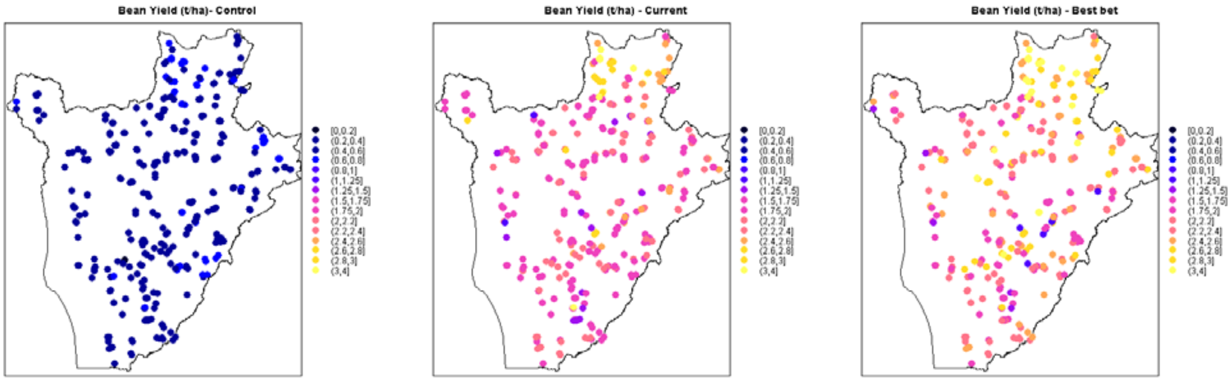


Figure 7. *Yield response to the fertilizer treatments for beans.*

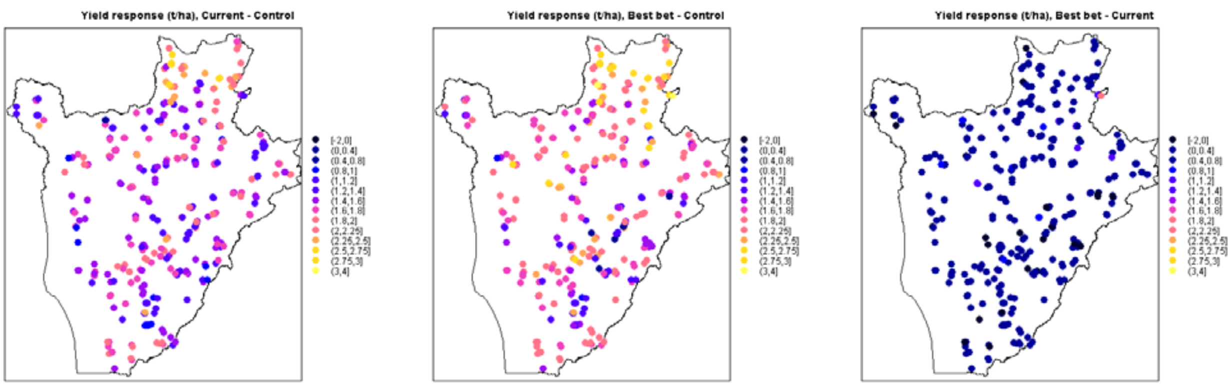


Figure 8. *Difference in yield response between the fertilizer treatments for beans.*

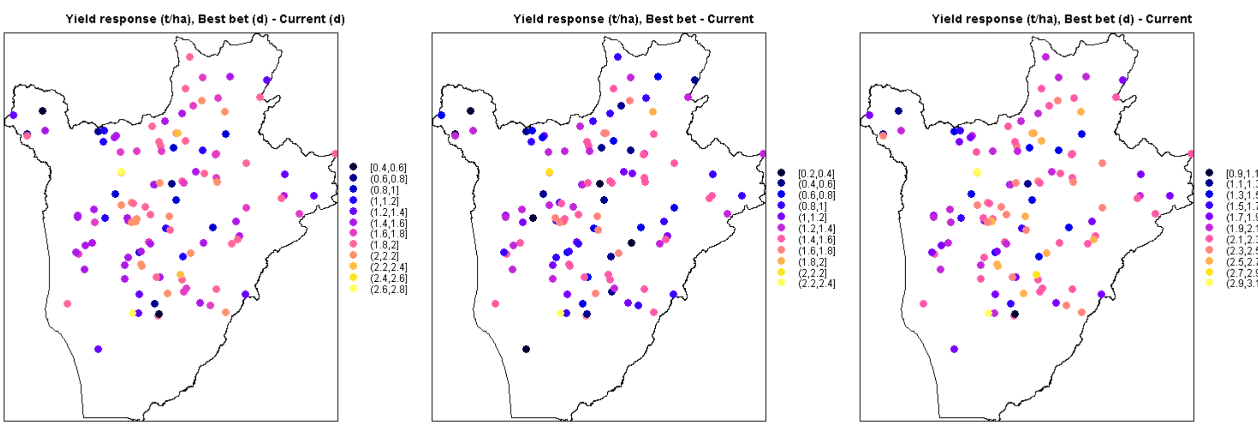


Figure 9. *Difference in yield response between the fertilizer treatments for maize.*

Results

Random forest modeling

The variable important measures show striking differences in covariate relevance between Burundi and Rwanda. For each model we tallied the five most relevant covariates, i.e. the five covariates that result in the largest increase in prediction accuracy and grouped them according to covariate categories: climate, soil, vegetation, physiography, land surface reflectance. The results are summarized in Table 13.

Table 13. *General covariate importance based on the five most important covariates for each soil property (11 properties times five covariates gives 55 counts).*

	Burundi	Rwanda
Climate	25	7
Soil	13	19
Vegetation	6	3
Physiography	8	19
Land surface reflectance	3	7

For Burundi, covariates related to climate were generally the most important. These mainly related to daytime temperature (11). For the soil group, soil moisture status was the most important one (12). For the vegetation group the NDVI (4), for physiography e-SOTER terrain units (4). The most relevant covariates for Rwanda were related soil and physiography. For the soil group, the dominant soil class according to SOTERNEA (8) was the most important covariate followed by soil moisture status (7). For the physiography group, the e-SOTER terrain units (8), followed by landform class (5) and lithology (4). For the climate group, daytime temperature (3) and precipitation (3), for vegetation the NDVI (3) and for land surface reflectance the MODIS mid-infrared band (3). A reason for this difference can be that the spatial variation in soils is larger in Rwanda than in Burundi, which is evidenced by the (very general) soil map in Figure 4 and the soil property summary statistics in Table 12. Variation between soils classes, and related to this landform and lithology classes, might therefore be larger for Rwanda and Burundi, resulting in a stronger association between these covariates and the target soil properties. The strong association between daytime temperature and soil properties in Burundi might represent an orographic effect since daytime temperature is strongly correlated with elevation (Pearson's r = -0.72). Because of the complexities of the random forest models it is very hard to unravel the exact effects of the covariates on the soil properties.

Semivariogram modeling and kriging

The random forest residuals were examined for spatial correlation by fitting semivariogram models to the residuals. The semivariograms for Burundi are shown in Figure 10, those for Rwanda in Figure 11. The Burundi semivariograms show that all properties, except Cu and OM, lack spatial structure in the residuals evidenced by a (almost) flat line. Semivariograms fitted to the raw data (not shown here) showed that for all soil properties there was a degree of spatial correlation. Apparently, the random forest models were able to capture almost all spatial structure in the data. Kriging was carried out only for Cu and OM. The other properties were mapped using the random forest model only. The semivariograms for Rwanda show that the residuals of the random forest models were spatially correlated for B, Ca, Cu, ECEC, Mg. For these properties the residuals were kriged and added to the random forest predictions. For OM, estimating the nugget effect proved to be problematic; the semivariogram model in Figure 6 has nugget of zero, which is unrealistic. Therefore we decided not to carry out the kriging step for OM.

Figure 12 shows the soil maps of Burundi, Figure 13 those of Rwanda. Both countries show distinctive spatial patterns. The range of predicted values is larger for Rwanda than for Burundi for most properties. Soil nutrient status is more favorable for Rwanda than for Burundi with on average larger absolute nutrient concentrations. The largest Ca, Mg concentrations and ECEC values are found in the alluvial plains north of Bujumbura and in Kirundo and Muyinga provinces in the north-central part of the country. These are also the locations where relatively large B concentrations are mapped. The largest OM values are predicted in the Burundi Highlands mountain range in the western part of the country that show that the spatial pattern is dominantly influenced by elevation and temperature effects. Also for S, the largest concentrations are found along the Burundi Highlands, which likely is a geological effect. The most important covariates for modeling S were elevation that likely serves a proxy for geology, and the SOTERNEA soil map. The pH map shows three distinct areas: the Burundi Highlands with typically pH<5, the alluvial plains north of Bujumbura with pH>6, and the rest of the country with pH between 5 and 6. Spatial patterns of P and Zn show similarities. In Rwanda there is a clear distinction between the mountainous western part of the country and more hilly eastern part as a result of the dominating effect of soil and geography on soil nutrient conditions. In the western part, soils are relatively acid with low ECEC and with relatively low concentrations of nutrients, except for S and OM. The eastern part is characterized by larger pH values and more favorable nutrient concentrations.

Cross-validation

The cross-validation statistics are given in Table 14. The plots in Figures 14 and 15 show the relationship between predicted and observed soil property values. The statistics show similar trends for Burundi and Rwanda, except for Ca, Mg and pH that are predicted substantially better for Rwanda than for Burundi based on R^2. P, K and Zn prove the most challenging to predict based on the R^2 values that are typically between 0.20 and 0.25. K for Burundi is even as low as 0.12. These were also the properties with the weakest spatial correlation structure in the raw data (small sill/nugget ratio, small range; data not shown). It is therefore not surprising that the relationship with spatial covariates is limited. OM is best predicted for Burundi with an R^2 of 0.45. The R^2 of the other properties range between 0.31 (pH) and 0.39 (B). The Rwanda statistics show that Ca, Mg, pH, Cu and ECEC are predicted best with R^2 values ranging between 0.40 and 0.45.

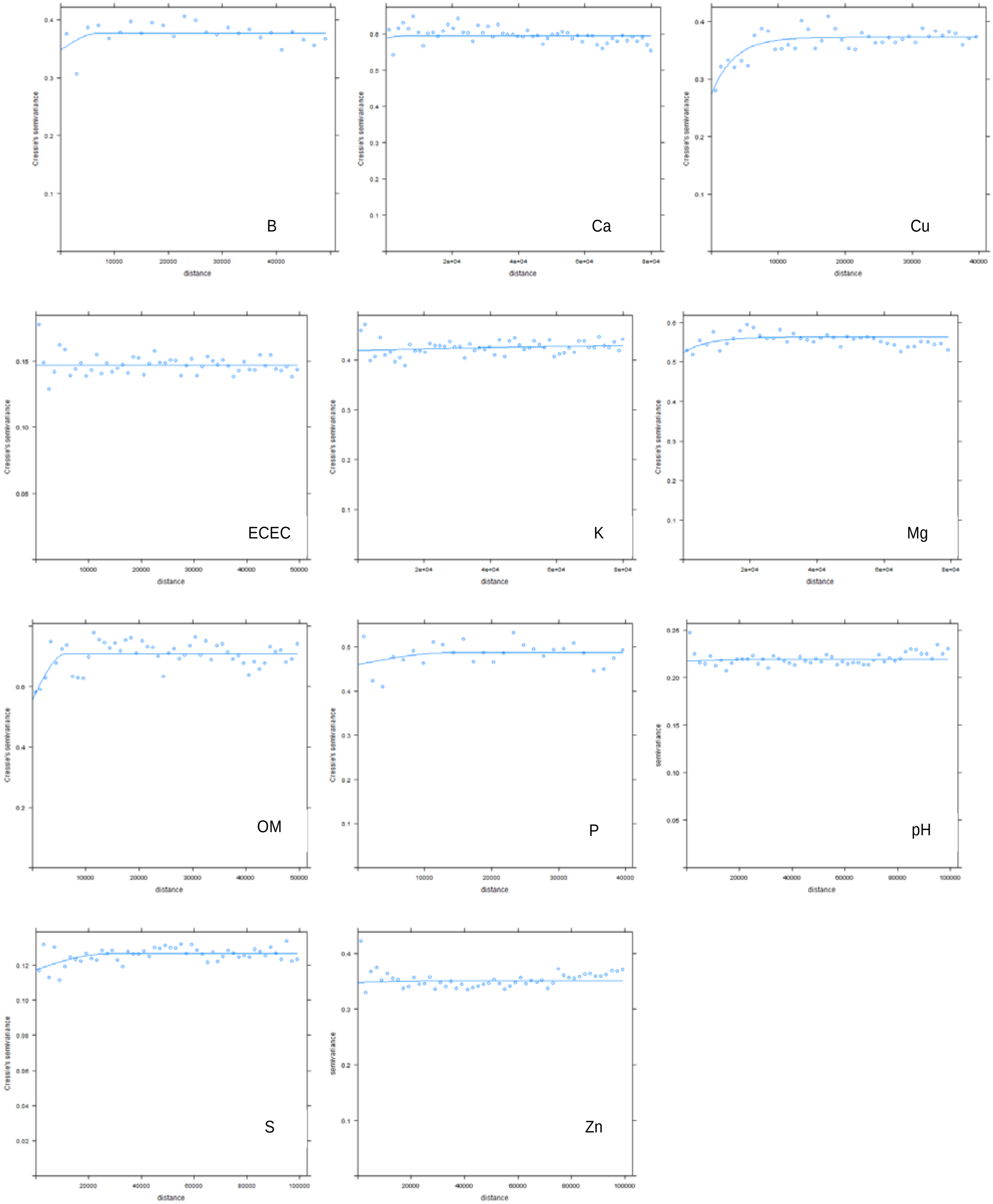


Figure 10. *Sample semivariograms (dots) and fitted semivariogram models (solid lines) of the Burundi random forest residuals.*

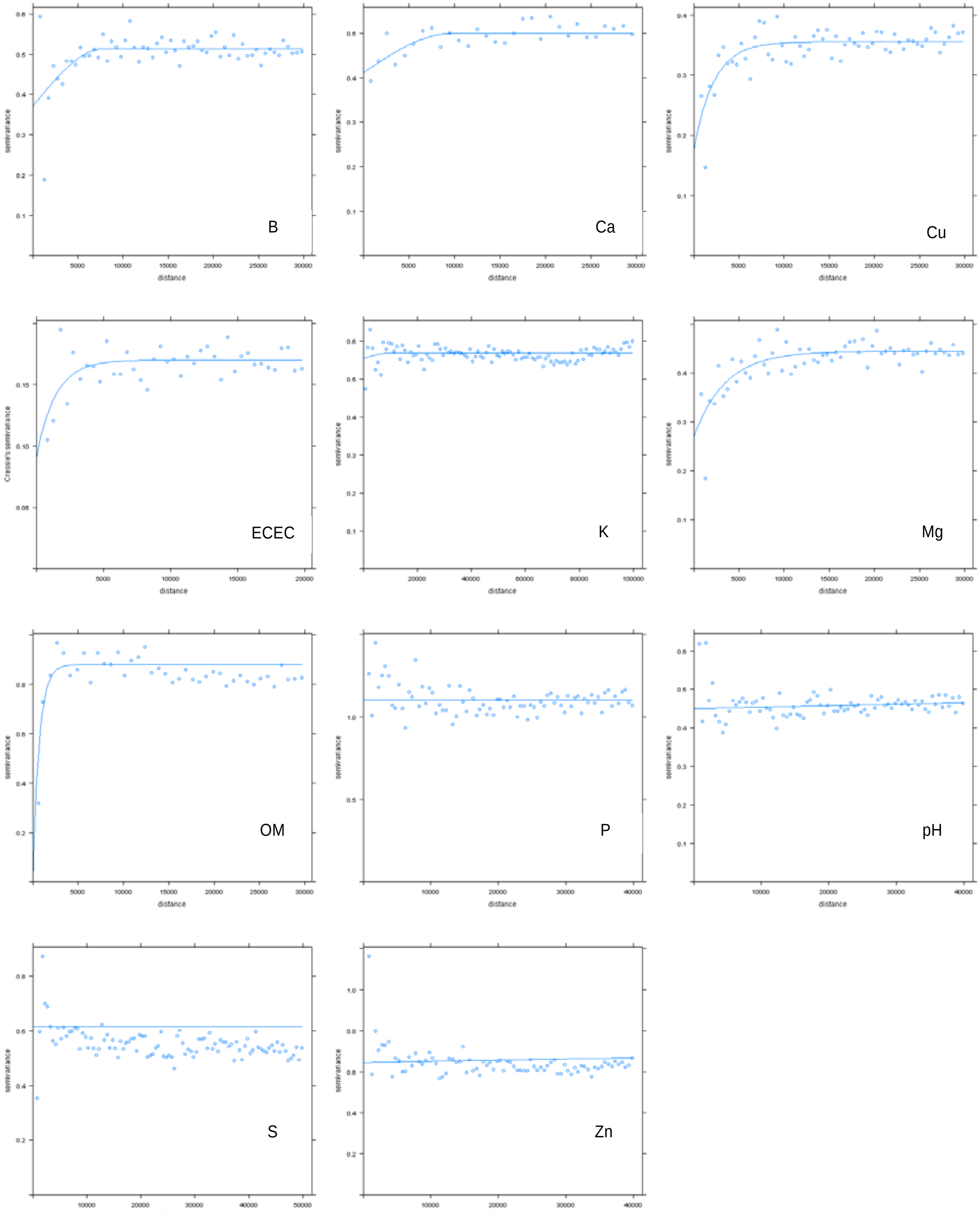


Figure 11. *Sample semivariograms (dots) and fitted semivariogram models (solid lines) of the Rwanda random forest residuals.*

