

**THE ROLE OF TOPOGRAPHY AND COVER CROPS IN MICHIGAN
AGRICULTURAL ECOSYSTEMS AND ITS POTENTIAL EFFECT UNDER
FUTURE CLIMATE SCENARIOS**

By

Juan David Munoz-Robayo

A DISSERTATION

**Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of**

Crop and Soil Sciences - Doctor of Philosophy

2014

ABSTRACT

THE ROLE OF TOPOGRAPHY AND COVER CROPS IN MICHIGAN AGRICULTURAL ECOSYSTEMS AND ITS POTENTIAL EFFECT UNDER FUTURE CLIMATE SCENARIOS

By

Juan David Munoz-Robayo

The use of cover crops is reported to enhance agro-ecological services in rotational crop systems, however their adoption by farmers has remained limited. A challenge to farmer uptake is high spatial and temporal variability in cover crop establishment and growth. Since the benefits of cover crop use are a function of the amount of cover crop biomass that enters the soil, it is important to quantify cover crop biomass production across the field. The ability to easily and inexpensively quantify the spatial variability of cover crop biomass is needed to better understand and predict its potential as an input to agricultural systems. My study demonstrated that hierarchical nonlinear models can adequately predict biomass of the cover crop from the easily measured Normalized Difference Vegetation Index (NDVI) data thus providing a relatively inexpensive method of obtaining dense cover crop biomass measurements.

Topography plays an important role in spatial processes that ultimately affect plant performance, it could be used to quantify and predict cover crop spatial variability and cover crop contribution to a subsequent cash crop. However, my results show that the utility of topographical information in plant performance predictions depends on the analysis scale. I explored the relationships between cover crop biomass and topography at different scales of derivation, and demonstrated that neighborhood size greatly affects the strength of the prediction performance in multiple regression models. Equipped with resulting information on the optimal analyses scales I then studied the effects of topography and cover crop biomass on corn yields. Topographic attributes were found to contribute significantly to

explaining the variability in both red clover biomass and corn yields and red clover biomass positively influenced corn yield, however, the magnitude of that effect varied both temporally and spatially. Resulting better understanding of how variations in topography affect cover crops and row-crops will contribute to increased cover crop adoption, allowing producers to tailor management to site-specific features of their fields.

The combined effect of topography and cover crops in row-crops can be implemented in a process-based crop simulation model (SALUS) to predict the performance of conventional and cover crop-enhanced managements under future climate scenarios. Projections of crop performance under 100 years of future climate scenarios showed a significant decline in corn yields, in particular for the organic-based treatment. On the other hand, soybean and wheat yields showed a slight increment. These simulation outcomes then can be used to identify potential strategies for climate change mitigation.

Este trabajo está dedicado a mi familia, Aura Cely, Jorge Adelmo y Omar Felipe, por su apoyo incondicional.

This work is dedicated to my family, Aura Cely, Jorge Adelmo, and Omar Felipe, for their unconditional support.

ACKNOWLEDGMENTS

I want to express my sincere gratitude to my advisor Dr. Alexandra (Sasha) Kravchenko for her valuable advice and guidance. Her advice had a profound effect in my education at MSU. With Sasha I have learned very important skills in the areas of applied statistics and soil science. However, this is just a fraction of the priceless lessons I have received from her, that are not limited to the technical aspects of research. It was a pleasure to work with her; I will keep her advice for the years to come. Thank you Sasha for sharing your wise thoughts and opportune comments!!

I would like to thank the members of my PhD. committee Dr. Bruno Basso, Dr. Andrew Finley, Dr. Sieg Snapp, and Dr. Jiaguo Qi for their guidance, suggestions, and valuable comments in my research. Thanks to Dr. Basso for his valuable lessons and insightful discussions about crop modeling, and also for his opportune instructions in the use of SALUS. Thanks to Dr. Finley for teaching me and introducing me to the interesting topics of hierarchical models and Bayesian statistics. Thanks to Dr. Snapp for her suggestions during the execution of my experiments and also for sharing her thoughts on agroecology. Thanks to Dr. Qi for teaching me the theory and the application of remote sensing, I have enjoyed his class and I will benefit from his lectures in my future career.

I would also like to thank Jim Bronson and Suzanne Sippel from the W.K. Kellogg Biological Station for their assistance in the LTER experiments. Thanks to Brian Graff and Tim Boring from the Agriculture Farm, for their help in data collection and sample processing. Also, special thanks to Brian Baer for his very helpful instructions with the use of the software SALUS. Thanks to Dr. Juan Pedro Steibel for his thoughts and advice in the use of WinBUGS to fit path analysis. Thanks to Dr. Jeff Andresen to provide me with the future climate scenarios that I used in the crop simulations.

This work could not be done with the help of many people, specially my lab mates. Thanks to Wei Wang, Kateryna Ananyeva, Moslem Ladoni, Andrew Worth, Melissa Erickson, and Richard Price. I expend valuable time with these colleagues, thanks for the friendship, support and motivation.

My deepest gratitude to my parents, Aura and Jorge, and my brother Felipe, for their support during all these years.

TABLE OF CONTENTS

LIST OF TABLES	ix
LIST OF FIGURES	x
Chapter 1: Introduction	1
REFERENCES	6
Chapter 2: Predicting cover crop biomass using Normalized Difference Vegetation	
Index and nonlinear hierarchical models	10
Abstract	10
Keywords	10
2.1. Introduction	11
2.1.1. Importance of cover crops in agricultural systems	11
2.1.2. Methods for quantifying biomass	11
2.1.3. Study objectives	14
2.2. Materials and methods	14
2.2.1. Data	14
2.2.2. Data pooling by field	16
2.2.3. Hierarchical models	18
2.2.4. Bayesian implementation	21
2.2.5. Prediction	21
2.2.6. Model Selection	22
2.2.7. Software implementation	23
2.3. Results and discussion	23
2.4. Conclusion	33
REFERENCES	34
Chapter 3: Deriving the optimal scale for relating topographic attributes and cover	
crop plant biomass	37
Abstract	37
Keywords	38
3.1. Introduction	38
3.2. Materials and methods	42
3.2.1. Study site	42
3.2.2. Data collection and pre-processing	45
3.2.3. Statistical analysis	47
3.3. Results and discussion	49
3.3.1. Effect of scaling on the topographic factors	49
3.3.2. Effect of scaling on the sensor-based variable (NDVI)	54
3.3.3. Relationship between topographic factors and NDVI at different scales	56
3.3.4. Indicators of optimal topographic scale	63
3.4. Conclusions	69
REFERENCES	71

Chapter 4: Cover crop effect on corn growth and yield in agricultural fields with diverse terrain	75
Abstract	75
Keywords.....	76
4.1. Introduction	76
4.2. Objectives	81
4.3. Materials and methods.....	81
4.3.1. Study site	81
4.3.2. Data collection	82
4.3.3. Landscape classification	86
4.3.4. Statistical Analysis-Hierarchical path analysis	87
4.3.5. Statistical Analysis-Analysis of covariance	94
4.4. Results and discussion.....	94
4.4.1. Relationship between topography, red clover biomass and corn yield	94
4.4.2. Effect of red clover biomass on corn yield at different topographical positions	103
4.5. Conclusions	108
APPENDIX	110
REFERENCES	112
Chapter 5: Modeling performance of row crop systems with and without cover crops across topographically diverse terrain in future climate	119
Abstract	119
5.1. Introduction	120
5.2. Objectives	123
5.3. Materials and Methods	124
5.3.1. Study site and management systems.....	124
5.3.2. Landscape classification and experimental plots	125
5.3.3. Field data for model validation	128
5.3.3.1. Soil data	128
5.3.3.2. Crop data	129
5.3.3.3. Weather data	130
5.3.4. Future climate scenarios	134
5.3.5. Crop model and simulation experiments	136
5.3.6. Statistical analysis	137
5.4. Results	141
5.4.1. Model validation-measures of error	141
5.4.2. Model validation-ANOVA.....	144
5.4.3. Simulation with future climate scenarios.....	149
5.4.3.1. Projections of crop yield	149
5.4.3.2. Comparison of crop yields in the climate period 3 (2070-2099).....	158
5.5. Conclusions	161
REFERENCES	162
Chapter 6: Conclusions	166

LIST OF TABLES

Table 2.1. Parameter credible intervals for each candidate model. Parameter estimates expressed as 2.5%50% 97.5% percentiles.	25
Table 2.2. Comparison of model complexity and fit using the effective number of parameters (P_D) and Deviance Information Criterion (DIC).	26
Table 3.1. List of the fields used in the study along with the year when <i>NDVI</i> data were collected from the field, field size, and the range of elevation values.	45
Table 3.2. List of topographical variables that were statistically significant predictors ($p < 0.05$) for <i>NDVI</i> in multiple regression analysis for each neighborhood size and each studied field.	58
Table 4.1. List of the fields used in the study, along with years when the fields were sampled, their agricultural management treatments (T3=Low input and T4=Organic), field sizes, and the numbers of samples collected per field.	83
Table 4.2. Summary of the statistical models used to analyze the relationships between corn, red clover cover crops and topographical variables.	90
Table 4.3. Fit statistics for each studied model. MSE: mean squared error. p_D : penalized deviance. DIC: deviance information criteria.	97
Table 4.4. Parameter estimates and credible intervals for the model with varying intercepts per field and varying slopes per year (M-4). Estimates expressed as 2.5%50%97.5% percentiles.	100
Table 5.1. Number of fields and experimental plots used in the validation of the crop model. Numbers of plots are presented per agricultural management system and topographical position.	128
Table 5.2. Relevant model parameters used as input in SALUS during model validation. Units are presented in parenthesis.	140
Table 5.3. Measures of model error for the initial and final simulations presented by crop.	142
Table 5.4. ANOVA F-value and associated probability ($Pr > F$) for the statistical term in the linear mixed model. F-values are presented for the models with measured yield and simulated yields respectively.	145
Table 5.5. ANOVA F-value and associated probability ($Pr > F$) for the linear mixed model on simulated future yields.	158