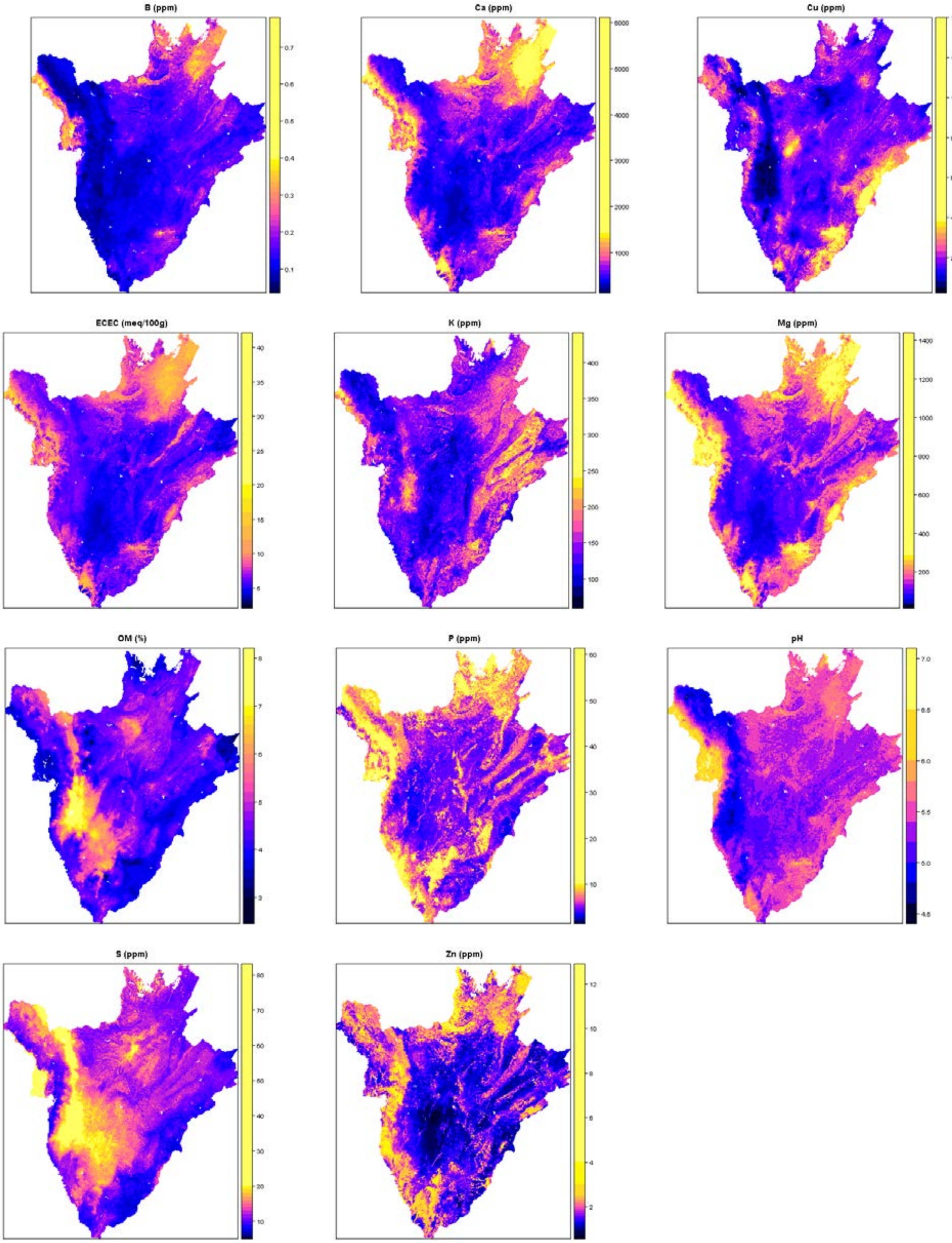


Figure 12. *Soil maps of Burundi generated with RF modeling (B, Ca, ECEC, K, Mg, P, pH, S, Zn) or RF-kriging (Cu, OM).*

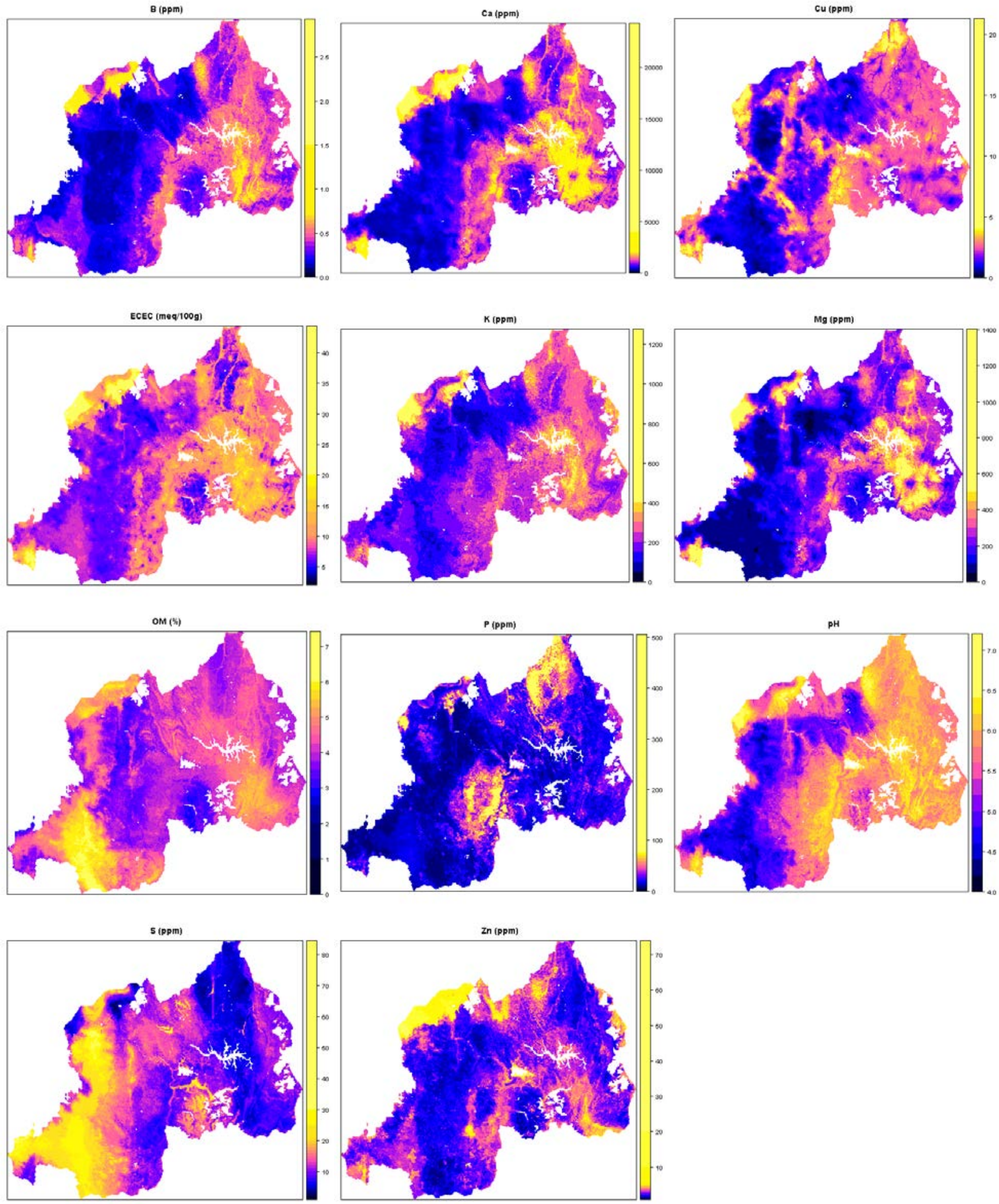


Figure 13. *Soil maps of Rwanda generated with RF modeling (K, OM, P, pH, S, Zn) or RF-kriging (B, Ca, Cu, ECEC, Mg).*

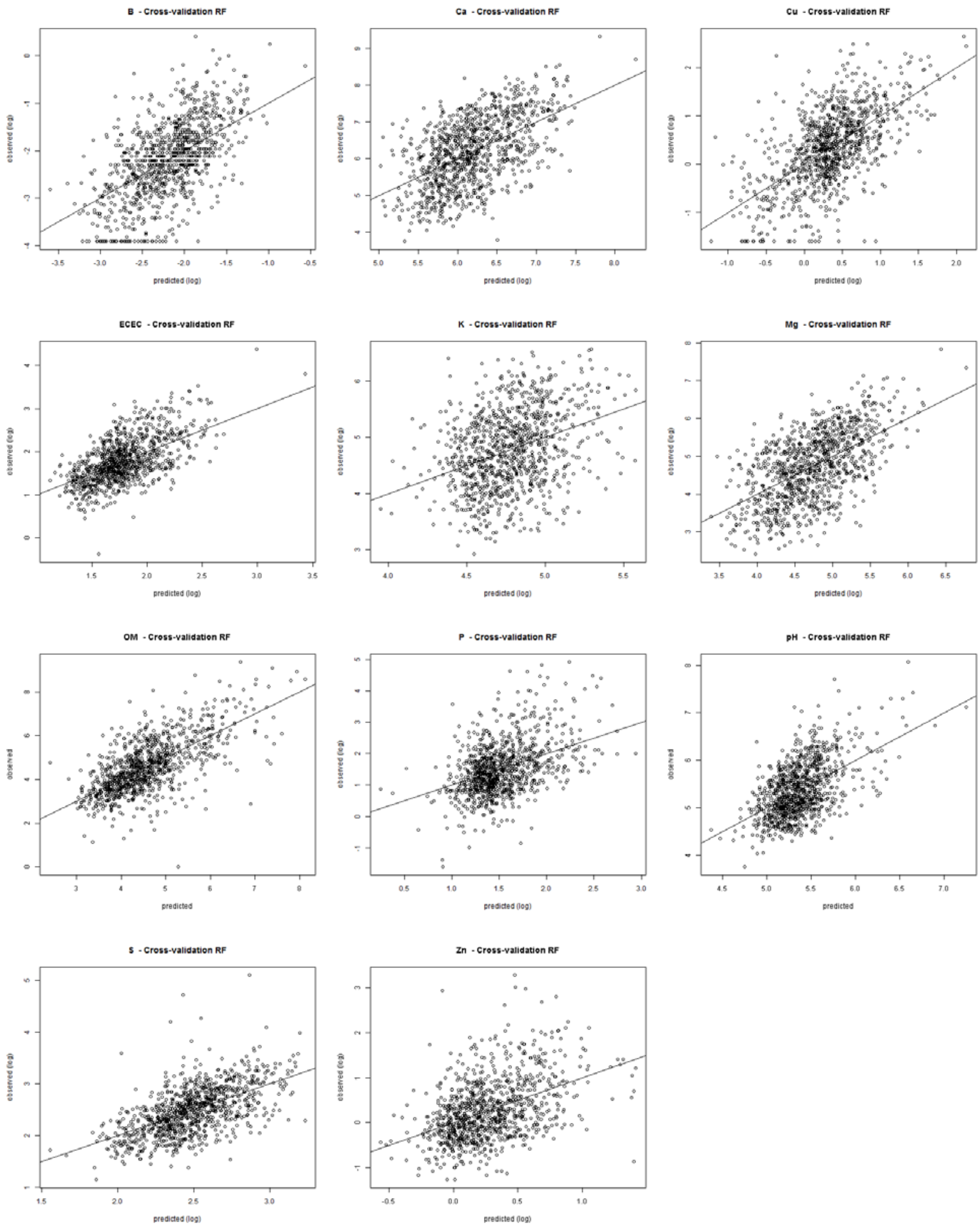


Figure 14. *Correlation plots of predicted versus observed values for Burundi.*

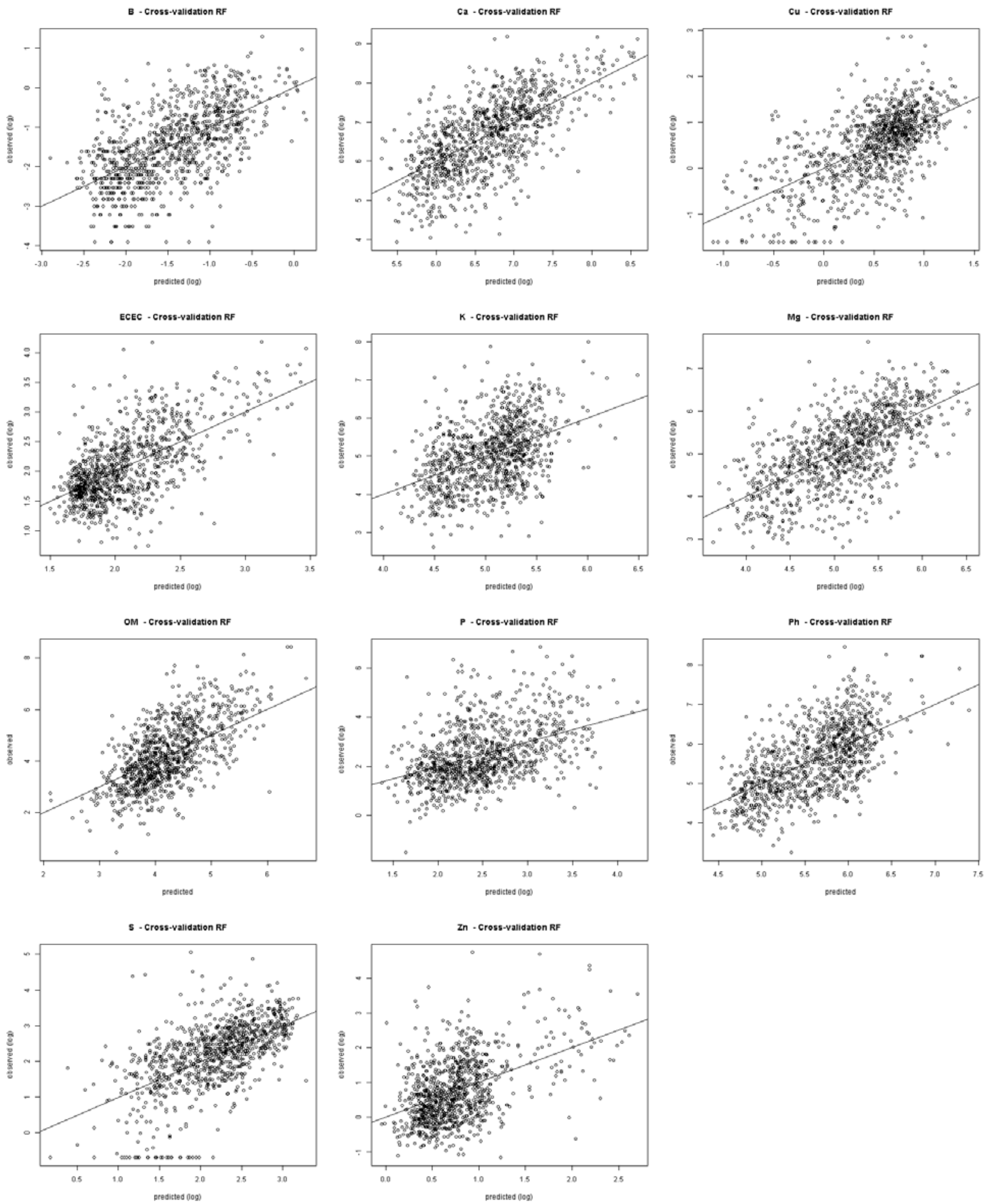


Figure 15. *Correlation plots of predicted versus observed values for Rwanda.*

Table 14. *Cross-validation statistics. ME is the mean error, MAE the mean absolute error, RMSE the root mean squared error, RMedSE the root median squared error and R^2 the coefficient of determination. The R^2 is computed on log-scale for the log-transformed properties.*

Burundi						
	pH	OM (%)	P (ppm)	K (ppm)	Ca (ppm)	Mg (ppm)
ME	0.00	-0.01	-0.7	0.9	0.5	-0.4
MAE	0.39	0.67	4.7	76	418	92
RMSE	0.48	0.90	10.9	104	629	141
RMedSE	0.34	0.50	2.5	63	29	64
R^2	0.31	0.45	0.22	0.12	0.37	0.35

Burundi					
	S (ppm)	Zn (ppm)	B (ppm)	Cu (ppm)	ECEC (meq $100g^{-1}$)
ME	-0.2	-0.08	0.00	0.00	-0.2
MAE	4.0	0.91	0.07	0.88	2.3
RMSE	8.0	1.86	0.12	1.36	3.9
RMedSE	2.8	0.56	0.05	0.62	1.5
R^2	0.37	0.20	0.39	0.36	0.38

Rwanda						
	pH	OM (%)	P (ppm)	K (ppm)	Ca (ppm)	Mg (ppm)
ME	0.00	-0.03	-5.6	-5.8	12.2	3.3
MAE	0.54	0.75	26.6	137	582	111
RMSE	0.68	0.94	69.8	232	917	171
RMedSE	0.47	0.63	12.3	92	367	75
R^2	0.40	0.42	0.22	0.21	0.45	0.42

Rwanda					
	S (ppm)	Zn (ppm)	B (ppm)	Cu (ppm)	ECEC (meq $100g^{-1}$)
ME	0.2	-0.35	0.00	0.01	-0.5
MAE	6.3	2.53	0.20	0.86	3.6
RMSE	11.3	6.82	0.31	1.43	5.8
RMedSE	4.3	1.34	0.12	0.61	2.2
R^2	0.37	0.25	0.37	0.40	0.40

The models roughly explain between 20 and 40% of the variation the data. This does not seem like much. However, the semivariogram of the RF residuals in Figures 10 and 11 show there is not much spatial structure left. The RF models captured most of the spatial structure in the data. Most of the unexplained variation is random noise. It is therefore unlikely that much improvement can be expected by extending the set of environmental covariates or collecting more data.

The mean errors are for most properties close to zero, which means predictions are unbiased: there is no systematic over- or underestimation. Predictions for P, K and Ca and Mg seem to be somewhat biased for Rwanda. However, if the MEs of K, Ca, and Mg are compared to the averages the predicted values (216 ppm for K, 1102 ppm for Ca, and 230 ppm for Mg) this bias is negligible and it can be questioned whether the MEs differ significantly from zero. For P the average predicted value is 25 ppm. A bias, in this case underestimation, of 5 ppm is substantial. This bias is caused by the presence of several very large P concentrations (>600 ppm) while the median P value in the sample data is 9 ppm. These large values are not well predicted by the RF model leading to a set of very large errors that cause bias. On log-scale predictions are unbiased.

The cross-validation results show that the RMedSE are in most cases substantially smaller than the RMSE. This means that for most properties some large errors are present that inflate the mean squared error (MSE; the square root of which is the RMSE). The RF and kriging predictors are smoothers. Extreme values in the data are often not reproduced by these predictors, which can lead to gross errors. Because the errors are squared to compute the MSE and RMSE presence of large errors can greatly inflate the MSE and RMSE. For example, in case of P, 25 sampling sites out of 999 had a concentration >200 ppm, 4 had a concentration >600 ppm, and 2 had a concentration >800 ppm. In case of such extreme concentrations (at least compared to the mean and median) one should check if such concentrations are realistic and can occur under local circumstances. Here we assumed that this is the case. By taking the natural logarithm of the measured values before modeling reduces the influence of extreme values on model fitting.

Uncertainty assessment

Figure 16 shows an example of quantitative information about uncertainty of a predictive soil map in the form of a prediction interval. In this case, it shows the lower and upper boundaries of the 90% PI for predicted B in Rwanda. From these boundaries the width of the prediction interval can be calculated, shown in Figure 17. This figure shows that the uncertainty associated to the predictions is not constant in space. In some areas the uncertainty is larger than in other areas. The spatial pattern resembles the pattern of predicted B (Figure 13). Areas with a relative high uncertainty are generally areas that have predicted concentrations that are relatively large. Prediction uncertainty can be reduced by predicting average concentrations for the pixels (block kriging) instead of predicting a concentration at one point (the center) in each pixel (point kriging as was done here).

Information about prediction uncertainty can be used to compute the probability that a certain critical threshold will be exceeded. Figure 18 gives an example of four possible realities of the spatial distribution of pH in Burundi that are all as likely. (Remember that the kriging map gives us only the most likely value for each prediction location. Reality will differ from these values because we are uncertain about these). Simulations are repeated a large number of times, typically 500 to 1,000. To construct a map that shows the probability of exceeding a certain critical threshold, one determines for each simulated value if the threshold is exceeded or not by assigning an indicator to each prediction location that takes value 1 if the threshold is exceeded and 0 otherwise. This results in 500 to 1,000 indicator values for each prediction location from which a probability can be calculated. An example for pH in Burundi is shown in Figure 19. Note that in this example, no covariates were used.

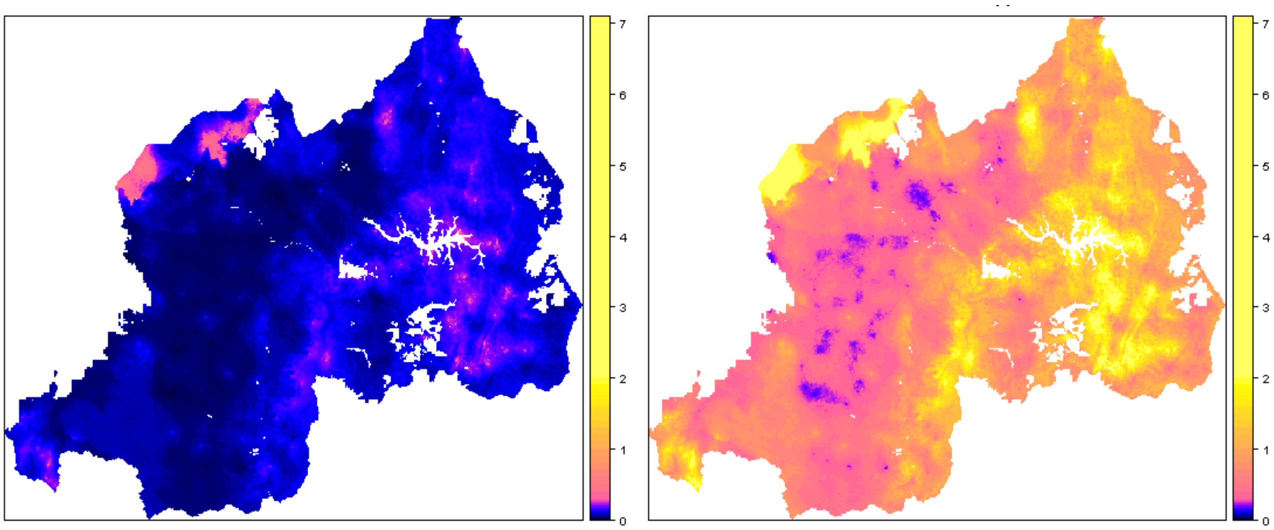


Figure 16. *Lower (left panel) and upper (right panel) boundary of the 90% prediction interval (ppm) for predicted B in Rwanda.*

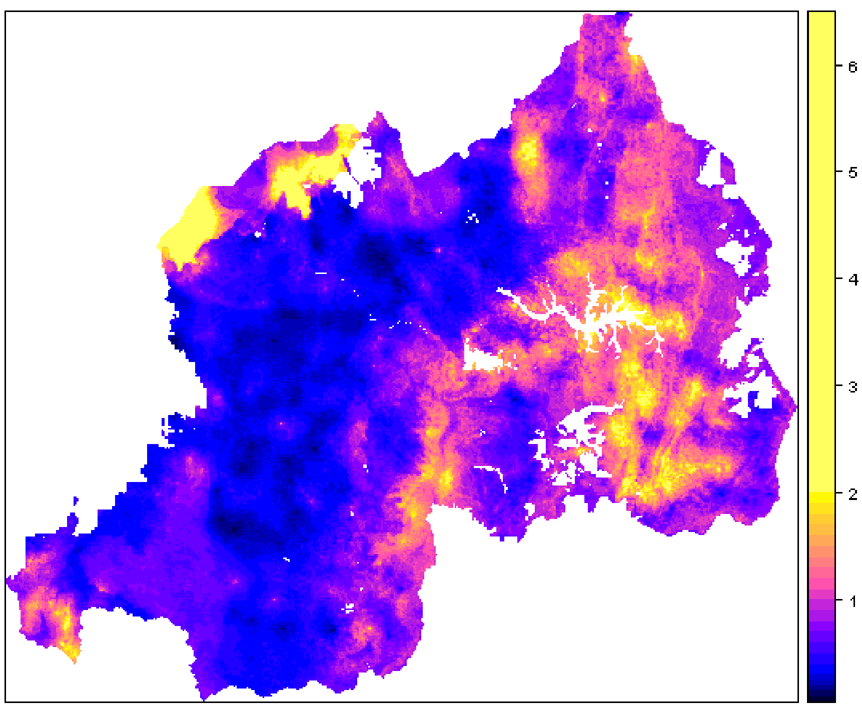


Figure 17. *The width of the 90% PI for B in ppm.*

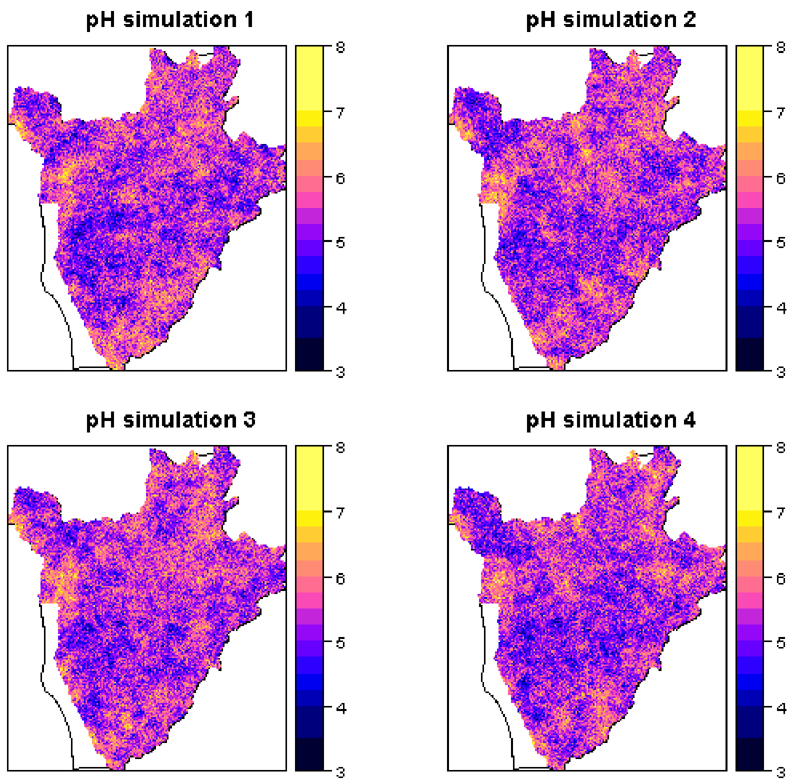


Figure 18. *Four possible realities of the spatial variation of pH in Burundi generated through geostatistical simulation.*

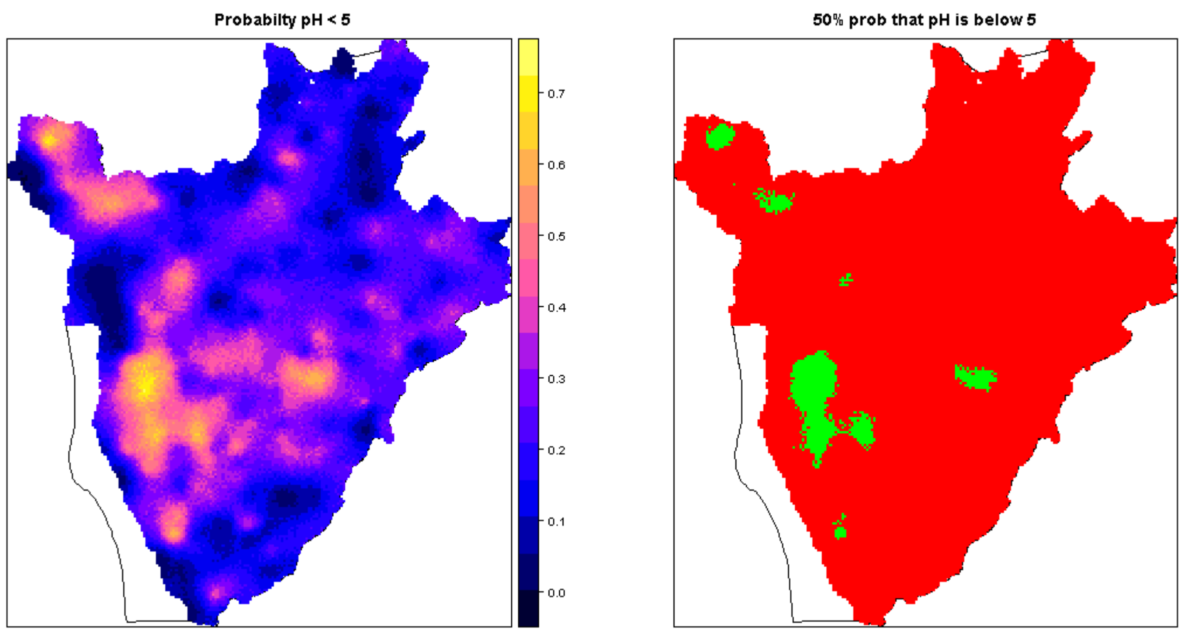


Figure 19. *Probability that pH is below 5 (left panel), and areas that have a 50% probability that the pH is below 5 (depicted in green; right panel).*

Spatial analysis of crop yield data

Figures 20 and 21 show semivariograms for yield responses and differences between yield responses for beans. Figures 22 shows the semivariograms for yield responses for maize. Figure 23 shows the semivariograms for the most relevant yield response differences for maize. The semivariograms show that there is spatial structure in both maize and beans yield responses but that spatial correlation is much stronger for the bean yields and bean yield differences (larger range, smaller nugget effect) than for the maize yields and yield differences. The fact that for maize far fewer data were available to estimate the semivariogram might contribute to this observation.

Figure 24 shows the response maps for the three fertilizer treatments for beans. The maps show clear spatial patterns. The orange and yellow colors indicate areas with yields larger than average, while blue and black colors indicate areas with yields smaller than average. The yield response difference maps depicted in Figure 25 show clear spatial patterns as well. Some areas have a stronger response to the addition of soil micronutrients (SMNs) to NPK than other areas (Figure 25, right panel). Similar maps for maize are shown in Figures 26 and 27. Effects of the sparse sampling density and less strong spatial correlation are clearly shown in the spatial patterns of these maps. Nevertheless, there are clear regional differences in response to the treatments.

Finally, we investigated if the spatial variation of the yield responses and yield response differences between treatments for the *maize* crop (season 2015a) could be correlated to the spatial variation in soil fertility parameters. We used both measured and mapped parameters for this analysis. The mapped parameters were extracted from the soil maps at the locations of the field trials through a spatial overlay. Correlation was quantified with the Pearson's correlation efficient r. Results are shown in Table 15.

The analysis showed that correlations between soil fertility parameters and absolute yield responses and yield response differences were generally very weak (r <0.2), except for pH and S. Relationships between the measured soil fertility parameters and yield responses were stronger than the relationships between the mapped parameters and yield responses. This is not surprising given that the mapped value is an interpolated value that is subject to error. Analyzing the relationship between mapped parameters and yield response is relevant, however, because if one wants to (spatially) upscale the results from field scale to regional or national scale one has to rely on mapped values since measurements are not available exhaustively.

Weak (0.2< r <0.3), moderate (0.3< r <0.4) or strong (0.4< r <0.6) positive relationships between pH and maize yields (Table 15). Correlation between pH and yield was stronger for the treatments where NPK or NPK+SMN were given than for the control treatment. This might indicate that when NPK or NPK+SMN are applied, then there is a stronger effect of pH on yield then when these fertilizers are not applied. In other words, NPK or NPK+SMN additions become more effective with increasing pH (at least in the pH 5-7 range).

S shows weak to strong negative relationships with yield response, meaning that yields tend to get smaller when S concentration increases (Table 15). The effect of S on yield might be confounded with climatic factors such as temperature. The S map in Figure 10 shows that relatively large S concentrations are found in the Burundi highlands where lower average temperatures might negatively affect yield that is reflected by the relationship with S. A similar effect might explain the negative correlation between OM and yield response.