LIST OF FIGURES

Figure 2.1. Model 2b residuals versus NDVI from the population fitted values.....	13
Figure 2.2: a) NDVI versus biomass by sample date. b) NDVI versus biomass by field.....	17
Figure 2.3: a) Nonlinear models for NDVI-biomass relationship. Mean of the posterior fitted value distribution for the exponential Model 1a (solid line) and Richards Model 1b (dashed line). b) Mean of the posterior fitted values field specific random effects (dashed lines) and population (solid line) from Richards model (Model 2b).....	27
Figure 2.4: a) Upper and lower 95% credible interval for the population’s posterior predictive distribution (dashed lines) along with the associated mean (solid line) for Model 2b. b) Model 2b, constant variance model, mean residual variance (solid points) calculated within fixed-width NDVI intervals (vertical dashed lines) with fitted curve from the Model 3b variance function. c) Model 3b upper and lower 95% credible interval for the population’s posterior predictive distribution (dashed lines) along with the associated mean (solid line).	28
Figure 2.5. Spatial interpolation of observed NDVI and predicted biomass for Field-98 on August 9. Mean and standard deviation a) and b) respectively, of the biomass predictive distribution estimated with Richards mean function and constant variance (Model 2b). Mean and standard deviation c) and d) respectively, of the biomass predictive distribution estimated with Richards mean function and functional variance (Model 4b).....	32
Figure 3.1. Location of the Kellogg Biological Station (KBS) in Southwest Michigan. Studied fields are marked by red outlines; field ID labels are shown within each field.....	43
Figure 3.2. NDVI values of the studied fields plotted over fields’ digital elevation models (DEMs). The DEMs show the landscape configuration of each field, whereas NDVI represents the spatial variability of cover crop biomass. The map resolution is 1×1 m.	44
Figure 3.3. Histograms of the original topographic features and NDVI values overlaid with their probability density functions at three studied neighborhood sizes (original, 5 m and 40 m). A) Relative elevation. B) Slope. C) Square root of the flow length. D) Potential solar radiation index. E) Curvature. F) NDVI. G) Log of flow accumulation. H) Illustration of the effect of the neighborhood size on the flow accumulation maps for original (ORI), 5 and 40 m neighborhood sizes in a portion of field 792. Values in the map are the numbers of contributing pixels.	51
Figure 3.4. Mean correlation coefficients between the studied variable values obtained at different neighborhood sizes and the original layer. Error bars indicate the standard deviation across the fields. A) Variables with minor decrease in the correlation coefficient (i.e. relative elevation, slope, solar radiation and NDVI). B) Variables with major decrease in the correlation coefficient (i.e. flow accumulation, flow length, curvature and wetness index)...	55
Figure 3.5. Adjusted-R ² for multiple regression models. (A) Adjusted-R ² as a function of neighborhood size in each individual field. B) Mean adjusted-R ² for multiple regression from all studied fields as a function of neighborhood size. Different uppercase letters mark	

significant differences between topographic scales ($p < 0.05$). Different lowercase letters mark significant differences between NDVI scales ($p < 0.05$). Values presented in (A) are for the cases where NDVI and topography were filtered at the same neighborhood sizes. Only significant predictors were used in the multiple regression models ($\alpha = 0.05$). 61

Figure 3.6. Indexes for selecting optimum neighborhood size. A) Linear regression between multiple regression adjusted-R² and the standardized ranges of solar radiation, relative elevation and slope in the studied fields. B) Linear regression between the change in adjusted-R² after filtering and semivariogram nugget/sill ratios from NDVI maps of the studied fields. C) Linear regression between the change in adjusted-R² after filtering and the semivariogram nugget/sill ratio from slope maps of the studied fields. D) Optimal neighborhood size for topography plotted versus the partial sill value from semivariograms of the slope maps. 66

Figure 3.7. Correlation between score values of the first principal component (PC1) and the mean adjusted-R² values from multiple regression analyses. Black squares: Before filtering. Fields with relatively flat topography had positive values of PC1, while fields with more diverse topography had negative values of PC1. Red circles: After filtering. Adjusted-R² values from multiple regressions in the original layers were plotted as a reference with their respective change values (White squares). Numbers indicate the field labels. 67

Figure 4.1. Example of topographical and yield information available at two of the studied fields (i.e. field-97 on the left and field-38 on the right). A) Map showing field topography classified as depression (red), slope (green), summit (blue) positions; vertical gradient corresponds to field elevation. B) Sampling points where the cover crop biomass was collected before plowing down in along with map of corn yield in 2008. 84

Figure 4.2. Path diagram for the relationship between topographical attributes, red clover biomass and corn yield. A) Basic model structure (M-1). B) Hierarchical model structure with varying slopes per year (M-4). 88

Figure 4.3. Cumulative precipitation in KBS for the years 2008, 2009 and 2011. The black line is the 24-year cumulative precipitation computed as the daily average from 1988 to 2011. Vertical gray lines are approximate day of planting, end of vegetative stage and harvesting of corn. Bars in the bottom of the figure represent values of daily precipitation. 99

Figure 4.4. Red clover biomass during the growing season. Biomass values at 3 landscape positions combined across studied years (A). Biomass values in different years combined across topographical positions (B). Error bars represent standard errors of the mean. Bars with different letter within the same sampling date indicate significant differences at $p < 0.05$ 105

Figure 4.5. Effect of red clover biomass on corn traits at three different landscape positions. Relationship between corn biomass at stage V4 and red clover biomass (A). Relationship between corn yield and red clover biomass (B). Comparisons between landscape positions were performed at the lower quartile, median, and upper quartile of the covariate (Vertical gray lines). Different letters within the same value of red clover biomass indicate statistically significant difference between topographical positions at $p < 0.05$ 106

Figure 5.1. Location of the experimental fields at KBS (Top). Example of the topographical classification for 9 fields (Bottom). Black dots indicate the location of experimental plots for soil and yield data collection..... 127

Figure 5.2. Distribution of soil texture in the first soil depth (0-20 cm) across all experimental plots. Texture triangle is resented by topographical position: Summit, Slope, and Depression. 131

Figure 5.3. Soil properties used as input in the crop simulation model presented by soil depth: a) sand, silt, clay content, and organic carbon, b) bulk density, and water limits. Median values from all experimental plots are presented by topographical position. Medial value (line) is bounded by the 25th and 75th percentiles (shaded area). DUL is drained upper limit. 132

Figure 5.4. Projection of year mean temperature and precipitation in the future climate scenarios A2 and B2 for the weather station in Greenville, Michigan. 135

Figure 5.5. Measured yield vs. simulated yield from the final validated model for: Corn, Soybean, and Wheat. The 1:1 line is presented in red. Dot color represents the agricultural management treatment: T1 (Red), T3 (Green), and T4 (Blue). 143

Figure 5.6. Simulated yield and measured yield by topographic position and management treatment for: a) corn, b) soybean, and c) wheat. Mean separation based on Fisher’s LSD ($\alpha=0.05$). Upper-case letters compare management treatments within a particular topographic position. Lower-case letter compare topographic positions within a particular management treatment. 146

Figure 5.7. SALUS projection of crop yields for a) corn, b) soybean, and c) wheat under future climate scenarios. Topographic position is color-coded: depression (green), summit (blue), and slope (red). Projection of mean yield in management treatments are plotted with different line type: continuous (T1), dash (T3), and dotted (T4). 152

Figure 5.8. Relevant stress factors by management treatment, position and climate period in a) nitrogen stress in corn, b) drought stress in soybean, and c) nitrogen stress in wheat. Mean stress values with different letters within same treatment and position indicate differences between climate periods after Tukey’s HSD ($\alpha=0.05$). Stress factors range from 0 to 1, where 0 indicates the maximum stress and 1 indicates no stress. 155

Figure 5.9. Future simulated yields by topographic position and management treatment in a) corn, b) soybean, and c) wheat. Mean separation is based on Tukey’s HSD ($\alpha=0.05$). Upper-case letters compare management treatments within a particular topographic position. Lower-case letter compare topographic positions within a particular management treatment. 160

Chapter 1: Introduction

Agricultural ecosystems use substantial amounts of natural resources to supply food, fiber, and fuel to the growing global population. Agricultural ecosystems are managed to optimize the yield productivity which is demanded by the markets. However, the intensive use of resources in agricultural systems has raised concerns regarding its sustainability (Vitousek et al., 1997; Millennium Ecosystem Assessment, 2005; Zhang et al., 2007). For example, it has been shown that intensive cultivation tends to deteriorate the biophysical capacity of agricultural ecosystems by reducing soil fertility and water quality; and increasing soil erosion and nutrient losses (Isse et al., 1999; Dinesh et al., 2001; Snapp et al., 2010). Degradation of water quality and soil resources are becoming urgent problems because they reduce the potential to achieve optimal yields and the ability to provide other ecosystem services. Maintaining ecosystem services in agricultural systems is of major relevance especially in upcoming years when facing future climate changes (IPCC, 1995). Examples of important ecosystem services that could help mitigate the negative impact caused by climate change are soil carbon sequestration and water quality regulation (IPCC, 1995; Lal, 2004).

The use of cover crops in rotation with the main row crop has been proposed as a solution for maintaining and even enhancing ecosystem services (Lal, 2004; Snapp et al., 2005). The benefits of using cover crops are multiple; it has been shown that cover crops increase carbon sequestration, water retention and soil fertility, while reducing soil erosion (Kerr et al., 2007; Bhardwaj et al., 2011; Fageria et al., 2011; Dabney et al., 2012). Among available cover crops, legume covers are considered the most appropriate plants to enhance soil productivity. For example, rotation with leguminous cover crops has been shown to

increase availability of soil nitrogen thus lessening the need for chemical fertilizers (Meisinger et al. 1991; Miguez and Bollero, 2005). Leguminous cover crops also have a positive impact on soil organic matter and thus increasing soil carbon sequestration (Dinesh et al. 2001; Fageria et al., 2011). Red clover (*Trifolium pratense* L.) is among the most commonly used legume cover crop in the northeastern United States (Singer and Cox, 1998). Red clover is recognized as one of the most effective species in producing substantial amounts of residues under Michigan conditions (Snapp et al., 2005); and has been shown to have beneficial effects not only in soil fertility, but also on soil porosity and soil structure (Papadopoulos et al., 2006).

Although many benefits can be derived from legume cover crop use, the adoption by Midwest producers has been relatively slow (Snapp et al., 2005; Knowler and Bradshaw, 2007). One of the main reasons of this slow adoption by farmers is the high spatial and temporal variability in the growth and performance of legume cover crops. For example, high spatial and temporal variability increases management challenges of cover crop-based systems. In addition, the expected contribution of cover crops to row crops becomes more difficult to assess because the potential nutrient supply to the main crops varies in space and time as well (Nykanen et al., 2008; Hauggaard-Nielsen et al., 2010).

In order to increase cover crop adoption we first must understand the causes of its variability. Several works have shown that topography is a key driver in many biophysical processes across the landscape. For example, topography has been shown to affect the spatial distribution of soil properties and water redistribution (Boyer et al., 1996, Green and Erskine, 2004). In addition, the role that topographic attributes play on soil dynamics substantially contributes to the spatial distribution of yields in row crops (Kravchenko et al., 2000; Kravchenko et al., 2005). Therefore, I consider that topographic attributes can be used to explain the spatial variability in cover crop performance. A better understanding of how

variations in topography and soil properties affect cover crop performance would contribute to more effective cover crop implementation allowing producers to tailor cover crop recommendations to site-specific features of their fields.

However, before the main drivers of this spatial and temporal variability in cover crops can be investigated, the variability in production of cover crop biomass across the landscape must be quantified first. Unfortunately, not much information is available regarding the total biomass production and its spatial variability in agricultural fields; and assessment of cover crop biomass is a time- and a labor-consuming data collection process. As a consequence our knowledge of the factors affecting cover crop performance in row crops still remains limited. Use of non-destructive methods in cover crop biomass assessment can greatly decrease data collection costs and time. For example, remote sensing techniques have been successfully used for plant biomass estimation on grazing rangelands and in row crops (Tood et al., 1998; Flynn et al., 2008). Use of these methods can greatly benefit further cover crop research by producing continuous spatial estimations of biomass across the landscape. However, there is no standard procedure during the processes of filtering, computing, and analyzing remote sensing data. A first challenge in this work was the generation of continuous maps that accurately represent the spatial variability in cover crop biomass. Since the scale of representation influence the strength of the relationship between topography and cover crop biomass (Roecker and Thompson, 2010), an immediate second challenge was to find the optimal scale of derivation and representation of these continuous maps so that the relationships between topography and cover crop biomass are the most favorable. The meaningful relationships that we could discover between cover crop biomass, topographic attributes, and row-crop yields will depend on the optimal quantification and representation of the spatial information.