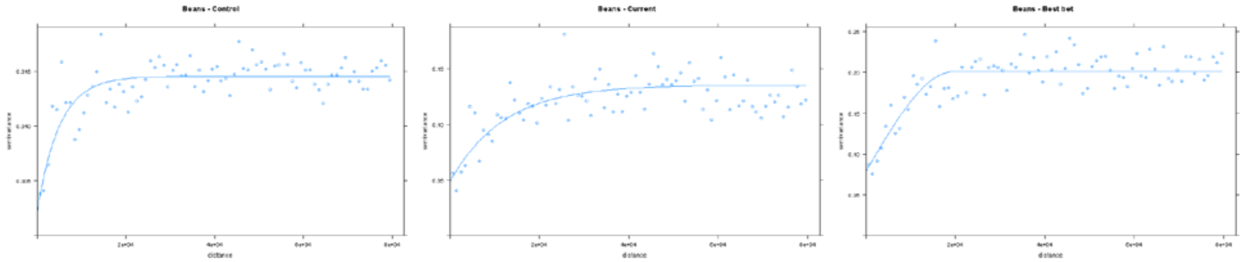


Figure 20. *Semivariograms for yield responses to the control, current and best bet treatment for beans.*

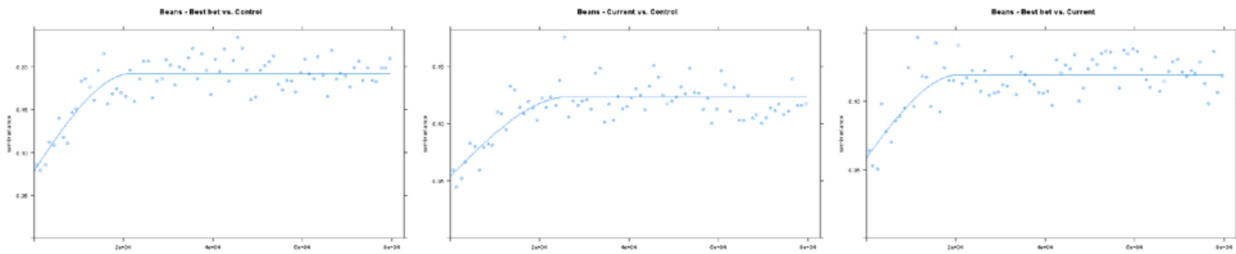


Figure 21. *Semivariograms for differences in yield responses between the fertilizer treatments for beans.*

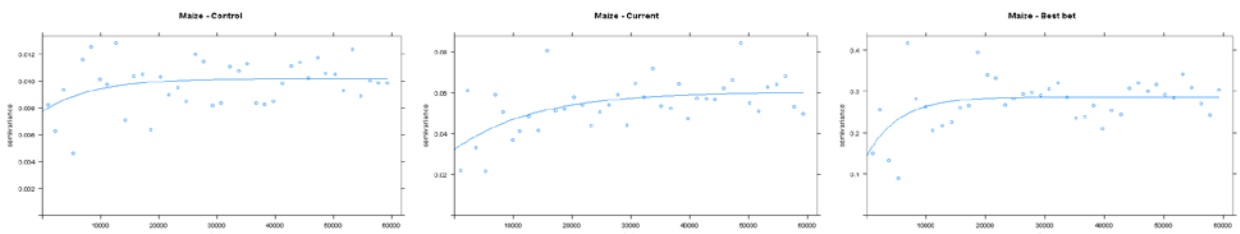


Figure 22. *Semivariograms for yield responses to the control, current and best bet treatment (without dolomite application) for maize. The semivariograms of the treatments with dolomite show a similar shape.*

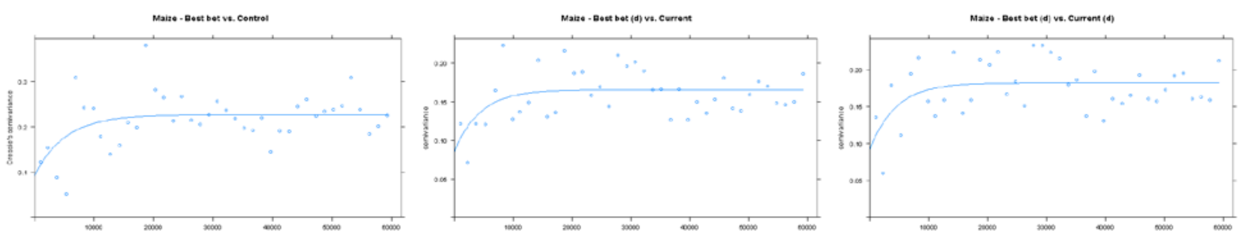


Figure 23. *Semivariograms for differences between yield responses for fertilizer treatments for maize. The '(d)' in the title means dolomite application.*

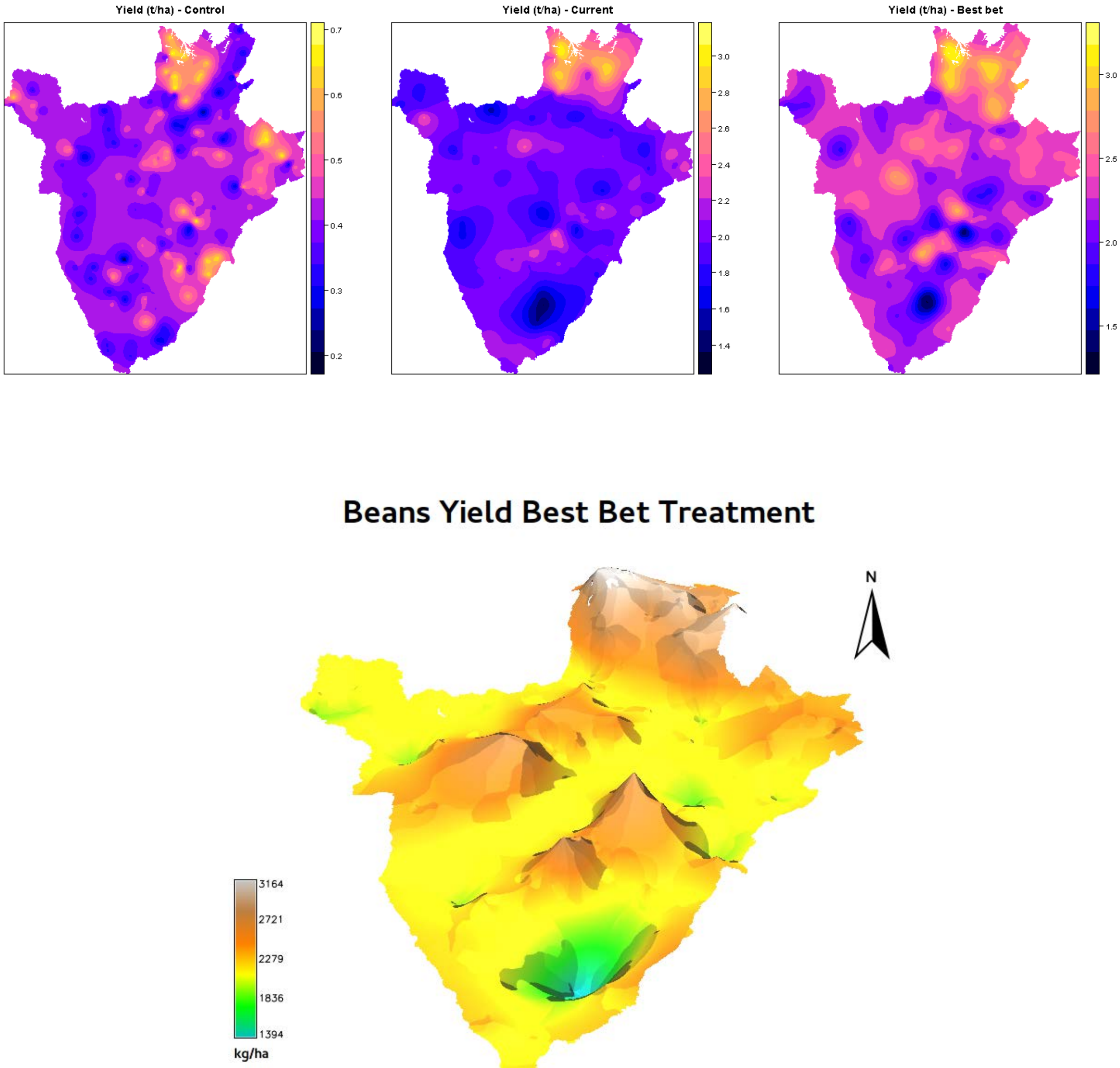


Figure 24. *2D (top) and 3D (bottom) visualizations of interpolated bean yields.*

Note that the color scales are different and are chosen to maximize contrast in each map. Note that the bean maps do not take into account the differences between bean varieties. Based on yield evaluations, the bush variety generally yields less than the climbing variety.

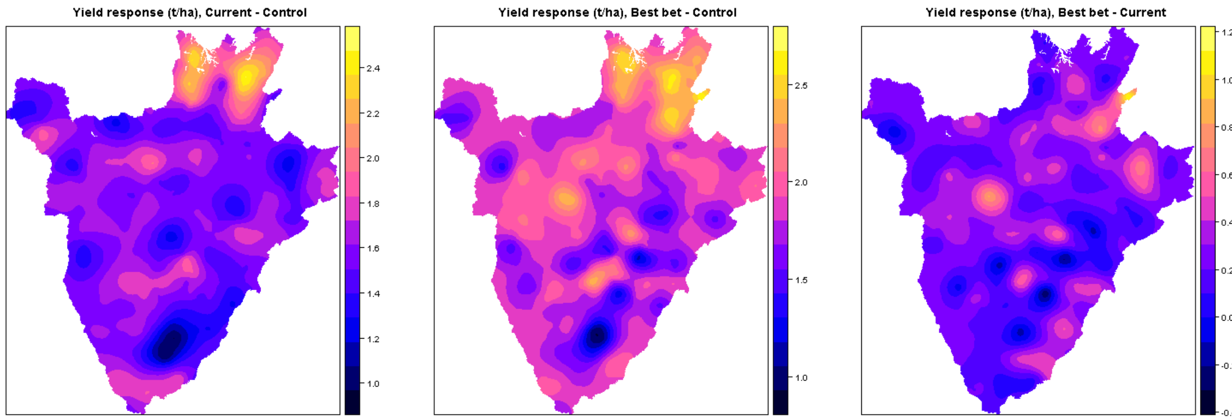


__Figure 25.__ Maps of yield response differences for beans.

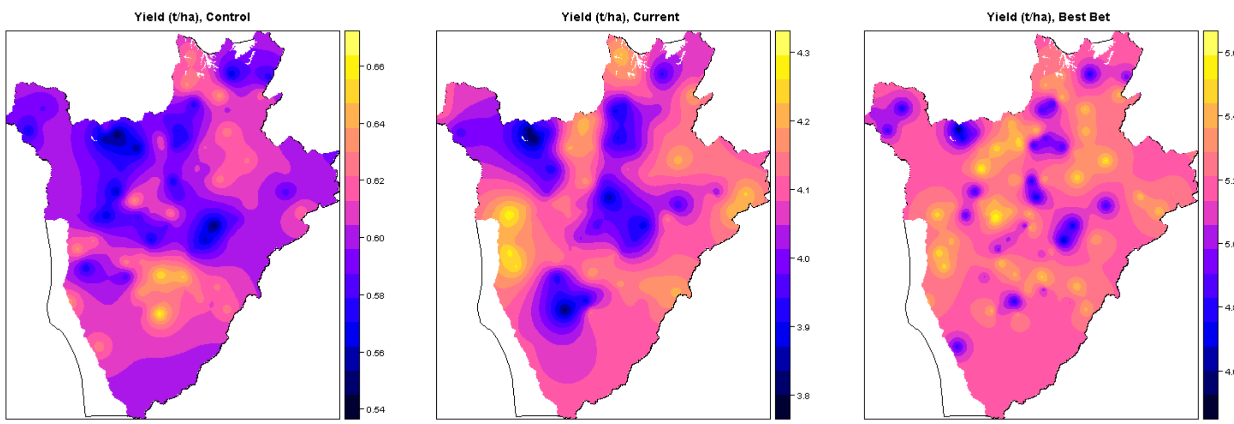


__Figure 26.__ Maps of interpolated yield response for maize. Note that the color scales are different and are chosen to maximize contrast in each map.

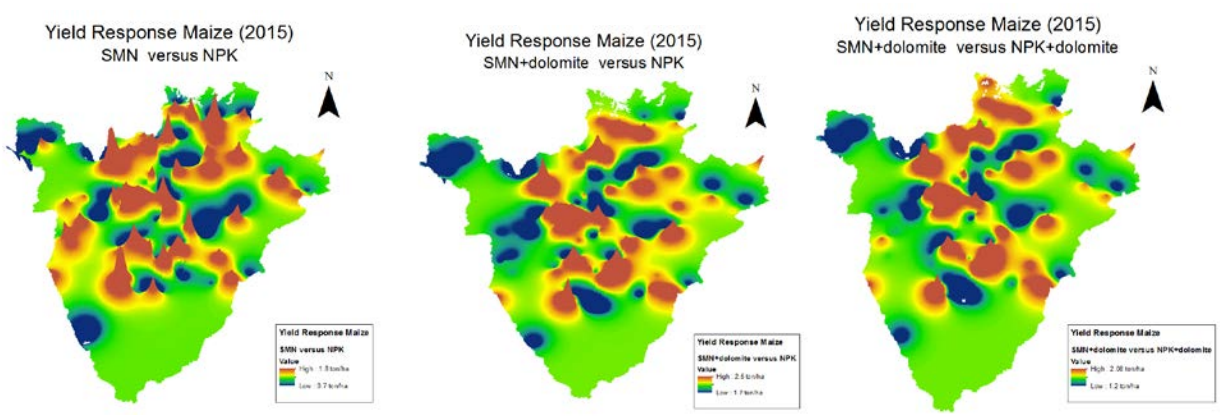


__Figure 27.__ 3D visualizations of yield response differences for maize.

Table 15. *Correlation coefficients between yield and yield responses (T_x-T_y) for maize and mapped and measured soil fertility parameters. '+D' indicates dolomite application.*

Mapped	T0 Control	T1 Control+D	T4 NPK	T5 NPK+d	T2 NPK+SMN	T3 NPK+SMN+d
B	0.11	0.08	0.20	0.23	0.24	0.14
Ca	0.19	0.15	0.23	0.23	0.29	0.16
Cu	0.09	0.07	0.09	0.07	0.09	0.05
ECEC	0.09	0.07	0.07	0.09	0.19	0.09
K	0.06	0.02	0.04	0.01	0.16	0.07
Mg	0.17	0.14	0.18	0.19	0.33	0.19
OM	-0.08	-0.05	-0.18	-0.19	-0.19	-0.05
P	0.13	0.13	-0.09	-0.03	0.01	-0.01
pH	0.29	0.24	0.37	0.36	0.43	0.32
S	-0.29	-0.24	-0.31	-0.31	-0.41	-0.29
Zn	0.01	0	0.06	0.01	0.09	-0.02

Measured	T0 Control	T1 Control+D	T4 NPK	T5 NPK+D	T2 NPK+SMN	T3 NPK+SMN+D
B	0.11	0.08	0.19	0.06	0.17	0.13
Ca	0.27	0.23	0.33	0.23	0.27	0.26
Cu	0.09	0.06	0.07	0.03	0.07	0.06
CEC	-0.01	-0.03	0	-0.08	0.02	-0.04
K	0.09	0.04	0.17	0.05	0.06	0.01
Mg	0.24	0.2	0.38	0.24	0.23	0.22
OM	-0.08	-0.08	-0.13	-0.08	-0.11	-0.16
P	0.08	0.06	0	-0.05	-0.05	-0.02
pH	0.39	0.32	0.52	0.44	0.47	0.45
S	-0.37	-0.31	-0.47	-0.31	-0.35	-0.32
Zn	-0.01	-0.03	0.02	-0.09	0.04	-0.02

Mapped	T3-T5	T2-T4	T3-T1	T2-T0	T5-T1	T4-T0	T5-T4	T3-T2
B	0.19	0.11	0.23	0.2	0.11	0.23	-0.26	0.01
Ca	0.17	0.11	0.21	0.21	0.11	0.25	-0.32	-0.05
Cu	0.06	0.06	0.06	0.08	0.02	0.06	-0.10	-0.04
ECEC	0.05	0.03	0.08	0.06	0.06	0.18	-0.24	0.02
K	-0.02	-0.04	0	0.03	0.06	0.15	-0.22	-0.06
Mg	0.12	0.04	0.17	0.16	0.14	0.30	-0.35	0.03
OM	-0.18	-0.12	-0.19	-0.18	-0.03	-0.18	0.30	0.04
P	-0.02	-0.12	-0.05	-0.12	-0.07	-0.05	-0.04	0.13
pH	0.25	0.22	0.34	0.34	0.23	0.35	-0.35	-0.09
S	-0.21	-0.15	-0.28	-0.27	-0.19	-0.33	0.36	0.06
Zn	0.02	0.03	0.02	0.07	-0.02	0.10	-0.21	-0.10

Measured	T3-T5	T2-T4	T3-T1	T2-T0	T5-T1	T4-T0	T5-T4	T3-T2
B	0.12	0.11	0.12	0.17	0.03	0.16	-0.28	-0.11
Ca	0.18	0.15	0.23	0.24	0.13	0.25	-0.30	-0.08
Cu	0.06	0.05	0.06	0.06	0	0.04	-0.09	-0.03
CEC	-0.01	0.02	-0.04	0.02	-0.07	0.01	-0.15	-0.11
K	-0.02	-0.03	0	0.04	0.04	0.15	-0.25	-0.09
Mg	0.13	0.07	0.19	0.20	0.16	0.32	-0.37	-0.07
OM	-0.14	-0.07	-0.16	-0.11	-0.05	-0.11	0.13	-0.06
P	0	-0.06	-0.03	-0.07	-0.08	-0.04	-0.07	0.06
pH	0.30	0.30	0.41	0.43	0.31	0.41	-0.32	-0.14
S	-0.21	-0.17	-0.28	-0.31	-0.19	-0.37	0.44	0.14
Zn	0.01	0.04	-0.03	0.04	-0.09	0.02	-0.19	-0.12

Concluding remarks

This report presents the results of exploratory analyses of the use of geospatial information and modeling approaches to refine soil fertility recommendations. The first part of this report focused on the use of linear mixed models to explore if observed differences between farms in yield response to a specific fertilizer treatment are correlated to soil or climate conditions. A (significant) correlation allows the relations to be used to predict yield responses to fertilizer treatments at other locations given their soil and climate conditions. The second part of this report focused on spatial analysis and interpolation of soil and yield data.