A particular challenge in studying the role of topography in agricultural ecosystems with cover crops lays in the fact that the effects of topographic attributes on row crop yields can be both direct and indirect. For example, topography has a direct effect on main crop performance by controlling soil and water dynamic across the landscape, but also an indirect effect by controlling the spatial performance of cover crops biomass. In practice it is difficult to separate the direct effect of topography on crop yields from the effect of cover crop on crop yields. In addition, the magnitudes of these direct and indirect topographical influences on main row crops have not been addressed before. To assess direct and indirect effects of topography on row crop yields I made use of Path Analysis (PA), a procedure that allows testing direct and indirect relationships between topography, red clover and corn yield (Hoyle, 1995). With PA I studied the contribution of cover crop biomass to the main crop while accounting for confounding effects of topographic attributes.

The main hypothesis in this study is that use of legume cover crops in agricultural ecosystems help to enhance or maintain ecosystems services while maintaining the row crop productivity. Therefore I consider that legume cover crops could be a promising strategy to mitigate the negative effect of climate change. However, almost no testing has been carried out on the efficacy of this mitigation strategy. The complex interactions between multiple factors that influence agricultural ecosystem make particularly difficult the evaluation of mitigation strategies. Crop simulation models are a relevant tool to predict and understand the consequences of weather and management perturbations to biophysical systems (Rosenzweig and Parry, 1994; White et al., 2008). Therefore, simulation models could play an important role in the evaluation of mitigation strategies for agricultural production under changing climate (Mall et al., 2004). Evaluation of the role of topography and cover crops through simulation could help farmers, resource managers, and decision-makers to be better informed on the potential mitigation strategies for agricultural production in future years.

Unfortunately, crop simulation models were developed to execute specific tasks; therefore most of the available models are not capable to integrate topographic information, rotation systems, and cover crop use. System Approach to Land Use Sustainability (SALUS) is a program designed to model continuous crop, soil, water and nutrient conditions under different management strategies for multiple years (Basso and Ritchie, 2012). I used SALUS to integrate the factors of relevance in my study and reproduce the actual performance of conventional and organic agricultural ecosystems. I expanded the use of these crop simulations by using future climate scenarios to predict the potential role of topography and cover crops under climate change.

The overall goals of my study are to examine the effect of topographic attributes and cover crop biomass in row crop agricultural ecosystems; and to assess its potential role under future climate scenarios. The dissertation is organized as follows: In Chapter 2, I presented a methodology to predict cover crop biomass using Normalized Difference Vegetation Index. Chapter 3 highlights the importance of deriving the optimal scale for relating topographic attributes and cover crop plant biomass. Chapter 4 studies the effect of cover crop biomass on corn growth and yield in agricultural fields with diverse terrain. Chapter 5 presents the results of modeling crop rotations under future climate scenarios at different management systems and landscape positions. Chapter 6 states the most relevant conclusions from this study.

REFERENCES

REFERENCES

- Basso, B. and Ritchie, J., 2012. Assessing the impact of management strategies on water use efficiency using soil-plant-atmosphere models. *Vadose Zone J.* 11(3)
- Bhardwaj, A., Jasrotia, P., Hamilton, S., and Robertson, G. 2011. Ecological management of intensively cropped agro-ecosystems improves soil quality with sustained productivity. *Agri. Ecos. Environ.* 140: 419-429.
- Boyer, D., Wright, R., Feldhake, C. and Bligh, D., 1996. Soil spatial variability relationships in a steeply sloping acid soil environment. *Soil Sci.*, 161(5): 278-287.
- Dabney, S., Delgado, J., and Reeves, D. 2001. Using winter cover crops to improve soil and water quality. *Comm. Soil Sci. Plant Anal.* 32: 1221-1250.
- Dinesh, R., Suryanarayana, M., Nair, A., Ghoshal, S. 2001. Leguminous cover crop effects on nitrogen mineralization rates and kinetics in soils. *J. Agron. Crop Sci.* 187: 161-166
- Fageria, N., Baligar, V., and Bailey, B. 2005. Role of cover crops in improving soil and row crop productivity. *Comm. Soil Sci. Plant Anal.*, 36: 19, 2733-2757
- Flynn, E., Dougherty, C. and Wendroth, O., 2008. Assessment of pasture biomass with the normalized difference vegetation index from active ground-based sensors. *Agron. J.*, 100: 114-121.
- Green, T. and Erskine, R., 2004. Measurement, scaling, and topographic analyses of spatial crop yield and soil water content. *Hydrol. Processes*, 18: 1447-1465.
- Hauggaard-Nielsen, H., Holdensen, L., Wulfsohn, D., and Jensen, E. 2010. Spatial variation of N₂-fixation in field pea (*Pisum sativum* L.) at the field scale determined by the 15N natural abundance method. *Plant Soil.* 327:167-184
- Hoyle, R. 1995. The structural equation modeling approach: Basic concepts and fundamental issues. In *Structural equation modeling: Concepts, issues, and applications*, R. Hoyle (Ed). Thousand Oaks, CA: Sage Publications, Inc., pp. 1-15.
- Isse, A., MacKenzie, A., Stewart, K., Cloutier, D., Smith, D. 1999. Cover crops and nutrient retention for subsequent sweet corn production. *Agron. J.* 91:934-939
- IPCC, 1995. *Climate change 1995. Working Group 1. Inter-Government Panel on Climate Change*, University Press, U. K., pp. 879.
- Kerr, R., Snapp, S., Chirwa, M., Shumba, L. and Msachi, R., 2007. Participatory research on legume diversification with Malawian smallholder farmers for improved human nutrition and soil fertility. *Expl Agric.*, 43: 437-453.
- Knowler, K., Bradshaw, B. 2007. Farmers' adoption of conservation agriculture: A review and synthesis of recent research. *Food Policy* 32: 25-48

- Kravchenko, A. and Bullock, D., 2000. Correlation of corn and soybean grain yield with topography and soil properties. *Agron. J.*, 92: 75-83.
- Kravchenko, A., Robertson, G., Thelen, K. and Harwood, R., 2005. Management, topographical, and weather effects on spatial variability of crop grain yields. *Agron. J.*, 97: 514-523.
- Lal, R. 2004. Soil carbon sequestration impacts on global climate change and food security. *Science*. 304: 1623-1627.
- Mall, R., Lal, M., Bhatia, v., Rathore, L., Singh, R. 2004. Mitigating climate change impact on soybean productivity in India: a simulation study. *Agri. Forest Meteor.* 121 : 113–125
- Meisinger, J., Hargrove, W., Mikkelsen, R., Williams, J., and Benson, V. 1991. Effects of cover crops on groundwater quality. pp. 57–68. In: Hargrove, W. (Ed.) *Cover Crops for Clean Water*. SWCS. Ankeny, IA.
- Miguez, F., and Bollero. G. 2005. Review of corn yield response under winter cover cropping systems using meta-analytic methods. *Crop Sci.* 45:2318-2329
- Millennium Ecosystem Assessment, 2005. *Ecosystems and Human Well-being: Desertification Synthesis*. World Resources Institute, Washington, DC.
- Nykanen, A., Jauhiainen, L., Kemppainen, J., and Lindström, K. 2008. Field-scale spatial variation in yields and nitrogen fixation of clover-grass leys and in soil nutrients. *Agric. Food Sci.* 17: 376-393
- Papadopoulos, A., Mooney, S. and Bird, N., 2006. Quantification of the effects of contrasting crops in the development of soil structure: an organic conversion. *Soil Use and Management.*, 22: 172-179.
- Roecker, S., and Thompson, J. 2010. Scale Effects on Terrain Attribute Calculation and Their Use as Environmental Covariates for Digital Soil Mapping. In J.L. Boettinger et al. (eds.), *Digital Soil Mapping, Progress in Soil Science 2*. Springer.
- Rosenzweig, C., Parry, M. 1994. Potential impact of climate change on world food supply. *Nature*. 367: 133-138
- Singer, J. and Cox, W. 1998. Agronomics of corn production under different crop rotations. *J. Prod. Agri.*, 11: 462–468.
- Snapp, S., Gentry, L., Harwood, R. 2010. Management intensity – not biodiversity – the driver of ecosystem services in a long-term row crop experiment. *Agri. Ecos. Environ.* 138: 242–248
- Snapp, S., Swinton, S., Labarta, R., Mutch, D., Black, J., Leep, R., Nyiraneza, J., and O’Neil, K. 2005. Evaluating Cover Crops for Benefits, Costs and Performance within Cropping System Niches. *Agron. J.* 97:322–332
- Tood, S., Hoffer, R. and Milchunas, D., 1998. Biomass estimation on grazed and ungrazed rangelands using spectral indices. *Int. J. Remote Sensing.*, 19(3): 427-438.

Vitousek, P., Aber, J., Howarth, R., Likens, G., Matson, P., Schindler, D., Schlesinger, W., Tilman, D. 1997. Human Alteration of the Global Nitrogen Cycle: Sources and Consequences. *Ecol. Applications*, Vol. 7(3): 737-750

White, T., Johnson, I. and Snow, V., 2008. Comparison of outputs of a biophysical simulation model for pasture growth and composition with measured data under dryland and irrigated conditions in New Zealand. *Grass and Forage Science*, 63: 339–349.

Zhang, W., Ricketts, t., Kremen, C., Carney, K., Swinton, S. 2007. Ecosystem services and dis-services to agriculture. *Ecol. Economics*. 64: 253-260

Chapter 2: Predicting cover crop biomass using Normalized Difference Vegetation Index and nonlinear hierarchical models

Abstract

Incorporating cover crops into agricultural systems can improve soil structural properties, increase nutrient availability, reduce erosion and loss of agrochemicals, and suppress weeds. These benefits are a function of the amount of cover crop biomass that enters the soil. The ability to easily and inexpensively quantify the spatial variability of cover crop biomass is needed to better understand and predict its potential as an input to agricultural systems. Here, I explore the use of Normalized Difference Vegetation Index (NDVI) as a source of information for improving accuracy and precision of cover crop biomass prediction. I focus on developing models that account for biomass variability within and among fields. These models are used to produce digital data layers of predicted biomass and associated uncertainty. I propose hierarchical nonlinear models with field random effects and a residual variance function to accommodate strong heteroscedasticity. These models are motivated using aboveground biomass of red clover (*Trifolium pratense* L.) measured on three different dates in five fields in southwest Michigan. Model adequacy was assessed using the Deviance Information Criterion. Given this criterion, the “best” fitting model included field effects and a polynomial function to account for non-constant residual variance. Importantly, I demonstrate that accounting for heteroscedasticity in the model fitting is critical for capturing uncertainty in subsequent biomass prediction.

Keywords: Red clover, cover crop, NDVI, Bayesian, MCMC, heteroscedasticity.