Based on the results presented here, we cannot draw any general conclusions about the merit or added value of geospatial analysis of yield and soil fertility data to refine soil fertility recommendations. This study had limited scope: linear mixed models being fitted for only a few crops using yields of only one harvest season for one specific country (Burundi) and one specific trial type (omission trials). Furthermore, there were some limitations with respect to the data used. Not all relevant data were available. The effect of crop variety was, for instance, unaccounted for since these data were not available for this study. A more complete dataset with variables to be considered should be designed prior to experimentation and data collections. The experimental setup is unbalanced because the fertilizer trials were not designed for the type of analyses presented here. The soil and climate data used as covariates in the linear mixed models were relatively coarse in terms of spatial resolution. Dolomite was not consistently applied to the treatments, which confounds the modeling results, only to those for which the soil pH was less than 5, though this was changed for the 2015 field trials where lime was applied irrespective of soil pH. There were errors in the georeferencing of the field locations that add noise to the linear mixed models and random forest models used for soil mapping. The recorded province of the trial site did, for instance, not match the province according to its spatial location for several sites.

Despite these limitations, some observations can be made from the results of this study for the Burundi case. The results of the modeling exercise indicate that there is a general lack of correlation between soil type and climatic variables and yield. Based on these results we are not able to refine recommendations for Burundi. It indicates that fertilizer formulations suggested by the omission trials are valid throughout the study zone and do not need to be altered to suit certain soils or climate zones. The soil nutrient maps do not appear to be indicative of crop nutrient response, and their utility to refine fertilizer recommendations might be limited. The soil maps, however, are indicative of nutrient deficiencies that might be anticipated and are useful for that. We note, however, that this situation may prove different over more variable environments such as Rwanda, where soils are considerably more variable.

The evidence of added value of geospatial analysis for improving fertilizer recommendations might not be very convincing on basis of the analyses presented here. Nevertheless, we cannot conclude that it has no added value at all given the scope of this study and limitations with respect to the available data. This study was a pilot study and its results and the lessons learnt can be used to scope a more robust and thorough strategy for further geospatial data gathering and analyses.

Covariates for linear mixed modeling

The linear mixed model analysis showed that soil and climate classes had a significant effect on potato yield only, though not sufficiently important to result in alteration of formulations for different regions. No significant effects were found for maize and beans. Soil and climate class data were coarse. The SOTER soil map used with scale 1:1 million represents a minimum mapping unit of approximately 25 km2. Soils were classified at the level of the FAO major soil groups. Variation in soil properties within these groups can be large. The climate data layers had a spatial resolution of 1 km and represented long-term averages. These coarse data, in combination with an unbalanced experimental design, may have been inadequate to significantly disentangling cause and effect.

The more detailed national soil map of Burundi with scale 1:50,000 was not available in a usable GIS-format for this study. This map is likely to better reflect local soil conditions than the SOTER map. One linear mixed model analyzed the effect of local soil conditions in the form of a set of seventeen soil fertility parameters, summarized in principal components. Principal components analysis (PCA) is an effective method to reduce the dimensions of a dataset. PCA generally captures over 90% of the variation in the data in three variables. The components could be regarded as a summary of the soil fertility status of the trial plots. Because of data transformation, the results are difficult to interpret and to associate with a particular piece of land and, therefore, would not be useful for targeting purposes. The first principal component (PC), and the only relevant one with respect to yield data, can be interpreted as the contrast between acid and non-acid soil conditions (Appendix 4). However, the effect of this component is likely to be confounded by dolomite application.

An alternative to PCs would be to evaluate the soil fertility parameters individually in a linear mixed model, but a single parameter might not be very informative about a possible yield response given synergistic and antagonistic nutrient effects (Rietra et al., 2015). Pairwise comparisons revealed that correlations between the individual soil fertility parameters and yield responses were largely absent, at least for the three crops in Burundi considered here, while some weak relations could be related to edaphic processes such as between Ca and Mg with pH and some other nutrients (Appendix 4). Thus, for Burundi, the soil maps offer very limited utility in predicting crop response. This may be due to the generally low levels of soil nutrients in Burundi, and the situation could be different in regions of higher nutrient levels.

Long-term averages of temperature and precipitation might not very well represent climate conditions during the fertilizer trial season. Water availability during the trial season will have a large impact on crop yield, directly through transpiration and indirectly through availability of nutrients. Climate has very complex interactions with nutrient behavior in the soil, and through these on crop response. These interactions could be disentangled by using data from additional growing seasons. Weather data measured during the trial season at local stations would contain the most effective climate variables to elicit a climate effect on crop yield.

Observed versus mapped data as covariates

Local (observed) as well as predicted (mapped) environmental data can be used as fixed effects in linear mixed models. In ‘Part 1’ of this study, both are used: soil class and climate data were derived from maps while soil property (transformed to PCs) and altitude data came from the trial plots. If our interest would be purely inferential, i.e. we are only interested in understanding the factors that control yield response through statistical modeling, then local covariates are preferred over covariates derived from maps. Local covariates better reflect local conditions that control yield responses than mapped covariates, and will have a stronger correlation with these responses than mapped covariates. This is illustrated by the results in Table 15 that shows larger Pearson's correlation coefficients for soil fertility parameters measured at the trial sites than those derived from soil maps. If our interest, however, would be predictive (spatial), i.e. we are interested at which locations in a country we might find similar yield responses because of similar environmental conditions at these locations, then it is better to use covariate data derived from maps for statistical modeling. If we would fit a model using observed covariate data and subsequently use this model to predict yield response at an observed location, then one will have to rely on mapped covariate data for prediction since local environmental conditions are typically not known at the prediction locations. Kempen et al. (2010) showed that if one would fit a model using local covariate data and then use mapped covariate data to predict, one will underestimate the uncertainty associated to the predictions.

In 'Part 2' of this study, the soil maps of Burundi were used to correlate to yield responses of the maize comparison trial of season 2015a. Of the SMNs, the relationship between Cu and Zn concentration and yield response is negligible. This might be explained by the fact that total amount in the soil might not be indicative of the amount available for uptake by plants, as was recently shown for Zn (Duffner et al., 2014). Only B concentration shows a weak relationship with yield response when NPK or NPK+SMN are applied. The relationships between nutrient concentrations and yield response for the control treatment are negligible. The lack of correlation between SMN and other soil fertility parameters and yield responses might also result from lack of impactful soil physical characteristics in the dataset, such as soil texture and bulk density, that control soil water availability and rooting depth, and rainfall data that control duration of the growing season. These data might have a stronger effect on yield response than (natural) soil fertility conditions.

The correlation analysis in this mapping exercise of the effect of soil properties on yield response is a relatively simple analysis. For more rigorously, more advanced techniques such as linear mixed modeling should be applied. Furthermore, the correlation analysis assumes a linear relation between the soil fertility parameters and yield response. In reality, the relationships may be non-linear. A better approach would then be to convert the soil parameters to factors, for example by quantile splitting, and then analyze the relationship between the factor levels and yield response.

Spatial analysis of yield data

The semivariograms presented in Figures 20 to 23 shows that there is spatial structure in the yield response data for beans (season 2014a) and maize (season 2015a) for the comparison trial treatments. For maize, the semivariograms are somewhat noisy and spatial correlation is not very strong. Beans show much stronger spatial correlation that result in more continuous spatial patterns in the maps (Figures 24 and 25). We should note, however, that bean response is strongly affected by the variety (which was not accounted for here) and dolomite that was applied only to acidic sites.

These findings suggest that the spatial variation in yield response and difference in response is not random for these crops and perhaps for others as well. These spatial patterns suggest that treatments generate a stronger response in area than another, and with that the existence of spatial variation in effectiveness of fertilizer treatment. A next step could be to translate the mapped yield responses to return-on-investment maps. These might help indicate in which areas specific fertilizer treatments are economically viable. These maps could then be further used to estimate how much yield increase could be expected and how much of a specific fertilizer type would be required to achieve this.

Yield responses were mapped for entire Burundi. Some crops, like potato, are only grown in specific areas and the mapping area could then be limited to the regions where the crops are grown, resulting in more realistic estimates of expected yield increase and fertilizer requirement.

These findings warrant further investigation. Increasingly, farmers in developing countries experience lack of response to application of macronutrient fertilizers due a suspected but hard to demonstrate micronutrient deficiency, other than by carrying out large number of field experiments to empirically determine crop responses to micronutrient application. Methodologically linking empirical findings to generic quantitative modeling approaches could allow geo-spatial assessment of micronutrient deficiency and anticipated yield responses at reduced costs. Such information may turn out to be valuable for fertilizer companies and other actors in the chain at regional scale where to sell which fertilizer composition.

Acknowledgments

The authors are grateful to support by staff from the Rwanda Agriculture Board (RAB) and the Institute of Agricultural Science of Burundi (ISABU) in the execution of on-farm trials, and the PAN-PNSEB and CATALIST-2 projects for making available the fertilizer field trial and soil data for this project.

References

Africa Soil Information Service (AfSIS), 2013. AfSIS Climate Collection: WorldClim, 2013 Release. Palisades, NY: Center for International Earth Science Information Network (CIESIN), Columbia University. ftp://africagrids.net/1000m/WorldClim/. Accessed 7 July 2014.

Africa Soil Information Service (AfSIS), 2014. AfSIS MODIS Collection: Land Surface Temperature, April 2014 Release. Palisades, NY: Center for International Earth Science Information Network (CIESIN), Columbia University. ftp://africagrids.net/1000m/MYD11A2/. Accessed 7 July 2014.

Breiman, L., 2001. Random forests. Machine Learning 45(1), 5-32.

Brus, D.J., Kempen, B., Heuvelink, G.B.M., 2011. Sampling for validation of digital soil maps. European Journal of Soil Science 62(3), 394-407.

Dobos, E., Daroussin, J., Montanarella, L., 2005. An SRTM-based procedure to delineate SOTER Terrain Units on 1:1 and 1:5 million scales, Office for Official Publications of the European Communities, Luxembourg.

Dobos, E., Daroussin, J., Montanarella, L., 2010. A quantitative procedure for building physiographic units supporting a global SOTER database. Foldrajzi Ertesito/Hungarian Geographical Bulletin 59(2), 181-205.

Duffner, A., Weng, L., Hoffland, E., Van Der Zee, S.E.A.T.M., 2014. Multi-surface modeling to predict free zinc ion concentrations in low-zinc soils. Environmental Science and Technology 48(10), 5700-5708.

FAO, 1997. The Soil and Terrain Database for northeastern Africa. FAO Land and Water Digital Media Series 2. CD-ROM, Rome.

Grimm, R., Behrens, T., Märker, M., Elsenbeer, H., 2008. Soil organic carbon concentrations and stocks on Barro Colorado Island - Digital soil mapping using Random Forests analysis. Geoderma 146(1-2), 102-113.

Häring, T., Dietz, E., Osenstetter, S., Koschitzki, T., Schröder, B., 2012. Spatial disaggregation of complex soil map units: A decision-tree based approach in Bavarian forest soils. Geoderma 185–186, 37-47.

Hastie, T., Tibshirani, R., Friedman, J., 2008. The Elements of Statistical Learning. Data Mining, Inference, and Prediction. Second Edition. Springer Series in Statistics. Springer, New York.

Hengl, T., Heuvelink, G.B.M., Rossiter, D.G., 2007. About regression-kriging: From equations to case studies. Computers & Geosciences 33(10), 1301-1315.

Hengl, T., Heuvelink, G.B.M., Stein, A., 2004. A generic framework for spatial prediction of soil variables based on regression-kriging. Geoderma 120(1-2), 75-93.

Isaaks, E.H., Srivastava, R.M., 1989. Applied Geostatistics. 1st ed. Oxford University Presss, Inc., New York, USA.

Kempen, B., Heuvelink, G.B.M., Brus, D.J., Stoorvogel, J.J., 2010. Pedometric mapping of soil organic matter using a soil map with quantified uncertainty. European Journal of Soil Science 61, 333-347.

Knotters, M., Heuvelink, G.B.M., Hoogland, T., Walvoort, D.J.J., 2010. A disposition of interpolation techniques. Werkdocument 190, Statutory Research Tasks Unit for Nature and the Environment, Wageningen, Netherlands.

Kuhn, M., 2015. A Short Introduction to the caret Package. Available at: https://cran.r-project.org/web/packages/caret/index.html.

Liaw, A., 2014. randomForest: Breiman and Cutler's random forests for classification and regression.

Nauman, T.W., Thompson, J.A., 2014. Semi-automated disaggregation of conventional soil maps using knowledge driven data mining and classification trees. Geoderma 213, 385-399.

Nduwimana, O., Nzohabonayo, Z., Hicintuka, C., Nibasumba, M., 2013. Cartographie de la fertilite des sols du Burundi et des besoins des principales cultures vivrieres en elements nutritifs, Ministère de l'Agriculture et de l'Elevage, Institute des Sciences Agronomiques du Burundi, International Fertilizer Development Center, Bujumbura.

Odeh, I.O.A., McBratney, A.B., Chittleborough, D.J., 1995. Further results on prediction of soil properties from terrain attributes: heterotopic cokriging and regression-kriging. Geoderma 67(3-4), 215-226.

Pebesma, E.J., 2004. Multivariable geostatistics in S: the gstat package. Computers \& Geosciences 30, 683-691.

R Development Core Team, 2015. R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria.

Rietra, R.P.J.J., M. Heinen, C. Dimkpa and P.S. Bindraban, 2015. Effects of nutrient antagonism and synergism on fertilizer use efficiency. VFRC Report 2015/5. Virtual Fertilizer Research Center, Washington, D.C. 42 pp.; 16 tables; 1 fig.; 229 ref.

Rivas-Martinez, S., Rivas Saenz, S., Penas, A., 2011. Worldwide bioclimatic classification system. Global Geobotany 1, 1-634.

Smith, P.F., Ganesh, S., Liu, P., 2013. A comparison of random forest regression and multiple linear regression for prediction in neuroscience. Journal of Neuroscience Methods 220(1), 85-91.

Strobl, C., Malley, J., Tutz, G., 2009. An Introduction to Recursive Partitioning: Rationale, Application, and Characteristics of Classification and Regression Trees, Bagging, and Random Forests. Psychological Methods 14(4), 323-348.

van Engelen, V.W.P., Verdoodt, A., Dijkshoorn, K., van Ranst, E., 2006. Soil and terrain database of Central Africa - DR of Congo, Burundi and Rwanda (SOTERCAF, version 1.0), Report 2006/07, ISRIC - World Soil Information. 28 pp. Available through: www.isric.org.

VSN International (2013). GenStat for Windows 16th Edition. VSN International, Hemel Hempstead, UK. GenStat.co.uk.

Webster, R., Oliver, M.A., 2007. Geostatistics for environmental scientists. Statistics in practice. Second edition ed. John Wiley & Sons, Chichester

Appendix 1. Estimated components of variance and results from testing fixed effects

The estimated components of variance σ^2_{farm} and σ^2_{plot} from the fitted mixed models are shown in Table 16.

Table 16. *Estimated components of variance σ^2_{farm} and σ^2_{plot} from mixed model analysis on the yields for beans, maize and potato.*

	Beans		Maize		Potato	
Model	σ^2_{farm}	σ^2_{plot}	σ^2_{farm}	σ^2_{plot}	σ^2_{farm}	σ^2_{plot}
Treatment	*0.15(0.03)*	*0.14(0.01)*	*0.31(0.06)*	*0.15(0.01)*	*2.99(0.80)*	*2.86(0.36)*
Soil Type + Treatment + Interaction	*0.14(0.03)*	*0.14(0.01)*	*0.33(0.07)*	*0.15(0.01)*	*3.13(0.85)*	*2.65(0.34)*
Climate	*0.14(0.03)*	*0.14(0.01)*	*0.28(0.06)*	*0.16(0.01)*	*2.84(0.80)*	*2.56(0.34)*
Altitude + Treatment + Interaction	*0.15(0.03)*	*0.14(0.01)*	*0.31(0.06)*	*0.15(0.01)*	*3.08(0.84)*	*2.87(0.37)*
Rainfall + Treatment + Interaction	*0.15(0.03)*	*0.14(0.01)*	*0.32(0.07)*	*0.15(0.01)*	*2.63(0.73)*	*2.67(0.34)*
Temperature + Treatment + Interaction	*0.15(0.03)*	*0.14(0.01)*	*0.29(0.06)*	*0.15(0.01)*	*2.64(0.74)*	*2.71(0.35)*
Night Temperature + Treatment + Interaction	*0.14(0.03)*	*0.14(0.01)*	*0.32(0.07)*	*0.15(0.01)*	*2.75(0.76)*	*2.64(0.34)*
Day Temperature + Treatment + Interaction	*0.15(0.03)*	*0.14(0.01)*	*0.32(0.07)*	*0.15(0.01)*	*3.10(0.85)*	*2.94(0.38)*
Dolomite + Treatment + Interaction	*0.07(0.02)*	*0.11(0.01)*	*0.14(0.03)*	*0.14(0.01)*	*2.99(0.82)*	*2.90(0.37)*
PC1 + PC2 + PC3 from soil Nutrients + Treatment + Interaction	*0.13(0.03)*	*0.14(0.01)*	*0.25(0.05)*	*0.14(0.01)*	*3.07(0.86)*	*3.02(0.39)*

Standard errors of the estimated components are between parentheses. The yellow highlighted cells show a substantial reduction of the size of the variance components compared to those in the model with treatment effects only. Calculated *p*-values from the *F*-tests for the Wald statistics for the fixed effects in the models are shown in Tables 17-19. Bold type indicates a significant effect at the 5% level.

Table 17. *Beans: p-values from F-tests for Wald statistics of fixed effects*

Beans: model	Cofactor effect	Treatment effect	Cofactor * treatment
Treatment	-	**<0.001**	-
Treatment + Soil Type + Interaction	0.200	**<0.001**	0.582
Treatment + Climate + Interaction	0.144	**<0.001**	0.644
Treatment + Altitude	0.857	**<0.001**	0.936
Treatment + Rainfall + Interaction	0.483	**<0.001**	0.175
Treatment + Temperature + Interaction	0.895	**<0.001**	0.706
Treatment + Day Temperature + Interaction	0.854	**<0.001**	0.409
Treatment + Night Temperature + Interaction)	0.062	**<0.001**	0.476
Treatment + Dolomite + Interaction	**<0.001**	**<0.001**	**<0.001**
Treatment + PC1 + PC2 + PC3 + Interaction			
-PC1	**0.021**		**0.054**
-PC2	0.169		0.560
-PC3	0.147	<0.001	0.076

Table 18. Maize: p-values from F-tests for Wald statistics of fixed effect

Maize: model	Cofactor	treatment	Cofactor * treatment
Treatment	-	**<0.001**	-
Treatment + Soil Type + Interaction	0.828	**<0.001**	0.676
Treatment + Climate + Interaction	0.092	**<0.001**	0.995
Treatment + Altitude + Interaction	0.307	**<0.001**	0.244
Treatment + Rainfall + Interaction	0.942	**<0.001**	0.096
Treatment +Temperature + Interaction	0.039	**<0.001**	0.562
Treatment + Day Temperature + Interaction	0.949	**<0.001**	0.826
Treatment + Night Temperature + Interaction	0.613	**<0.001**	0.140
Treatment + Dolomite + Interaction	**<0.001**	**<0.001**	**0.002**
Treatment + PC1 + PC2 +PC3 + Interaction			
-PC1	**<0.001**		**<0.001**
-PC2	0.596		0.612
-PC3	0.103	**<0.001**	0.731

Table 19. Potato: p-values from F-tests for Wald statistics of fixed effect

Potato: model	Cofactor	treatment	Cofactor * treatment
Treatment	-	**<0.001**	-
Treatment + Soil Type + Interaction	0.579	**<0.001**	**0.033**
Treatment + Climate + Interaction	0.200	**<0.001**	**0.022**
Treatment + Altitude + Interaction	0.636	**<0.001**	0.520
Treatment + Rainfall + Interaction	**0.034**	**<0.001**	**0.025**
Treatment + Temperature + Interaction	**0.040**	**<0.001**	0.053
Treatment + Day Temperature + Interaction	0.786	**<0.001**	0.880
Treatment + Night Temperature + Interaction	**0.068**	**<0.001**	**0.014**
Treatment + Dolomite + Interaction	0.362	**<0.001**	0.795
Treatment + PC1 +PC2 +PC3 + Interaction			
-PC1	0.442		0.835
-PC2	0.863		0.836
-PC3	0.319	**<0.001**	0.938

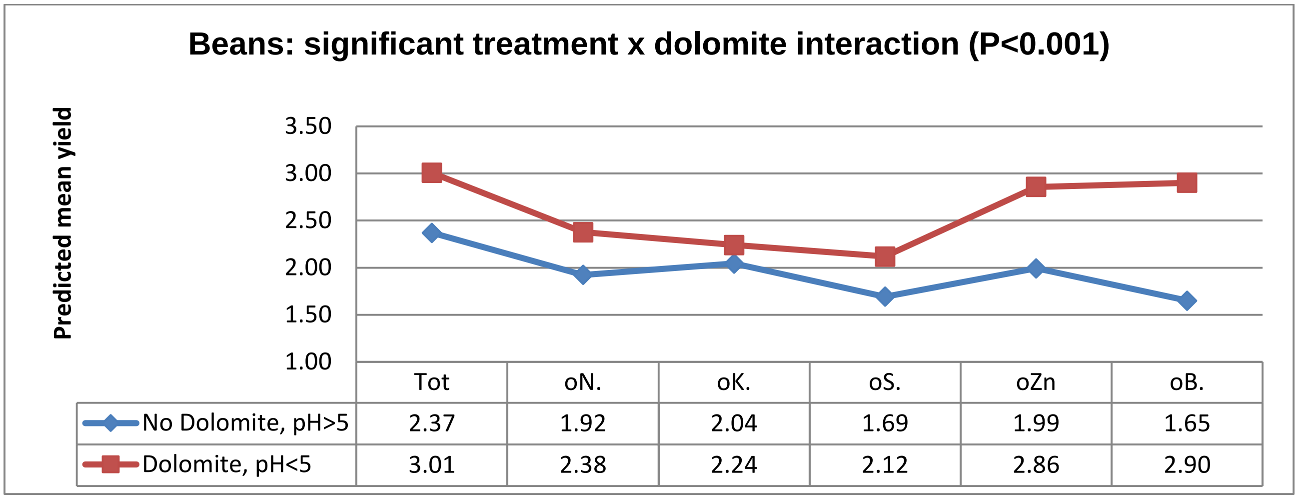
Conclusion: the results from the tests of fixed effect show:

- no significant effects or interaction for soil type or climate class, neither for the covariates for the yield data for beans and maize;
- significant interactions between the scores of PC1 and between dolomite and treatment for the yield data for beans and maize;
- significant interactions with treatment for soil type, climate class, rainfall, temperature and night temperature for the yield data from potato.

Appendix 2. Profile plots of predicted mean yields and results from testing pairwise differences

Significant effects for mean yield for beans.

Profile plots and diagrams with the results from testing pairwise differences between predicted mean yields are presented below.



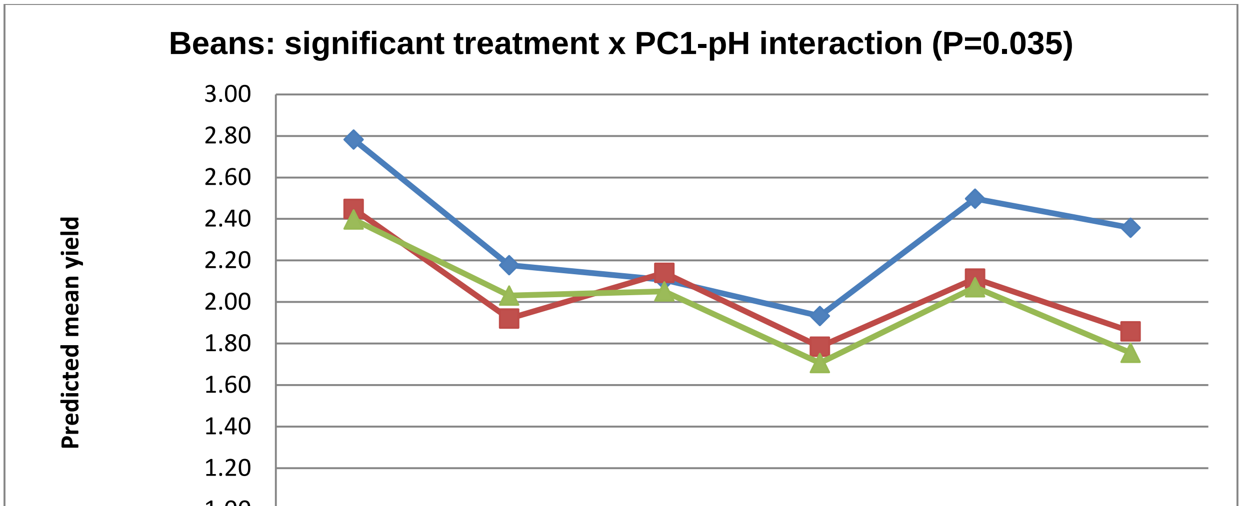
	Tot	oN.	oK.	oS.	oZn	oB.
No Dolomite, pH>5	2.37	1.92	2.04	1.69	1.99	1.65
Dolomite, pH<5	3.01	2.38	2.24	2.12	2.86	2.90

Beans: Pairwise testing: homogeneous groups, P=0.05. Treatments with the same letter are not significantly different from each other, but are significantly different from the other treatments. For example, the yields as a result of reducing N, omitting K, omitting Zn and omiting B when dolomite is given are not significantly different from each other, but are significantly different from the control ('Tot'), S omission, B omission and all dolomite treatments except S omission.

```
0 Tot   2.370   . . . d .
0 oN.   1.923   . b . . .
0 oK.   2.044   . b c . .
0 oS.   1.692   a . . . .
0 oZn   1.993   . b c . .
0 oB.   1.650   a . . . .
1 Tot   3.007   . . . . e
1 oN.   2.377   . . . d .
1 oK.   2.241   . . c d .
1 oS.   2.118   . b c . .
1 oZn   2.858   . . . . e
1 oB.   2.901   . . . . e
```

Conclusion: the test results in the diagram show:

- a significant negative effect on mean yield for all omission treatments except for omission of Zn and B when dolomite was supplied.



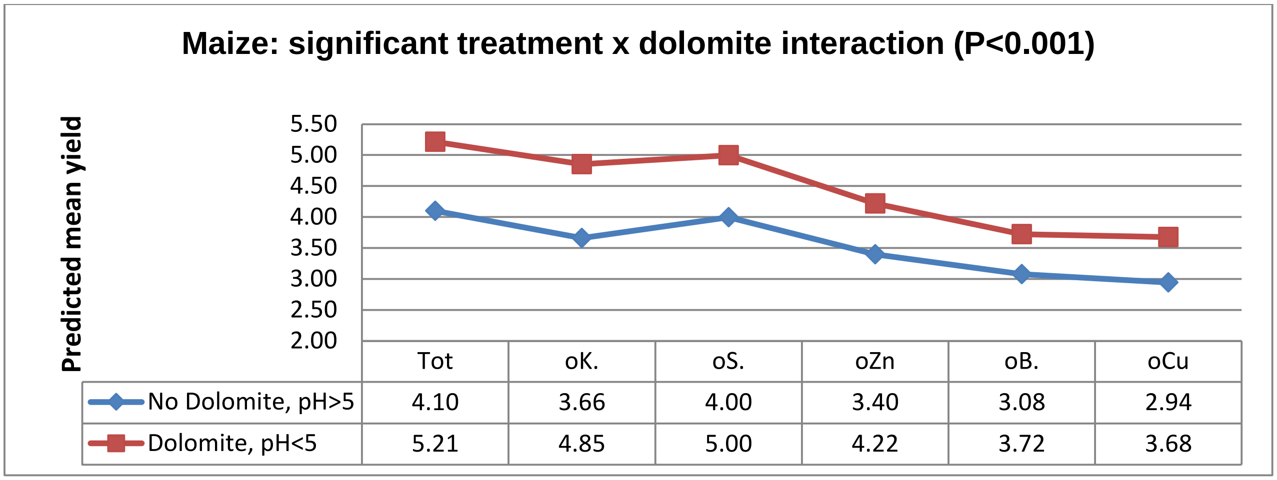
	Tot	oN.	oK.	oS.	oZn	oB.
Low:-2.11	2.78	2.18	2.11	1.93	2.50	2.36
Medium:-0.39	2.45	1.92	2.14	1.79	2.11	1.86
High:2.47	2.40	2.03	2.05	1.71	2.07	1.76

```
Pairwise testing: homogeneous groups , P=0.05

    -2.112 Tot    2.783    . . . . . . . h
    -2.112 oN.    2.178    . . c d e f . .
    -2.112 oK.    2.107    . b c d e . . .
    -2.112 oS.    1.933    a b c . . . . .
    -2.112 oZn    2.498    . . . . . . g .
    -2.112 oB.    2.357    . . . d e f g .
   -0.3925 Tot    2.448    . . . . . f g .
   -0.3925 oN.    1.921    a b c . . . . .
   -0.3925 oK.    2.141    . . c d e . . .
   -0.3925 oS.    1.786    a b . . . . . .
   -0.3925 oZn    2.113    . . c d e . . .
   -0.3925 oB.    1.859    a b c . . . . .
     2.466 Tot    2.397    . . . . e f g .
     2.466 oN.    2.032    . b c . . . . .
     2.466 oK.    2.051    . b c d . . . .
     2.466 oS.    1.707    a . . . . . . .
     2.466 oZn    2.072    . b c d . . . .
     2.466 oB.    1.756    a . . . . . . .
```

Conclusion: the test results show a significant negative effect on mean yield for beans for all nutrition omission treatments for the low, medium and high category of the first component PC1.

Significant effects for mean yield for maize.

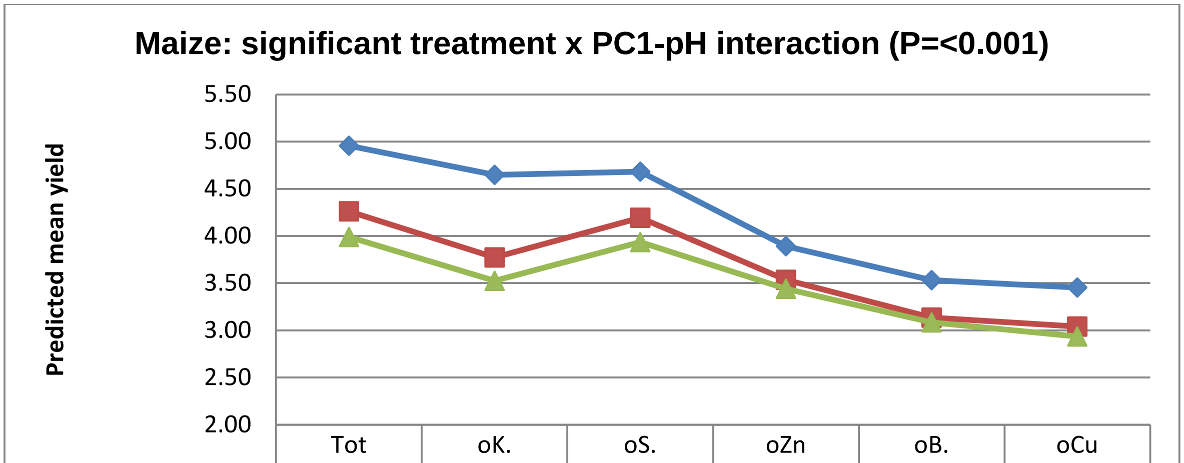



```
Maize: Pairwise testing: homogeneous groups, P=0.05

   0 Tot    4.099    . . . . e . .
   0 oK.    3.659    . b c . . . .
   0 oS.    3.996    . . . d e . .
   0 oZn    3.399    . b . . . . .
   0 oB.    3.075    a . . . . . .
   0 oCu    2.944    a . . . . . .
   1 Tot    5.214    . . . . . . g
   1 oK.    4.854    . . . . . f .
   1 oS.    4.999    . . . . . f g
   1 oZn    4.220    . . . . e . .
   1 oB.    3.724    . . c d . . .
   1 oCu    3.678    . b c . . . .
```

Conclusion: the test results show

- a negative effect on mean yield for all nutrient omission treatments. The reduction in mean yield however is not significant when S is omitted from the control treatment `Tot'.