2.1. Introduction

2.1.1. Importance of cover crops in agricultural systems

Legume cover crops are considered among the most appropriate plants to enhance soil productivity (Kerr et al., 2007). Benefits of cover crops include: increase in carbon sequestration and nutrients, reduction in erosion and loss of agrochemicals, and weed suppression (Guretzky et al., 2004; Papadopoulos et al., 2006). These benefits are a function of the amount of cover crop biomass that enters the soil. However, adoption of cover crop usage has been slow. This is primarily due to high temporal and spatial variability in the cover crop biomass that is incorporated into the soil (Snapp et al., 2005). Thus potential benefits to the main crops are also spatially and temporal variable by extension, making this activity unreliable for producers. The ability to easily and inexpensively quantify the spatial variability of cover crop biomass is needed to more fully understand and estimate its cost/benefit for agricultural systems.

2.1.2. Methods for quantifying biomass

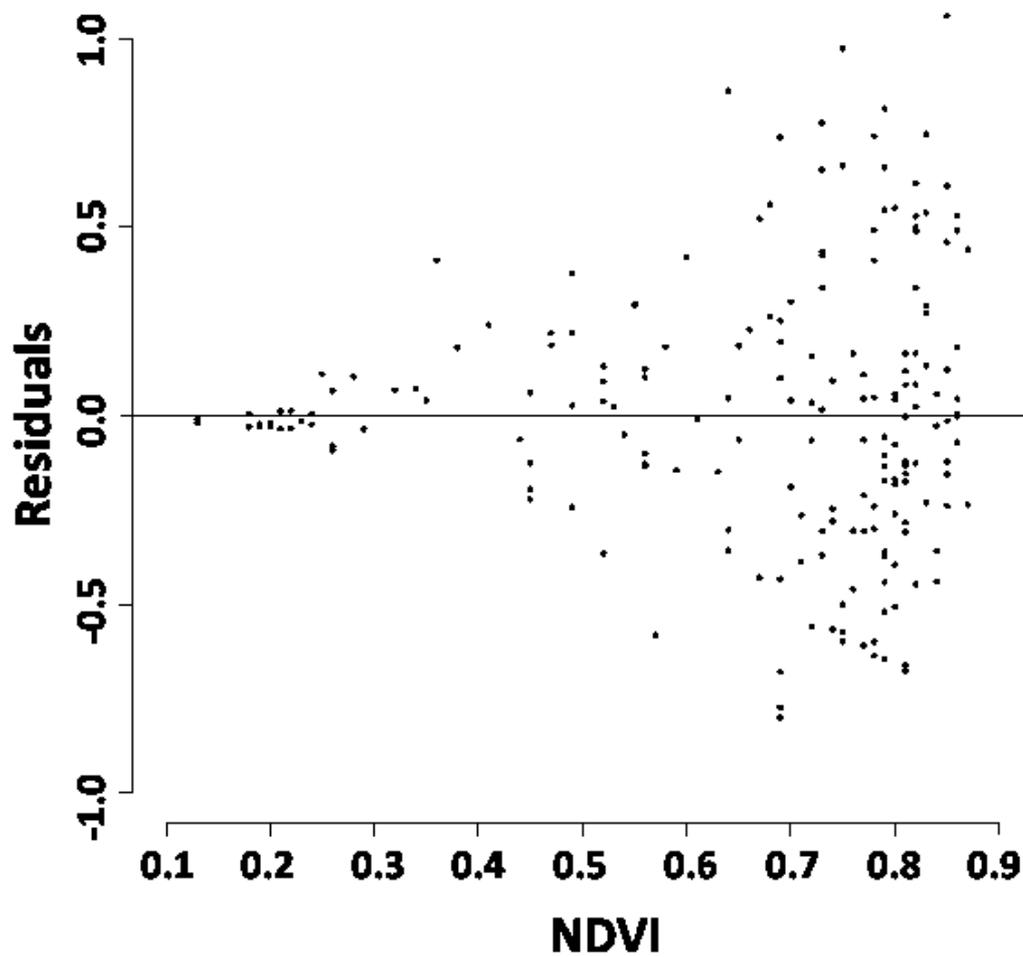
Assessment of cover crop biomass is a time- and labor-consuming process. Several authors compare the use of different techniques to measure biomass (Harmony et al., 1997; Whitbeck and Grace, 2006; Martin et al., 2005; Radloff and Mucina, 2007). Common conclusions from these studies are biomass estimates are extremely variable and measurement accuracy depends on the technique used, forage type, species, and stage of development; thus no single method is effective in all circumstances. Cover crop biomass quantification becomes even more challenging as we consider the natural spatial variability within fields. Producing spatially explicit measures of cover crop biomass over large fields using traditional measurement techniques is prohibitively expensive. Therefore, my focus is

on exploring efficient, inexpensive, and non-destructive methods that use indirect indicators to predict biomass. Remotely sensed measures of spectral reflectance and indices derived from these measures, such as Normalized Difference Vegetation Index (NDVI), have been used successfully to predict agricultural plant biomass. Several such studies use linear regression to predict biomass or similar biological parameters (Du Plessis, 1999; Calvao and Palmeirin, 2004), others use nonlinear models to explain the NDVI-biomass relationship with similar or even better results. Freeman et al. (2007) explored several common linear and nonlinear models that couple NDVI and plant height to predict biomass and forage yield. They found that an exponential nonlinear regression was useful for relating biomass to NDVI. Xu et al. (2008) compared six models to predict grass yield from NDVI for six regions in China. They compared linear, power, exponential, quadratic, cubic, and logarithmic function models, among which the exponential nonlinear offered the “best” fit based on their minimum residual error criterion. Flynn et al. (2008) also used regression models to describe the relationship between NDVI and crop dry matter.

Along with nonlinear trends in the relationships between NDVI and the biological response variable, many studies report greater variability in the response variable at higher NDVI – partially due sensor saturation (Thenkabail et al. 2000). In such cases, the characteristic “right opening megaphone” shape would be evident in plots of model residuals versus NDVI (see, e.g., Figure 2.1). This shape indicates the presence of non-constant variance, or heteroscedasticity, and suggests a possible violation of the model assumptions. When heteroscedasticity is present, inference about model parameters and the subsequent prediction could be wrong. To the best of my knowledge, heteroscedasticity has not been accounted for in biomass-NDVI models beyond simple, and often ineffective, transformation of the response variable. Even when transformation of the response variable is effective at stabilizing the variance, end-users are rarely interested in transformed expressions e.g., log or

square-root of biomass. Although, in some cases, back transformation can be used for the first moment of the predicted response (e.g., mean), it is often difficult or impractical to back transform the associated variance components.

Figure 2.1. Model 2b residuals versus NDVI from the population fitted values.



2.1.3. Study objectives

Three study objectives are addressed. First, explore the use of NDVI for predicting red clover (*Trifolium pratense* L.) biomass at a fine spatial resolution within multiple fields in southwest Michigan. Second, develop suitable hierarchical models that account for disparate sources of variability in biomass and explicitly accommodate heteroscedasticity. Third, develop digital data layers of predicted clover biomass with associated measures of uncertainty that can be subsequently used as input for growth and yield models in agricultural rotation systems.

2.2. Materials and methods

2.2.1. Data

The study was conducted at the Kellogg Biological Station (KBS) located in southwest Michigan at 42° 24' N, 85° 24' W. Annual rainfall averages 890 mm/y and mean annual temperature is 9.7 °C. The dominant soil series are Kalamazoo (fine-loamy, mixed, mesic Typic Hapludalfs) and Oshtemo (coarse-loamy, mixed, mesic Typic Hapludalfs). Data were collected from four large experimental fields (Field 30, 38, 79 and 97) of 5.9, 7.6, 5.8 and 5.4 ha in size and from 6 small experimental fields (1 ha each) at the KBS Long Term Ecological Research (LTER) site. For the purpose of this study, the small fields at the LTER site are viewed as a single field because they are located within 15 to 210 m of each other and received the same agriculture management. All studied fields were in corn-soybean-wheat rotation with leguminous cover crops and did not receive any input of commercial fertilizer. The four large fields were in organic management since 2005, the LTER fields were managed organically since 1988. Winter wheat was the main crop on the studied fields, planted in October 2006 and harvested in July 2007. Red clover was seeded in February of 2007 in existing wheat, when the ground was still frozen and finally mowed in May 2008.

Red clover biomass was destructively sampled at random locations in these fields on August 9, 2007, September 27, 2007, and on May 6, 2008. The total number of samples collected were 20, 56, 12, 42, and 68 in fields 30, 38, 79, 97 and LTER, respectively. Each sampling location was georeferenced using a global positioning system (GPS), and biomass was removed from a 0.5 x 0.5 m quadrant; all plants were cut at ground level. Biomass weight was recorded after drying the samples at 60 °C for 48 h (Corbin and VandelWulp, 2010). Just prior to biomass removal, NDVI measurements were taken at each sampling quadrant using a portable optical sensor device (The GreenSeekerTM optical sensor unit, model RT200; NTech industries, Inc., Ukia, CA, USA). The Green Seeker is an active sensor which emits and records red (656 nm ± 25 Full Width Half Magnitude) and infrared wavelength (774 nm ± 25 Full Width Half Magnitude). I used a sensing height of 1 m above the ground, which resulted in an approximate field of view of 0.93 by 61 cm at each reading. Because of the difference in support between the sensing field of view and the quadrant area, 15 equally spaced sensor readings were taken over the extent of the quadrant and then averaged to produce a single NDVI value. Ideally, the sensor view and extent of the biomass measurements would coincide exactly. However, I assume the error attributed to this change of support is small relative to other sources of uncertainty in the NDVI-biomass relationship.

Red clover was the dominant species in the majority of the biomass sampling locations. However, in five locations weeds comprised a substantial portion, 10%, of the sampled biomass. Because species other than red clover (e.g., grasses and weeds) tended to add noise to the NDVI values, these sample locations were marked as potential outliers and subsequently excluded from the dataset. The remaining samples were used to fit the candidate model defined in Section 2.2.3.

In addition to this sampling, the Green SeekerTM and GPS were mounted on a cart

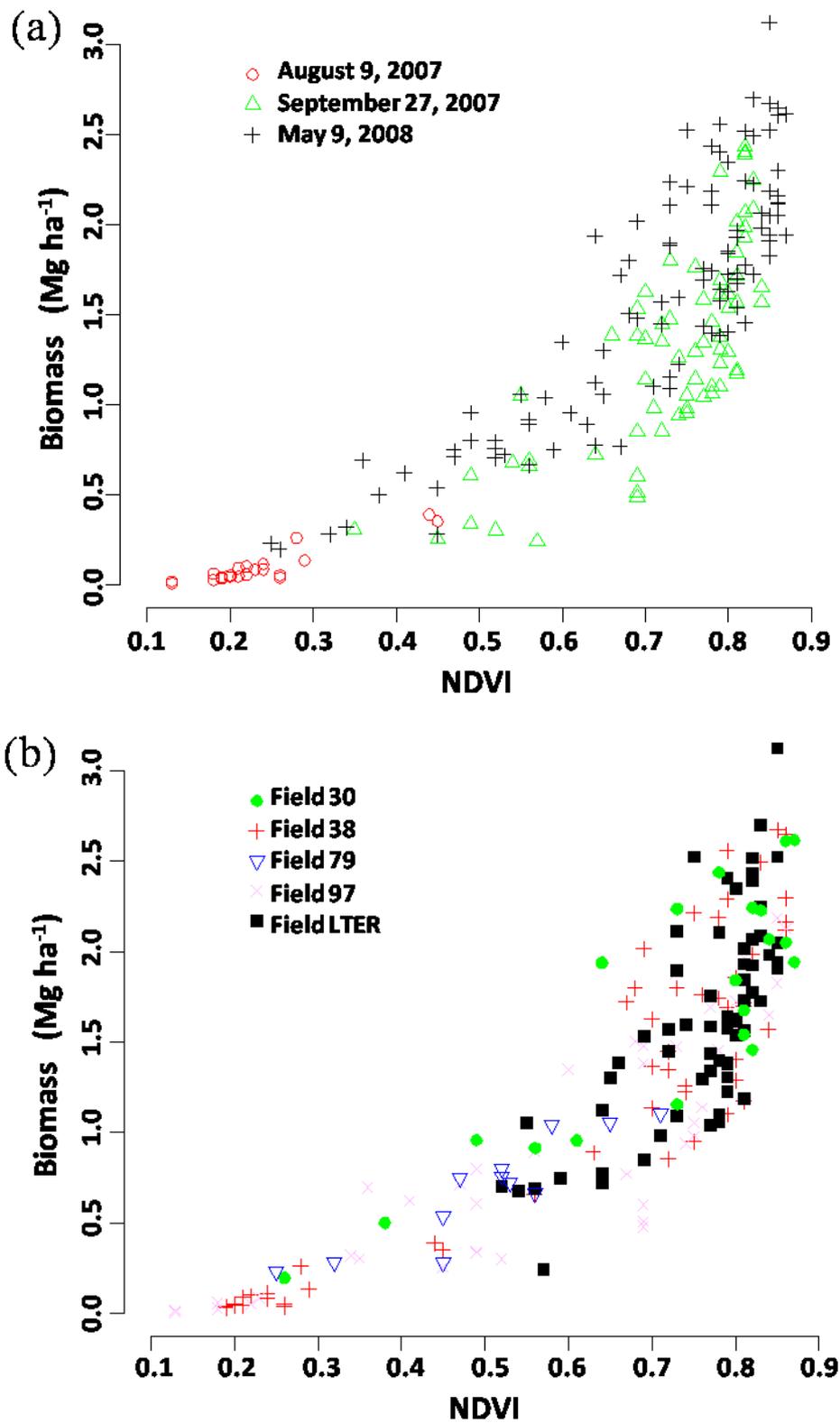
and pulled through the field twice in north-south and east-west direction. Thus, NDVI values were recorded across the entire field approximately every 3.5 m along cart passes with 10 m distance between the cart passes. This produced approximately 570 NDVI observations per ha. Total field biomass maps were produced by interpolating over predictions at these densely spaced NDVI observations.

2.2.2. Data pooling by field

As noted in Section 2.2.1, the data were collected from five fields and at three dates. Figure 2.2(a) shows NDVI versus biomass by date, pooled across field. Here we see that the range of NDVI and biomass is not well represented within any given sample date. The August date, corresponding to early stage of plant development, exhibits low biomass and corresponding NDVI while the September and May dates offer samples from the middle and end of the rotation. The lack of NDVI and biomass coverage within any individual date precluded the use of three independent models or allowing date to be a random effect in the model. Rather, to capture the NDVI-biomass relationship I chose to pool over date. Using a similar experimental design, Flynn et al. (2008) also chose to pool over date to reach a sufficient sample size and NDVI-biomass coverage.

Freeman et al. (2007) showed that fields with different soil conditions and other physical attributes can produce statistically different NDVI-biomass model parameter estimates. Figure 2.2(b) shows field specific NDVI versus biomass pooled across the sampling dates. Here we see the range of NDVI and biomass is fairly well represented within each field. By viewing observed fields as realizations from of an infinite population of fields, and modeling field as a random effect, I can explore field specific offsets from the population's mean NDVI-Biomass relationship.

Figure 2.2: a) NDVI versus biomass by sample date. b) NDVI versus biomass by field.



2.2.3. Hierarchical models

Suppose we observe total biomass and coinciding NDVI at sample locations within each of $j=1 \dots m$ fields and index these locations as $i=1 \dots n_j$, where n_j is the number of locations observed in the j -th field. Conditional upon a set of location-level measures of NDVI, I assume $y_{i,j}$, the i -th location's observed biomass in the j -th field, follows a normal distribution with a nonlinear mean function. Then modeling field as a random effect, the field specific parameters are assumed to be normally distributed and centered on the population parameters, $\alpha_j \sim Normal(\alpha, \Sigma)$, where

$$\Sigma = \text{Diag}[\sigma_i^2]_{i=0}^{p-1}$$

and p is the number of parameters. The initial model is given as

$$y_{i,j} = f(NDVI_{i,j}; \alpha_j) + \varepsilon_{i,j}; \quad \varepsilon_{i,j} \stackrel{iid}{\sim} Normal(0, \tau^2), \quad (1)$$

where $f(NDVI; \alpha)$ is a function that models the relationship between NDVI and biomass given a vector of parameters. I considered two scenarios for the distribution of the errors. First, as noted in Eq. (1), I assumed that model measurement errors are independent and normally distributed with variance τ^2 . Second, I model variance as a function of NDVI values:

$$y_{i,j} = f(NDVI_{i,j}; \alpha_j) + \varepsilon_{i,j}; \quad \varepsilon_{i,j} \stackrel{iid}{\sim} Normal(0, f_\varepsilon(NDVI_{i,j}; \Phi)), \quad (2)$$

where $f_\varepsilon(NDVI_{i,j};\Phi)$ returns a positive value given NDVI and Φ is a vector of q function parameters.

I considered two linear models for $f(NDVI;\alpha)$: the exponential model $\alpha_0 \exp(\alpha_1 NDVI)$ and the Richards model $\alpha_0 [1 - \exp(\alpha_1 NDVI)]^{1/(1-\alpha_2)}$, where $\alpha_1 \geq 0$ and $0 \leq \alpha_2 < 1$. Based on preliminary exploration I decided to use a polynomial function to describe the relationship between the variance and NDVI by setting $f_\varepsilon(NDVI_{i,j};\Phi)$ to $\Phi_0 + \Phi_1 NDVI^{\Phi_2}$, where

$$[\phi_i]_{i=1}^2 > 0.$$

For the collection $N = \sum_{j=1}^m n_j$ observations, the response vector $\mathbf{Y} = (y'_1, y'_2, \dots, y'_m)'$,

where y'_j is the $n_j \times 1$ vector of the j -th field's observations, is normally distributed,

Normal(μ_y, Σ_y). Here the $N \times 1$ vector is

$$\mu_y = [f(NDVI_{i,j}; \alpha_j)]_{i=1, j=1}^{n_j, m}$$

and the $N \times N$ dispersion matrix Σ_y is $\tau^2 I_N$ and

$$Diag[f_\varepsilon(NDVI_{i,j}; \phi)]_{i=1, j=1}^{n_j, m}$$

for models (1) and (2), respectively.

Given the different nonlinear models and residual structures I consider the following eight candidate models:

$$1a: f(NDVI_{i,j}; \alpha) = \alpha_0 \exp(\alpha_1 NDVI_{i,j})$$

$$1b: f(NDVI_{i,j}; \alpha) = \alpha_0 [1 - \exp(\alpha_1 NDVI_{i,j})]^{1/(1-\alpha_2)}$$

$$2a: f(NDVI_{i,j}; \alpha_j) = \alpha_{0,j} \exp(\alpha_{1,j} NDVI_{i,j})$$

$$2b: f(NDVI_{i,j}; \alpha_j) = \alpha_{0,j} [1 - \exp(\alpha_{1,j} NDVI_{i,j})]^{1/(1-\alpha_{2,j})}$$

$$3a: f(NDVI_{i,j}; \alpha) = \alpha_0 \exp(\alpha_1 NDVI_{i,j});$$

$$f_\varepsilon(NDVI_{i,j}; \phi) = \phi_0 + \phi_1 NDVI_{i,j}^{\phi_2}$$

$$3b: f(NDVI_{i,j}; \alpha) = \alpha_0 [1 - \exp(\alpha_1 NDVI_{i,j})]^{1/(1-\alpha_2)};$$

$$f_\varepsilon(NDVI_{i,j}; \phi) = \phi_0 + \phi_1 NDVI_{i,j}^{\phi_2}$$

$$4a: f(NDVI_{i,j}; \alpha_j) = \alpha_{0,j} \exp(\alpha_{1,j} NDVI_{i,j});$$

$$f_\varepsilon(NDVI_{i,j}; \phi) = \phi_0 + \phi_1 NDVI_{i,j}^{\phi_2}$$

$$4b: f(NDVI_{i,j}; \alpha_j) = \alpha_{0,j} [1 - \exp(\alpha_{1,j} NDVI_{i,j})]^{1/(1-\alpha_{2,j})};$$

$$f_\varepsilon(NDVI_{i,j}; \phi) = \phi_0 + \phi_1 NDVI_{i,j}^{\phi_2}$$

Models 1a and 1b use the exponential and Richards functions to model mean biomass, and assume constant residual variance. Models, 2a and 2b, assume a constant variance, but allow for field-level random effects associated with the parameters in the mean function. Models 3a and 3b use field invariant parameters in the mean function, same as 1a and 1b, but allow for residual variance to increase with increasing NDVI. Finally, models 4a and 4b, accommodate field specific offsets for parameters in the mean function and also allow for non-constant residual variance.

2.2.4. Bayesian implementation

Given the hierarchical structure of these models and my objective to properly account for uncertainty of all model parameters, I choose to use a Bayesian paradigm for model fitting and subsequent prediction (see, e.g., texts by Carlin and Louis, 2000; Gelman et al., 2004). To complete the Bayesian specification I assign a prior distribution to each model parameter. In the absence of prior information about the parameters I choose non-informative or diffuse prior distributions. For both the population and field random effect parameters I assumed α_0 and α_1 follow $Normal(0, 10^5)$ distribution for the exponential model and $\alpha_0 \sim Normal(0, 10^5)$, $\alpha_1 \sim Uniform(0, 10^5)$, and $\alpha_2 \sim Uniform(0, 1)$ for the Richards model. The inverse of the parameters τ^2 and σ^2 are assumed to follow a $Gamma(0.001, 0.001)$ distribution. Finally, the elements in follow a $Uniform(0, 10^5)$ distribution.

2.2.5. Prediction

My interest is to make inferences about model parameters and biomass over fields at a high spatial resolution. Inference, within a Bayesian paradigm, is based on Markov chain Monte Carlo (MCMC) samples (post *burn-in*) from the parameters' posterior distribution (Gelman et al., 2004). Given posterior samples for a generic parameter I can calculate any desired summary statistic, for example the parameter's posterior mean

$$\hat{\Sigma}(\theta | Y) = \frac{1}{M} \sum_{l=1}^M \theta^{(l)}$$

where l is used to index the M post burn-in MCMC samples. Second moment statistics are calculated in a similar fashion.