	Tot	oK.	oS.	oZn	oB.	oCu
Low:-2.32	4.96	4.65	4.68	3.89	3.53	3.45
Medium:-0.53	4.26	3.77	4.19	3.54	3.13	3.04
High: 2.8	3.99	3.52	3.93	3.44	3.08	2.94

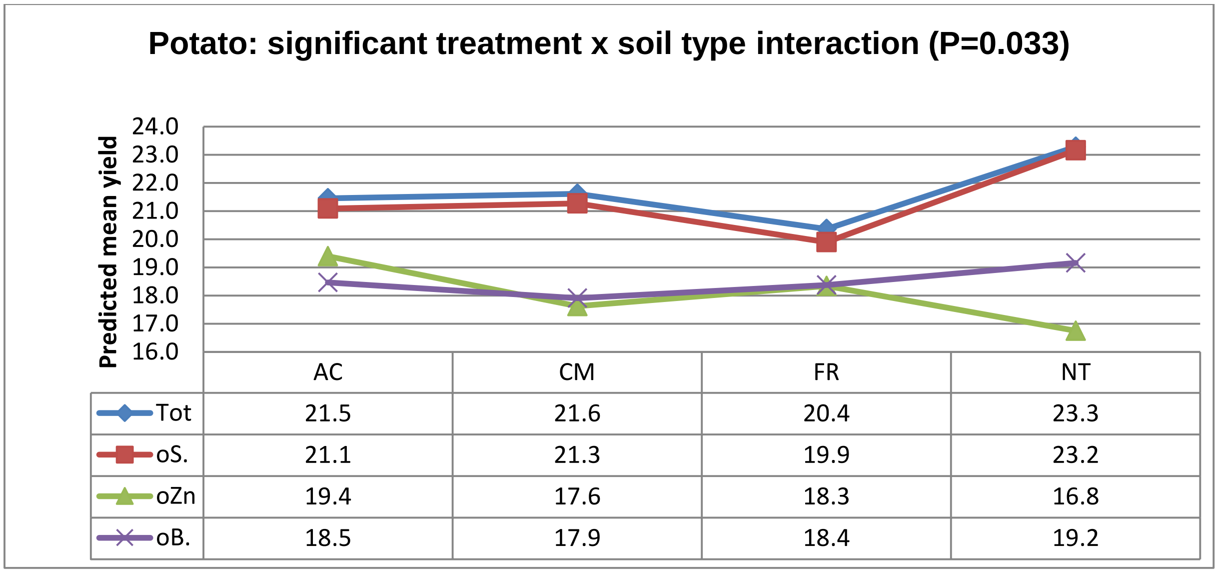
```
Pairwise testing: homogeneous groups, P=0.05

    -2.323 Tot    4.957    . . . . . . . h
    -2.323 oK.    4.647    . . . . . . g .
    -2.323 oS.    4.682    . . . . . . g .
    -2.323 oZn    3.891    . . . d e f . .
    -2.323 oB.    3.532    . . c d . . . .
    -2.323 oCu    3.454    . b c . . . . .
   -0.5261 Tot    4.260    . . . . . f . .
   -0.5261 oK.    3.774    . . c d e . . .
   -0.5261 oS.    4.194    . . . . . f . .
   -0.5261 oZn    3.537    . . c d . . . .
   -0.5261 oB.    3.134    a b . . . . . .
   -0.5261 oCu    3.040    a . . . . . . .
     2.797 Tot    3.989    . . . . e f . .
     2.797 oK.    3.524    . . c d . . . .
     2.797 oS.    3.934    . . . . e f . .
     2.797 oZn    3.440    . b c . . . . .
     2.797 oB.    3.084    a b . . . . . .
     2.797 oCu    2.936    a . . . . . . .
```

Conclusion: the test results show a significant reduction in mean yield for maize for all nutrition omission treatments except for omission of S in the medium and high PC1 category.

Mixed model analysis showed a significant larger mean maize yield at the higher temperature category compared to the medium and low temperature category. For the remaining cofactors neither main effects nor interaction with treatment was found significant.

Significant effects on mean yield for potato.



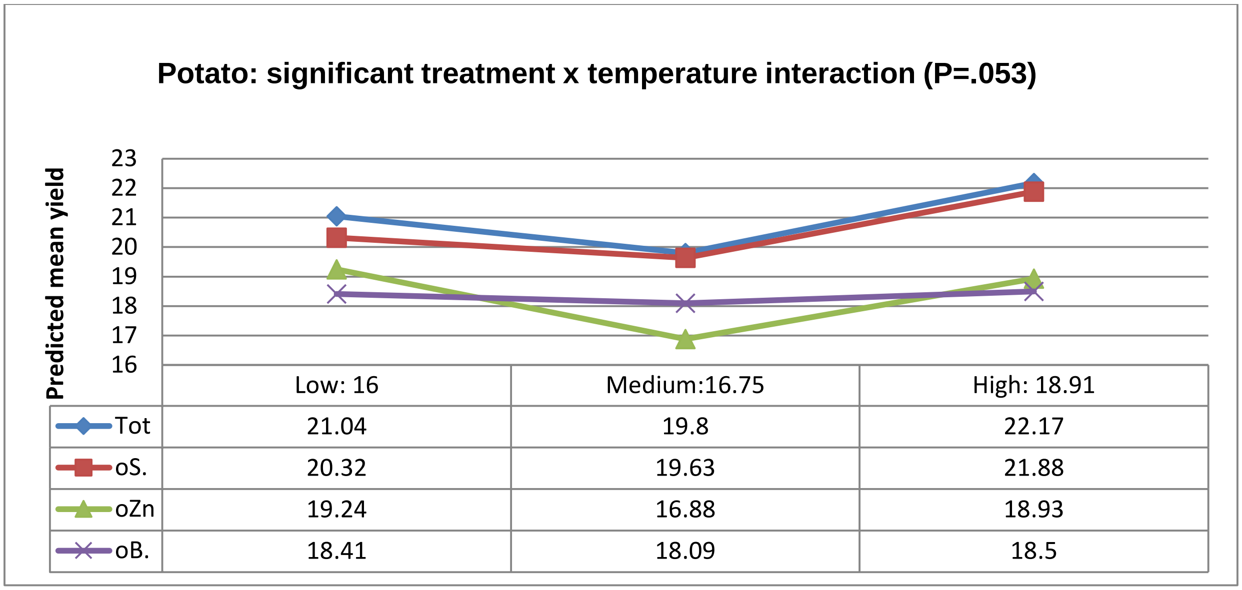
	AC	CM	FR	NT
Tot	21.5	21.6	20.4	23.3
oS.	21.1	21.3	19.9	23.2
oZn	19.4	17.6	18.3	16.8
oB.	18.5	17.9	18.4	19.2

```
Pairwise testing: homogeneous groups, P=0.05

   AC Tot    21.46    . b c
   AC oS.    21.10    . b c
   AC oZn    19.40    a b .
   AC oB.    18.47    a . .
   CM Tot    21.62    . b c
   CM oS.    21.28    . b c
   CM oZn    17.63    a . .
   CM oB.    17.91    a . .
   FR Tot    20.37    . b c
   FR oS.    19.90    a b c
   FR oZn    18.34    a . .
   FR oB.    18.38    a . .
   NT Tot    23.29    . . c
   NT oS.    23.16    . . c
   NT oZn    16.75    a . .
   NT oB.    19.16    a b .
```

Conclusion: the test results show:

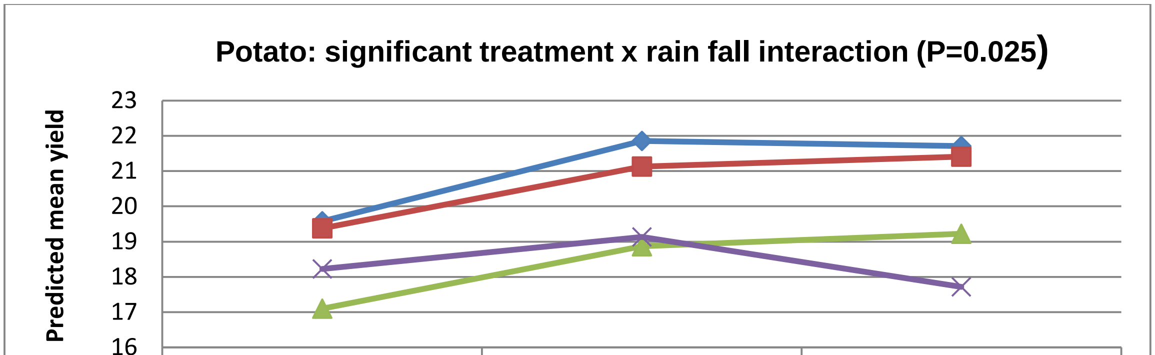
- a small not significant reduction in mean potato yield when nutrient S is omitted;
- a substantial significant reduction in mean potato yield when nutrient Zn or B is omitted, except for soil type AC where the negative effect on yield is not significant when Zn is omitted.




```
Pairwise testing: homogeneous groups, P=0.05

      16 Tot    21.04   . . . . e f g
      16 oS.    20.32   . . . d e f .
      16 oZn    19.24   . b c d . . .
      16 oB.    18.41   a b c . . . .
   16.75 Tot    19.80   . . c d e . .
   16.75 oS.    19.63   . . c d e . .
   16.75 oZn    16.88   a . . . . . .
   16.75 oB.    18.09   a b . . . . .
   18.91 Tot    22.17   . . . . . . g
   18.91 oS.    21.88   . . . . . f g
   18.91 oZn    18.93   . b c d . . .
   18.91 oB.    18.50   a b c . . . .
```

Conclusion: the test results show a significant reduction in mean potato yield for the omission fertilizer treatments oZn and oB but not for oS.

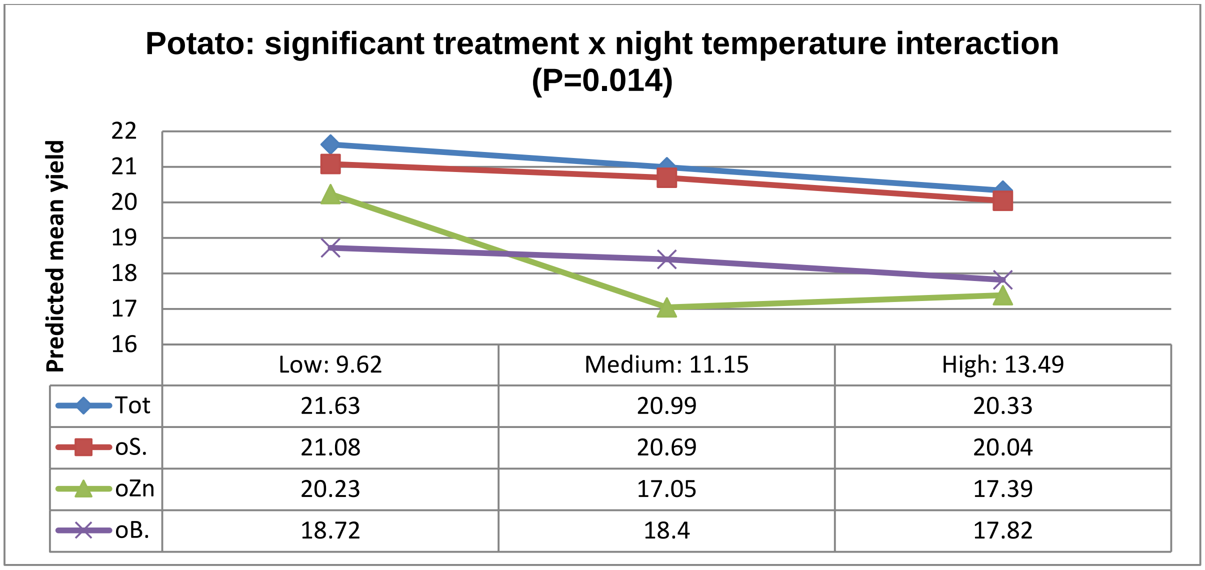


	Low:1340	Medium:1443	High: 1567
Tot	19.58	21.85	21.71
oS.	19.38	21.13	21.41
oZn	17.1	18.87	19.22
oB.	18.23	19.13	17.72

```
Pairwise testing: homogeneous groups, P=0.05

   1340 Tot   19.58   . . . d e .
   1340 oS.   19.38   . . c d . .
   1340 oZn   17.10   a . . . . .
   1340 oB.   18.23   a b c . . .
   1443 Tot   21.85   . . . . . f
   1443 oS.   21.13   . . . . e f
   1443 oZn   18.87   . b c d . .
   1443 oB.   19.13   . b c d . .
   1567 Tot   21.71   . . . . . f
   1567 oS.   21.41   . . . . . f
   1567 oZn   19.22   . . c d . .
   1567 oB.   17.72   a b . . . .
```

Conclusion: the test results in the diagram show a significant reduction in mean yield for the omission fertilizer treatments oZn and oB but not for oS.

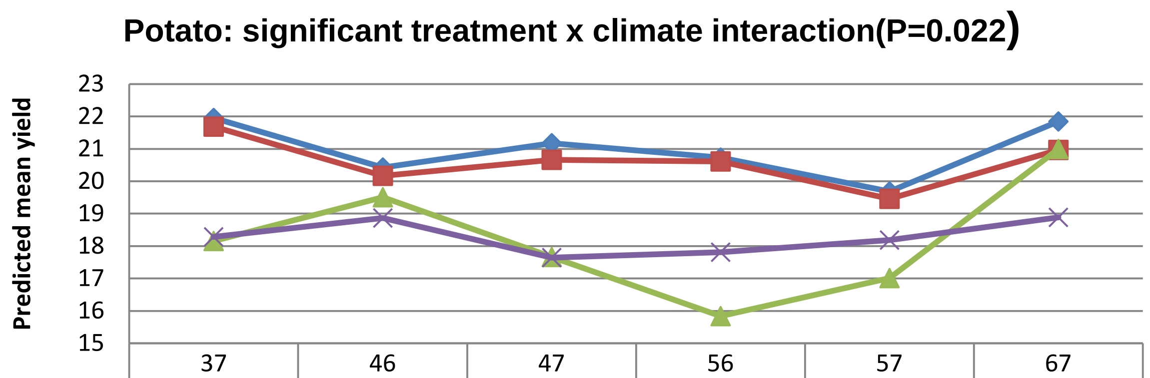



```
Pairwise testing: homogeneous groups, P=0.05

   Tot 9.624   21.63   . . . d
   Tot 11.15   20.99   . . . d
   Tot 13.49   20.33   . . c d
   oS. 9.624   21.08   . . . d
   oS. 11.15   20.69   . . . d
   oS. 13.49   20.04   . b c d
   oZn 9.624   20.23   . . c d
   oZn 11.15   17.05   a . . .
   oZn 13.49   17.39   a . . .
   oB. 9.624   18.72   a b c .
   oB. 11.15   18.40   a b . .
   oB. 13.49   17.82   a . . .
```

Conclusion: the test results show:

- no significant reduction in mean yield for the oS treatment;
- for the fertilizer omission treatment oB a significant reduction;
- for the fertilizer omission treatment oZn the reduction is significant for all night temperature categories except for the low night temperature category.



	37	46	47	56	57	67
Tot	21.95	20.42	21.17	20.73	19.69	21.84
oS.	21.69	20.17	20.66	20.61	19.46	20.97
oZn	18.16	19.51	17.66	15.84	17.01	21
oB.	18.28	18.87	17.64	17.81	18.19	18.89

```
Pairwise testing: homogeneous groups, P=0.05

   Tot 37    21.95    . . . d
   Tot 46    20.42    . b c d
   Tot 47    21.17    . . c d
   Tot 56    20.73    . b c d
   Tot 57    19.69    . b c d
   Tot 67    21.84    . . . d
   oS. 37    21.69    . . . d
   oS. 46    20.17    . b c d
   oS. 47    20.66    . b c d
   oS. 56    20.61    . b c d
   oS. 57    19.46    a b c .
   oS. 67    20.97    . . c d
   oZn 37    18.16    a b c .
   oZn 46    19.51    a b c d
   oZn 47    17.66    a b . .
   oZn 56    15.84    a . . .
   oZn 57    17.01    a . . .
   oZn 67    21.00    . . c d
   oB. 37    18.28    a b c .
   oB. 46    18.87    a b c .
   oB. 47    17.64    a b . .
   oB. 56    17.81    a b c .
   oB. 57    18.19    a b c .
   oB. 67    18.89    a b c .
```

Conclusion: the test results for climate classes show:

- no significant reduction in mean potato yield for the oS treatment;
- a significant substantial negative effect for climate class 37, 47, 56 and 57; when Zn is omitted;
- a significant substantial negative effect for climate class 37, 47 and 67 when B is omitted.

Appendix 3. Thresholds from calculated low, medium and high categories for the regional covariates

For the covariates altitude, rainfall, temperature, night temperature and day temperature the minimum, maximum, the 35 and 65 percentiles are shown in Table 20. The 3 categories used in the mixed model analysis are: Low with values below 35 percentile, Medium with values between 35 and 65 percentile and High with values large than the 65 percentile. The mean values for the Low, Medium and High category are presented as well. For the cofactors soil type and climate the categories are shown in Table 21.