Point prediction for a new NDVI measurement and associated measures of

uncertainty can be calculated from the posterior predictive distribution

$$p(y^* | Y, NDVI, NDVI^*) = \int p(y^* | \Omega, Y, NDVI^*) p(\Omega | Y, NDVI^*) d\Omega \quad (3)$$

where y^* is the biomass for a new measurement of NDVI, Y and $NDVI$ are the observed data used to fit the model, and Ω is the collection of model parameters. In practice I approximate Eq. (3) using composition sampling, drawing $\Omega^{(l)}$, the l -th MCMC sample from the parameters' posterior distribution, then a corresponding $y^{*(l)} \sim p(y^* | \Omega^{(l)}, Y, NDVI^*)$ (see, e.g., Banerjee et al. 2004). The subsequent collection of $\{y^{*(1)}, y^{*(2)}, \dots, y^{*(M)}\}$ samples from the posterior predictive distribution can be summarized using any point or dispersion statistic (e.g., the simple case might be the predicted mean and variance). As an illustration consider a prediction for a new NDVI measurement in a new field using the random effects mean function and functional variance model; I simply draw first for each parameter posterior $\alpha^{*(l)} \sim Normal(\alpha^{(l)}, \Sigma^{(l)})$, and then $y^{*(l)} \sim Normal(f(NDVI^*, \alpha^{*(l)}), f(NDVI^*, \Phi^{(l)}))$ for $l = 1 \dots M$ samples from the posterior predictive distribution of the new location.

2.2.6. Model Selection

To compare several alternative models, I used the Deviance Information Criterion (DIC) (Spiegelhalter et al., 2002). DIC compares models based on the trade-off between goodness of fit and model complexity. Goodness of fit is expressed as the deviance, whereas complexity is measured by an estimate of the effective number of parameters. DIC is easily calculated from the MCMC samples and provides a generally accepted, Akaike-like, criterion for model comparison. Letting Ω be the collection of parameters in each model; the expected

posterior deviance is computed as

$$\overline{D(\Omega)} = E_{\Omega|Y} \{-2 \log L(Data | \Omega)\},$$

where $L(Data|\Omega)$ is the likelihood from the respective model. The effective number of parameters (as a penalty) was computed as

$$p_D = \overline{D(\Omega)} - D(\Omega^*),$$

where Ω^* is the posterior mean of the model parameters. Finally, DIC is given by

$$DIC = \overline{D(\Omega)} + p_D$$

thus, lower DIC values indicate better models.

2.2.7. Software implementation

The candidate models were fit using the JAGS software (Plummer, 2010) run on a Linux workstation. Convergence diagnostic and posterior inferences were then completed using the coda package in R (<http://www.r-project.org>). Chain convergence was assessed using visual inspection of the chain trace plots and the Gelman-Rubin statistic (see Gelman and Rubin 1992 for details). There are several common methods for monitoring chain convergence (see e.g., Robert 1998). The Gelman-Rubin statistic uses the correlation between the within- and among-chain variation. Convergence is achieved when the Gelman-Rubin statistic is close to 1 (usually less than 1.2 in practice).

2.3. Results and discussion

The candidate model parameter estimates, offered in Table 2.1, were based on three MCMC chains of 500,000 post burn-in iterations (10,000 samples were discarded as burn-in).

The Gelman-Rubin statistic for all parameters was between 1 and 1.2 which indicates convergence. A visual inspection of chain trace plots confirmed that they did indeed converge (plots not shown).

Figure 2.2(b) shows a distinct nonlinear pattern in NDVI-biomass relationship, with a slow increase in biomass at lower values of NDVI (≤ 0.25 NDVI), a moderate increment in the middle range (from 0.25 to 0.65 NDVI), and a rapid increase in biomass for the largest values of NDVI. This plot also suggests that NDVI's ability to predict biomass weakens as NDVI becomes larger. This trend is also noted in other studies, e.g., Carlson and Ripley (1997) found NDVI's sensitivity to leaf area index (LAI) weakens with increasing LAI. This study also describes the saturation or asymptotic behavior of NDVI at high values of LAI, or in my case biomass.

I look first to DIC to assess model fit, Table 2.2. Similar to other goodness of fit metrics, lower DIC values suggest better fitting models. Here we see little difference between models 1a and 1b, with DIC of 152.1 and 152.3, respectively. Despite similar DIC values, Figure 2.3(a) shows the mean fitted curve of the exponential model (1a) overestimates biomass at low NDVI values (solid line), whereas Richards model (1b) more closely approximates this region (dashed line). This trend persists in the subsequent models, i.e., the exponential mean function in 2a, 3a, and 4a fails to adequately model biomass at low values of NDVI. In comparison, the Richards model consistently offers better fit both visually and based on the DIC criterion.

Table 2.1. Parameter credible intervals for each candidate model. Parameter estimates expressed as 2.5% 50% 97.5% percentiles.

Model	Parameters						
	α_0	α_1	α_2	τ^2	Φ_0	Φ_1	Φ_2
1a	0.070.100.14	3.203.624.07	-	0.110.130.16	-	-	-
1b	74.05756.842224.92	0.060.120.38	0.530.600.66	0.110.130.16	-	-	-
2a	0.080.110.16	3.033.463.93	-	0.100.130.16	-	-	-
2b	46.32582.861977.93	0.050.120.46	0.500.590.66	0.100.130.16	-	-	-
3a	0.050.070.10	3.664.084.52	-	-	2.9×10^{-3} 0.010.03	0.260.430.81	1.362.133.43
3b	16.49612.372197.07	0.060.120.79	0.550.590.66	-	4×10^{-5} 3×10^{-4} 1.6×10^{-3}	0.350.510.75	1.912.332.75
4a	0.050.080.11	3.523.984.49	-	-	2.8×10^{-3} 0.010.03	0.250.440.86	1.372.203.66
4b	34.20569.181931.60	0.060.120.50	0.550.590.63	-	3.6×10^{-5} 2.5×10^{-4} 1.5×10^{-3}	0.350.500.74	1.912.342.77

Table 2.2. Comparison of model complexity and fit using the effective number of parameters (P_D) and Deviance Information Criterion (DIC).

	Model							
	1a	1b	2a	2b	3a	3b	4a	4b
P_D	3.0	3.1	6.2	6.4	6.6	6.5	11.1	10.37
DIC	152.1	152.3	149.4	147.9	114.2	66.9	112.6	62.87

Results suggest there is little gain in including field-level random effects. Figure 2.3(b) shows the population mean curve (solid) surrounded by the field specific offsets. The strong similarities in these curves indicate the negligible differences in estimated random effects associated with the $[\alpha_i]_{i=0}^2$. This marginal gain in model fit afforded by including the random effects is also reflected in a slightly lower DIC score between model sets 1 and 2, i.e., 152.3 versus 147.9 for models 1b and 2b, respectively.

There is extreme non-constant variance seen in the residual diagnostic plot, Figure 2.1, which suggests assumptions for models 1 and 2 are not met and parameter estimates and subsequent prediction accuracy could be compromised. Figure 2.4(a) illustrates shortcomings of the constant variance assumption used in models 1 and 2. Here we see that first moment of the biomass-NDVI relationship is adequately captured by the model's mean function (solid line); however, the 95% credible interval produced using the scalar variance parameter, τ^2 , cannot accommodate the tapering second moment trend. That is, the model overestimates the variance for low biomass and underestimates variance for high biomass values. Explicitly modeling this residual heteroscedasticity by including the polynomial variance structure in models 3 and 4 provides the greatest gain in model fit, Table 2.2.

Figure 2.3: a) Nonlinear models for NDVI-biomass relationship. Mean of the posterior fitted value distribution for the exponential Model 1a (solid line) and Richards Model 1b (dashed line). b) Mean of the posterior fitted values field specific random effects (dashed lines) and population (solid line) from Richards model (Model 2b).

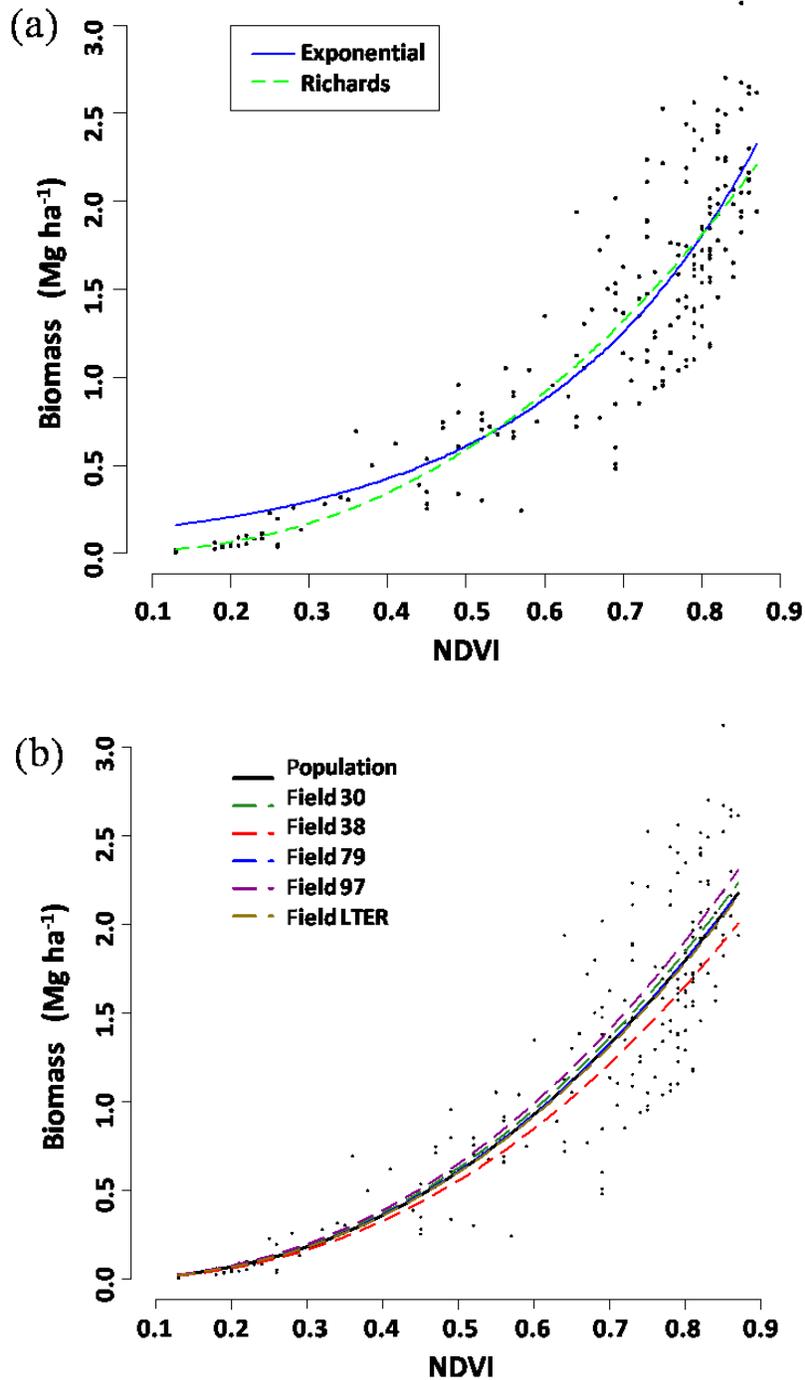


Figure 2.4: a) Upper and lower 95% credible interval for the population's posterior predictive distribution (dashed lines) along with the associated mean (solid line) for Model 2b. b) Model 2b, constant variance model, mean residual variance (solid points) calculated within fixed-width NDVI intervals (vertical dashed lines) with fitted curve from the Model 3b variance function. c) Model 3b upper and lower 95% credible interval for the population's posterior predictive distribution (dashed lines) along with the associated mean (solid line).

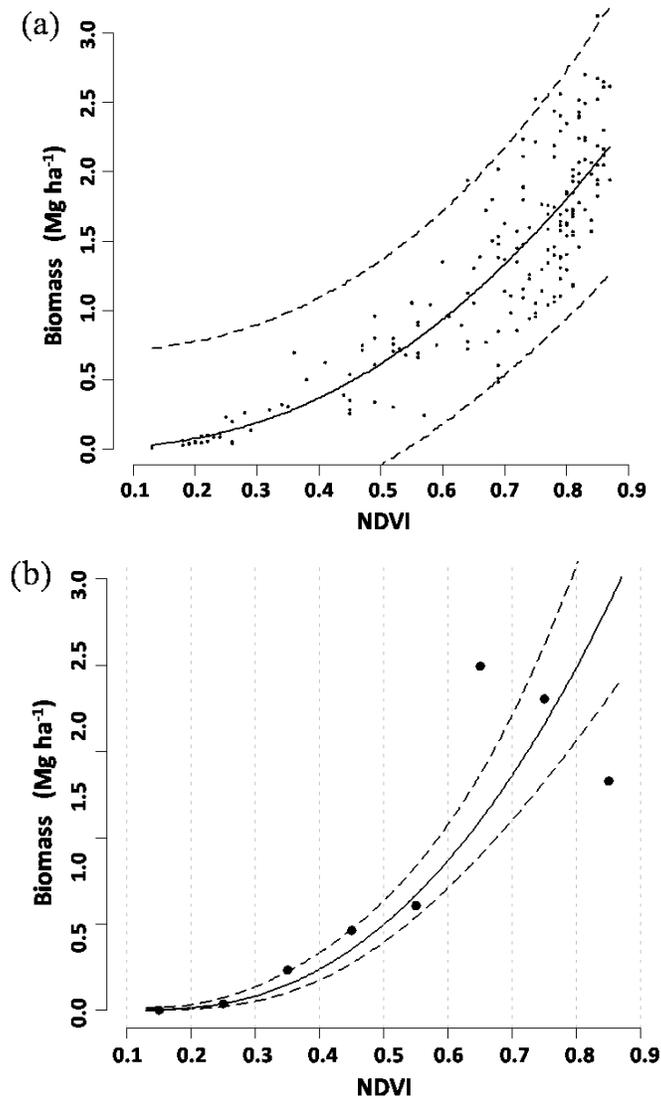
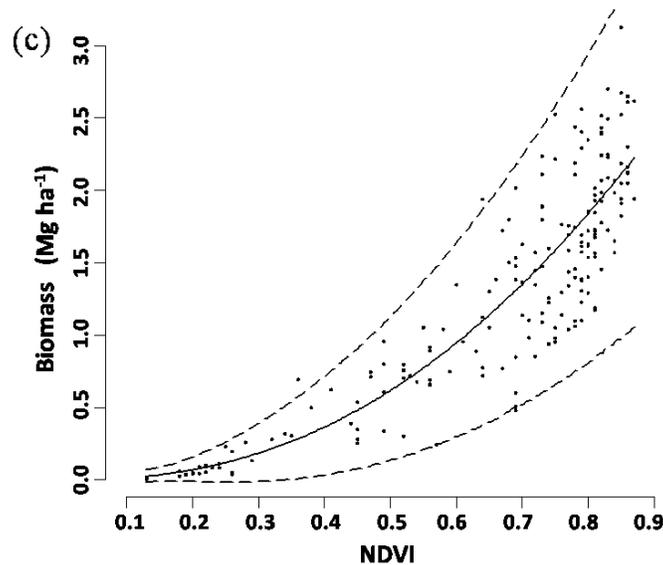


Figure 2.4. (cont'd).



My choice to use a polynomial function to model the non-constant variance was based on Figure 2.4(b). This figure was constructed by calculating the constant variance model's mean residual value within 0.1 unit width NDVI intervals. Here these values are illustrated as eight points. I noted a trend that could be approximated by a polynomial. A more formal approach could be undertaken; however, I felt this ad-hoc approach was adequate. Using the polynomial's parameter estimates, Table 1, I plot the mean and associated 95% credible interval over the points in Figure 2.4(b). These lines appear to capture the increasing residual variance and lend support to my choice of variance function.

Importantly, unlike the homogeneous variance model fit, Figure 2.4(a), the 95% credible interval of the heterogeneous variance model, Figure 2.4(c), exhibits a more reasonable estimate of residual variance, reducing uncertainty in biomass prediction at lower values of NDVI.

Beyond inferences about the models' parameters, correctly accounting for the non-constant variance is critical for predicting biomass for new NDVI values. The final study

objective was to produce digital data layers of predicted biomass with associated uncertainty at a fine resolution across a given field. These data layers will serve as input for subsequent predictions of the growth and yield of the main crop. Using the dense NDVI measurements taken across the entire field described in Section 2.2.1. and following the methods detailed in Section 2.2.5. I predicted biomass at each sample location for each field using model 2b and 4b. Figure 2.5 illustrates the mean predicted biomass and associated standard deviation at sample locations for field 38 on August 9 using the two models. The predicted values at these sample locations were then passed through a deterministic interpolation to produce the surface plots shown in Figure 2.5.

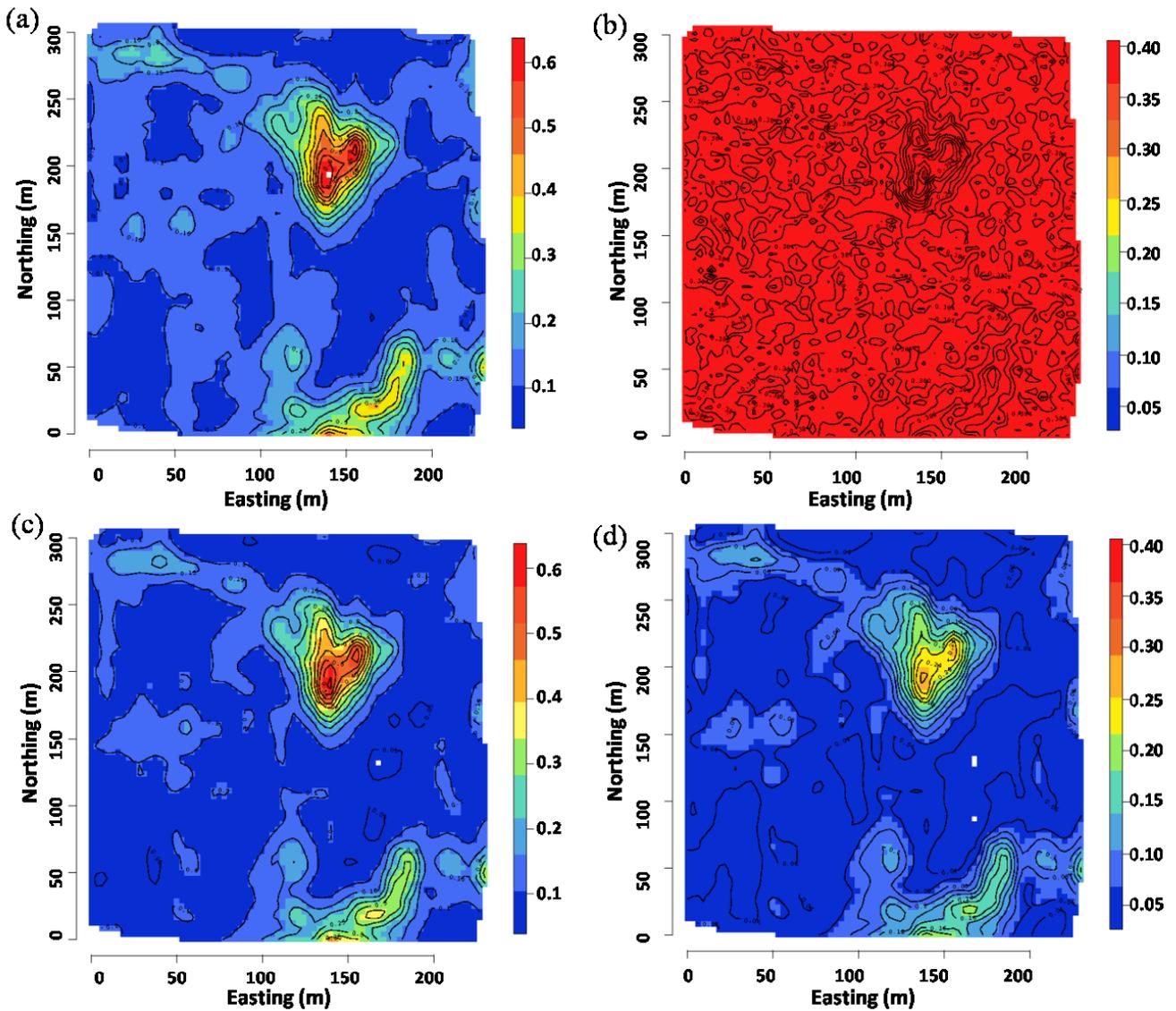
As expected, the two models produce nearly identical mean biomass surfaces, Figure 2.5(a) and (c). However, the constant variance model is again unable to describe the large difference in posterior predictive uncertainty associated with mean biomass across the field as illustrated by the near constant standard deviation surface (b). In comparison, the standard deviation surface (d) shows the functional variance model more fully express the uncertainty associated with the spatial-varying mean biomass.

I turn again to the discussion of data pooling. An obvious risk in pooling over date is the relationship between NDVI and biomass might vary over the growing season. For instance, changes in NDVI at different stages of plant maturity might be partially explained by leaf phenology, e.g., compact mesophyll in young leaf versus lacunate mesophyll in older leaves and subsequent increase in transmittance as crop matures. Raun et al. (2001) and Teal et al. (2006) showed NDVI versus grain yield can change across plant growth stages. To adjust for this growth stage dependence, they used a predictor variable called in-season estimated yield (INSEY) which is NDVI divided by days from planting or growing degree days. Following

these studies, I too modeled INSEY versus red clover biomass. The results showed a substantial decrease in model fit over the non-adjusted NDVI models. I attribute this to the large time lag between the 2nd and 3rd sampling visits which make my setting different from those of Raun et al. (2001) and Teal et al. (2006) where the sampling dates were shorter and near or at grain maturity. Further, red clover plant architecture is very different from wheat and corn, which will greatly influence light reflectance. Row-crops architecture results in substantial change in light interception at different phenological stages; whereas in red clover these changes are not as accentuated. Therefore I do not pursue the use of INSEY or similar adjustments in this study. However, the possibility remains that the relationship between NDVI and biomass has a temporal component. If this occurs in my setting, my ability to properly capture the true relationship between NDVI and biomass will be muddled and the postulated model fit and predictive ability will decrease.