Table 20. *Minimum, maximum, 35 and 65 percentiles*

Altitude	**Percentile**	**Beans**	**Maize**	**Potato**	**B mean**	**M mean**	**P mean**
	Minimum	772	828	1240	1265	1413	1728
	0.35	1514	1584	1932	1622	1654	2014
	0.65	1692	1762	2127	1840	1905	2200
	Maximum	2250	2282	2354			
Rainfall	**Percentile**	**Beans**	**Maize**	**Potato**	**B mean**	**M mean**	**P mean**
	Minimum	924	888	1208	1088	1081	1340
	0.35	1183	1172	1402	1217	1220	1443
	0.65	1244	1267	1487	1375	1360	1567
	Maximum	1656	1710	1725			
Temperature	**Percentile**	**Beans**	**Maize**	**Potato**	**B mean**	**M mean**	**P mean**
	Minimum	15.40	16.30	14.35	17.67	17.36	16
	0.35	18.98	18.70	16.49	19.22	18.96	16.75
	0.65	19.74	19.12	17.23	21.53	20.71	18.91
	Maximum	25.40	25.65	21.35			
Night temperature	**Percentile**	**Beans**	**Maize**	**Potato**	**B mean**	**M mean**	**P mean**
	Minimum	17.40	21.85	21.65	26.02	25.95	22.75
	0.35	27.78	27.68	24.09	28.37	28.26	25.27
	0.65	28.82	28.70	25.90	30.29	29.78	26.91
	Maximum	34.05	32.55	32.45			
Day temperature	**Percentile**	**Beans**	**Maize**	**Potato**	**B mean**	**M mean**	**P mean**
	Minimum	11.40	8.30	8.50	12.61	12.01	9.62
	0.35	13.54	12.9	10.5	14.23	13.66	11.15
	0.65	14.72	14.22	11.8	16.25	15.51	13.49
	Maximum	19.10	18.70	15.0			

Table 21. Categories for the cofactors dolomite, soil type and climate type.

Cofactor	Categories
Dolomite	0=no if pH≥5 and 1=yes if pH <5
Soil type: beans, maize, potato	37,46,47,56,57,67
Climate: beans	AC, CM, FR, GL, - , NT
Climate: maize	AC, CM, FR, GL, LP, NT
Climate : potato	AC, CM, FR, -, -, NT

Appendix 4. Pairwise correlations and calculated loading of first principal component PC1

The calculated correlations for beans, maize and potato are shown in Figures 28, 29 and 30. To highlight substantial correlation diagrams with correlations of 0.50 or more are presented as well.

Beans:

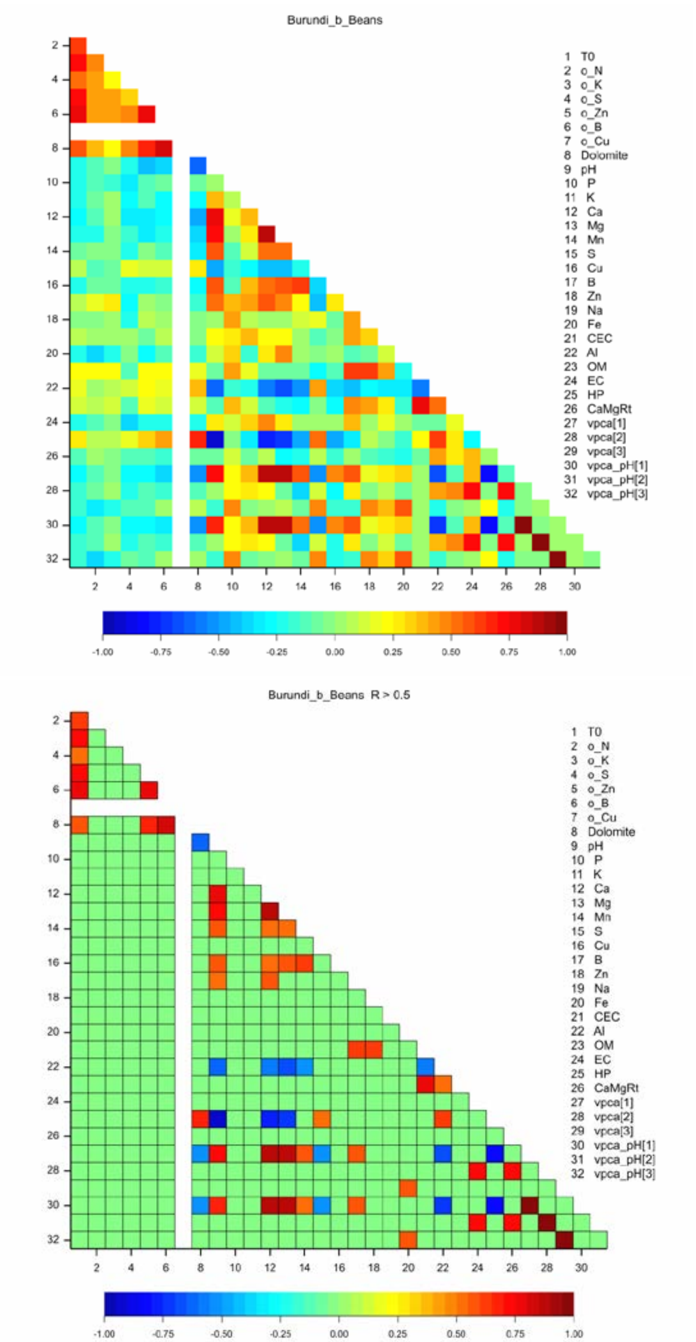


Figure 28. Pairwise correlations for yield and soil data for beans.

Maize:

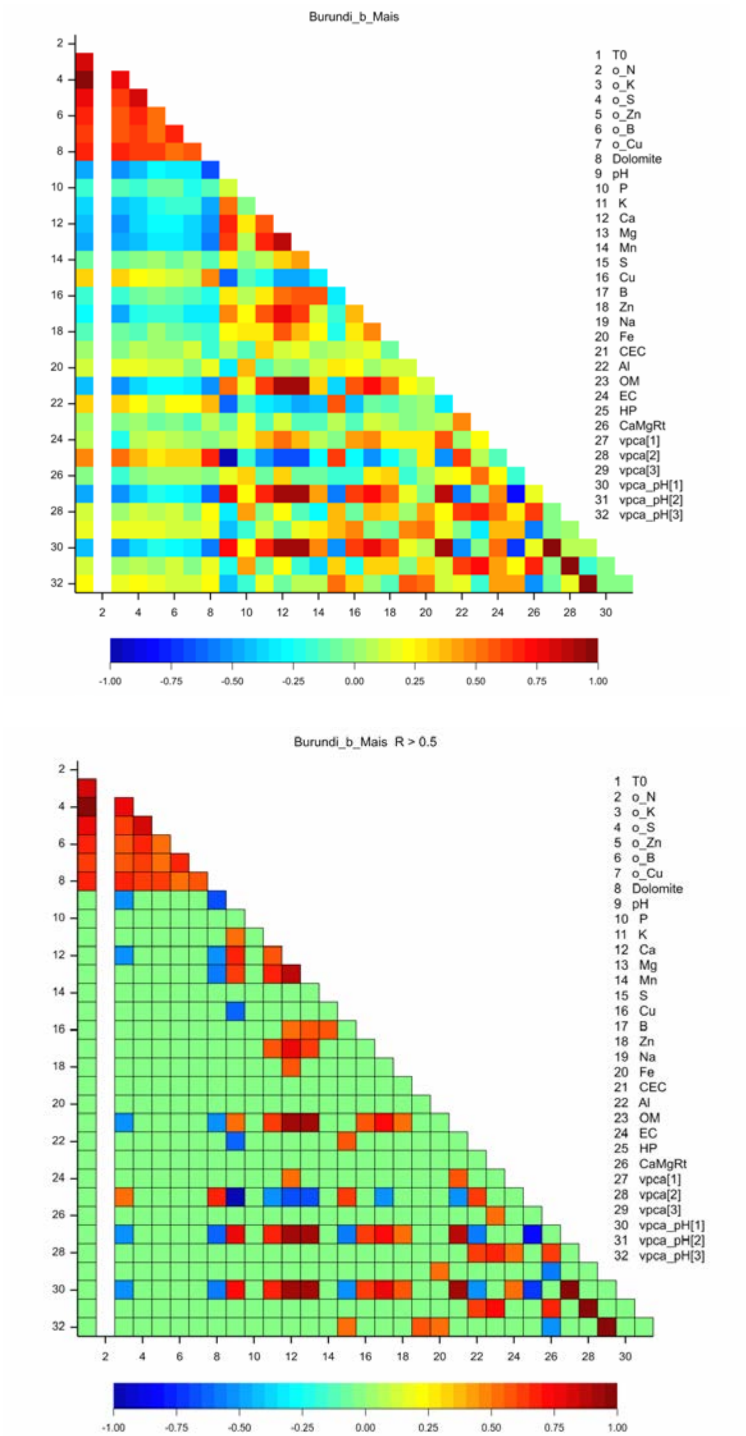
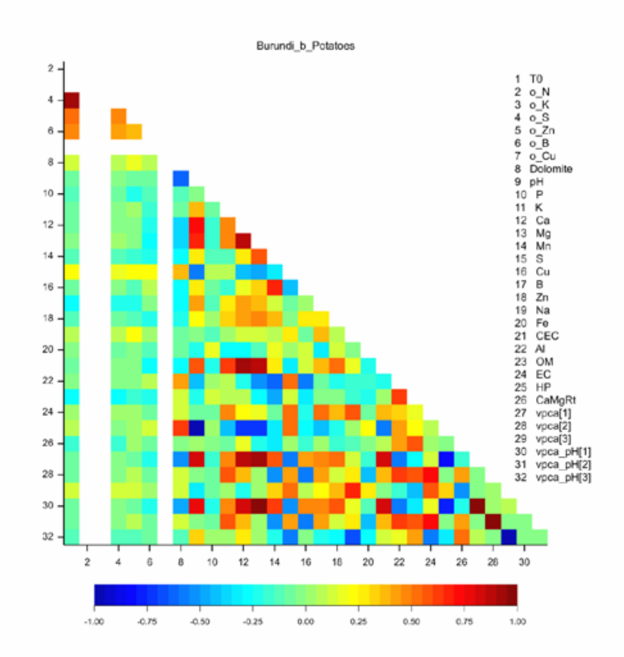


Figure 29. Pairwise correlations for yield and soil data for maize.

Potato:



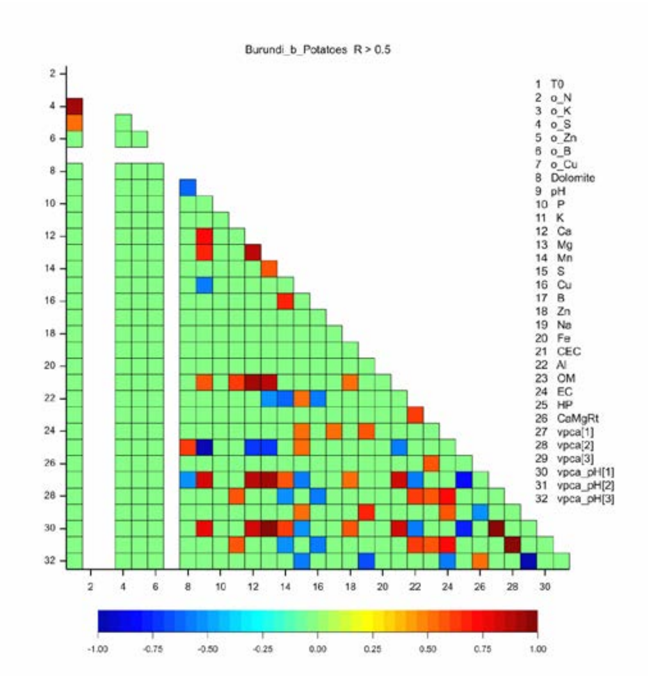


Figure 30. *Pairwise correlations for yield and soil data for potato.*

The correlation diagrams show that: pairwise correlation between crop yield and the soil parameters is low for beans and potato. High correlations are found between crop yields for maize when nutrient K is omitted with pH, Ca, CEC and HP. The calculated loadings (weights) of the soil nutrient parameters in the first principal component and the percentage of the variation in the data are shown in Table 22.

Table 22. Calculated loadings from the first component.

Soil Nutrient Parameter	Beans	Maize	Potato
P	0.10	0.10	-0.09
K	0.17	0.27	0.20
Ca	0.38	0.39	0.37
Mg	0.37	0.37	0.42
Mn	0.23	0.18	0.27
S	-0.21	-0.22	-0.23
Cu	0.26	0.26	0.21
B	0.30	0.30	0.21
Zn	0.08	0.23	0.25
Na	0.14	0.08	0.08
Fe	0.14	0.02	-0.06
CEC	0.38	0.38	0.37
Al	-0.29	-0.21	-0.26
OM	-0.09	-0.01	-0.14
EC	0.15	0.22	0.08
HP	-0.34	-0.30	-0.34
CaMgRt	-0.06	0.07	-0.09
% variation accounted	37	36	31

Cells with loadings in absolute value larger than 0.20 have been coloured yellow for an impression of the contrast the component is referring to. For beans, the first component refers to the contrast between the set: {Ca, Mg, Mn, Cu, B, CEC} with positive loadings versus the set {S, Al, HP} with negative loadings. For maize, the first component refers to the contrast between the set {K, Ca, Mg, Mn, Cu, B, Zn, CEC, EC} and the set {S, Al, HP}. For potato, the first component refers to the contrast between the set {K, Ca, Mg, Mn, Cu, B, Zn, CEC} and the set {S, Al, HP}.

For a visual impression of a possible geographical clustering of the scores from PC1 the appearance of the low, medium and high categories is shown in the Figure 31.

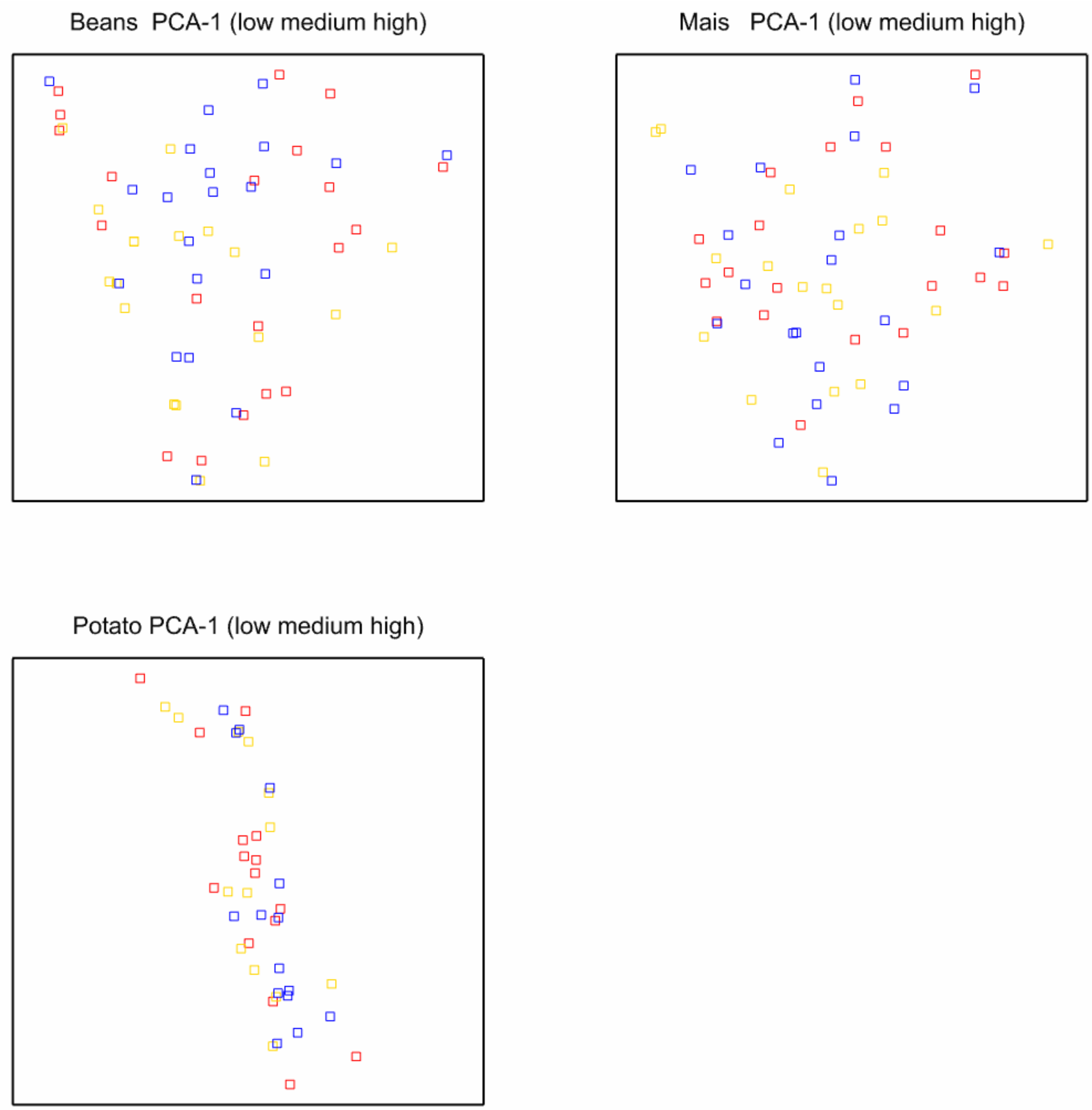


Figure 31. *Geographical layout of the low (blue), medium (yellow) and high (red) categories of pc1.*

VFRC is a semi-autonomous unit of the International Fertilizer Development Center (IFDC).
For more information visit www.ifdc.org.