A final, and very important, modeling consideration is that of spatial dependence. It is quite possible the candidate models' mean function is not able to account for all of the variability in the observed biomass. If this is true, then I can expect spatial dependence among the residuals. This would violate the assumptions of models (1) and (2). Ignoring residual spatial dependence may result in falsely precise estimates of model parameters and erroneous predictions. Hoeting (2009) offers a nice discussion on this topic. Field specific empirical semivariograms (plots not shown) did not suggest any spatially structured dependence among my candidate model residuals (see e.g., Cressie, 1993). This is likely because the NDVI variable itself exhibits strong spatial dependence and therefore accounts for the majority of spatial structured variability in biomass.

Figure 2.5. Spatial interpolation of observed NDVI and predicted biomass for Field-98 on August 9. Mean and standard deviation a) and b) respectively, of the biomass predictive distribution estimated with Richards mean function and constant variance (Model 2b). Mean and standard deviation c) and d) respectively, of the biomass predictive distribution estimated with Richards mean function and functional variance (Model 4b).



2.4. Conclusion

In this study I used NDVI to model red clover biomass. The NDVI-biomass data show a clear nonlinear relationship. Other studies have used the exponential function to model similar nonlinear trends. Here, I found a Richards nonlinear function more closely approximates the data scatter and also produces improved model fit over the exponential function. Further, I explored field-to-field variability using random effects on the model's mean function parameters. However, given the paucity of data or small differences in field-level productivity, there was little support for the inclusion of random effects. Despite this result, I encourage the use of hierarchical structures whenever field variability is relevant, particularly for experiments that include fields from different environmental and soil conditions.

The methods and models presented here are general and can be applied in other settings where remotely sensed data is used to describe the variability in biological variables of interest. From a statistical standpoint, it is important that we accommodate non-constant variance among the residuals. From a practitioners' perspective, and as seen in my results, addressing heteroscedasticity using a functional variance can improve model fit and predictions. These spatially explicit predictions of biomass and associated uncertainty are critical data products that will ultimately help to quantify the effects of cover crop residuals on main crop yields.

REFERENCES

REFERENCES

- Banerjee, S., Carlin, B.P., and Gelfand, A.E. 2004. Hierarchical Modeling and Analysis for Spatial Data, Boca Raton, FL: Chapman and Hall/CRC Press.
- Calvaio, T., and Palmeirin, J. 2004. Mapping Mediterranean scrub with satellite imagery: biomass estimation and spectral behaviour. *Int. J. of Rem. Sen.* 25: 16, 3113- 3126.
- Carlson, T., Ripley, D. 1997. On the relation between NDVI, fractional vegetation cover, and leaf area index. *Rem. Sen. of Environ* 62: 241-252.
- Carlin, B.P. and Louis, T.A. 2000. Bayes and Empirical Bayes Methods for Data Analysis. 2nd ed., Boca Raton, FL: Chapman and Hall/CRC Press.
- Corbin, D. and VandellWulp, S. 2010. KBS019: Aboveground Net Primary Productivity Protocol. Available: <http://lter.kbs.msu.edu/protocols/111> (Accessed: 2010, May 18)
- Cressie, N. A. C. 1993. Statistics for Spatial Data, 2nd edition. New York: Wiley.
- Du Plessis, W. 1999. Linear regression relationships between NDVI, vegetation and rainfall in Etosha National Park, Namibia. *J. of Arid Environ.*, 42: 235-260.
- Flynn, E., Dougherty, C. and Wendroth, O. 2008. Assessment of pasture biomass with the normalized difference vegetation index from active ground-based sensors. *Agron. J.*, 100: 114-121.
- Freeman, K., Girma, K., Arnall, D., Mullen, R., Martin, K., Teal, R., Raun, W. 2007. By-plant prediction of corn forage biomass and nitrogen uptake at various growth stages using remote sensing and plant height. *Agron. J.* 99: 530-536.
- Gelman, A., Carlin, J.B., Stern, H.S., and Rubin, D.B. 2004. Bayesian Data Analysis. Second Edition. Boca Raton, FL: Chapman and Hall/CRC Press.
- Gelman, A and Rubin, D. 1992. Inference from iterative simulation using multiple sequences, *Stat. Sci.*, 7, 457-511.
- Guretzky, J., Moore, K., Burras, C. and Brummer, E. 2004. Distribution of Legumes along Gradients of Slope and Soil Electrical Conductivity in Pastures. *Agron. J.*, 96: 547-555.
- Harmony, K., Moore, K., George, R., Brummer, C. and Russel, J. 1997. Determination of pasture biomass using four indirect methods. *Agron. J.*, 89: 665-672.
- Hoeting, J.A. 2009. The importance of accounting for spatial and temporal correlation in analyses of ecological data. *Ecol. Applications* 19:584–588.

- Kerr, R., Snapp, S., Chirwa, M., Shumba, L. and Msachi, R. 2007. Participatory research on legume diversification with Malawian smallholder farmers for improved human nutrition and soil fertility. *Expl Agric.*, 43: 437-453.
- Martin, R., Astatkie, T., Cooper, J. and Fredeen, A. 2005. A Comparison of Methods Used to Determine Biomass on Naturalized Swards. *J. Agronomy and Crop Science*, 191: 152-160.
- Papadopoulos, A., Mooney, S. and Bird, N. 2006. Quantification of the effects of contrasting crops in the development of soil structure: an organic conversion. *Soil Use and Manag.* 22: 172-179.
- Plummer, M. 2010. JAGS: Just Another GIBBS Sampler. (version 2.1.0.). Available: <http://www-fis.iarc.fr/martyn/software/jags/> (Accessed: 2010, May 18)
- Radloff, F. and Mucina, L. 2007. A quick and robust method for biomass estimation in structurally diverse vegetation. *J. of Veg. Sci.* 18: 719-724.
- Raun, W., Solie, J., Jonhson, G., Stone, M., Lukina, E., Thomason, W. and Schepers, J. 2001. In-season prediction of potential grain yield in winter wheat using canopy reflectance. *Agron J.*, 93: 131-138.
- Robert, C. 1998. Discretization and MCMC Convergence Assessment. Lecture Notes in Statistics. Springer, New York.
- Snapp, S., Swinton, S., Labarta, L., Mutch, D., Black, J., Leep, R., Nyiraneza, J., and O'Neil, K. 2005. Evaluating cover crops for benefits, costs and performance within cropping system niches. *Agron. J.*, 97: 322-332.
- Spiegelhalter, D.J., Best, N.G., Carlin, B.P., and van der Linde, A. 2002. Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society, Series B* 64, 583–639.
- Teal, R., Tubana, B., Girma, K., Freeman, K., Arnall, D., Walsh, O., Raun, W. 2006. In-season prediction of corn grain yield potential using normalized difference vegetation index. *Agron. J.*, 98: 1488-1494.
- Thenkabail, P.S., Smith, R.B., and De Pauw, E. 2000. Hyperspectral vegetation indices for determining agricultural crop characteristics. *Rem. Sen. of Environ.* 71: 158–182.
- Tood, S., Hoffer, R. and Milchunas, D. 1998. Biomass estimation on grazed and ungrazed rangelands using spectral indices. *Int. J. Rem. Sen.* 19(3): 427-438.
- Whitbeck, M. and Grace, J. 2006. Evaluation of non-destructive methods for estimating biomass in marshes of the upper texas, USA Coast. *Wetlands*, 26(1): 278-282.
- Xu, B., Yang, X., Tao, W., Qin, Z., Liu, H., Miao, J. and Bi, Y. 2008. MODIS-based remote sensing monitoring of grass production in China. *Int. J. of Rem. Sen.* 29:17, 5313-5327.

Chapter 3: Deriving the optimal scale for relating topographic attributes and cover crop plant biomass

Abstract

The use of cover crops generates a number of agro-ecological benefits for sustainable row-crop agriculture. However, their performance across agricultural fields is often highly spatially variable and there is insufficient information on factors affecting this variability and on tools to manage it. Topography is one of the main factors affecting spatial patterns of plant growth in the American Midwest. Digital elevation models are readily available for deriving topographic attributes; also sensor digital data can be used to indirectly assess cover crop biomass. However, processing procedures for identifying the proper scale of topographic and biomass representations are not well defined. The objectives of this study are to examine how relationships between cover crop biomass, assessed using the Normalized Difference Vegetation Index (*NDVI*), and topography depend on the neighborhood size used for deriving topographic attributes and creating *NDVI* maps; and identify the optimal neighborhood size for correlation and regression analyses. Slope, relative elevation and the potential solar radiation index were the variables that contributed the most to explaining variability in *NDVI* for raw data. However, other topographic attributes became significant predictors of *NDVI* at larger neighborhood sizes. I demonstrated that neighborhood size greatly affects some topographic attributes, i.e. curvature, flow accumulation, flow length and the wetness index; and changing the neighborhood size in both topography and *NDVI* considerably changes the strength of the prediction performance in

multiple regression models. I studied six neighborhood sizes from 1 to 40 m and the original raw data. On average, across all studied fields the best performance of multiple regression, as determined by the adjusted- R^2 , was obtained at neighborhood sizes 20 and 40 m. Parameters of semivariogram models for terrain slope, such as the spatial autocorrelation range and nugget/sill ratio, were found to be good indicators of prediction performance and optimum neighborhood size for filtering the raw data. The results demonstrate that topographic effects on growth and biomass production of cover crops are most pronounced at certain spatial scales, and topographic model predictions will be most accurate when used at the optimal scales.

Keywords: DEM, Topography, Spatial scale, Neighborhood size, *NDVI*.

3.1. Introduction

Cover crops are planted between periods of main crop growth typically not to produce food but to provide a variety of benefits including improvement in soil quality, weed and disease control and reduction in soil erosion. Cover crops are of great importance for today's sustainable agriculture (Snapp et al, 2005). However, they can be relatively difficult to establish and their temporal and spatial variability creates problems for achieving consistent performance across agricultural fields. In undulating terrain of the Midwestern US, topography is one of the main factors driving spatial and spatio-temporal variability of plant growth, along with weather patterns. A better understanding of how variations in topography affect cover crops will allow producers to tailor cover-crop management to site-specific features of their fields, and therefore will contribute to the increased cover-crop adoption.

Topographic attributes are widely used to model hydrological and pedological processes as well as to study biological and ecological phenomena at landscape scale (McBratney et al. 2003). Examples of the use of topographic information are found in soil water availability applications (Green and Erskine, 2004; Ticehurst et al., 2007), spatial variation of soil chemical properties (Goovaerts and Webster, 1994), soil carbon mapping (Huang et al., 2007), and vegetation cover properties and plant density (Florinsky and Kuryakova, 1996). Even though a large number of studies has been conducted on relationships between topography and a variety of row crops, i.e., corn, soybean, and wheat (Kravchenko and Bullock, 2000; Green and Erskine, 2004; Kravchenko et al., 2005; Huang et al., 2008) and the information about these relationships is being utilized in site-specific management (McCann et al., 1996; Lark, 1998; McMillan et al., 1998; Fraise et al., 2001), relatively little is known about relationships between topography and cover crops, e.g., red clover.

The following topographic attributes are most often used to represent the landscape surface: slope, aspect, curvature, flow accumulation, flow length and the wetness index. These attributes are commonly derived from digital elevation models (DEMs). However, the quality of derived attributes is highly dependent on a number of processing decisions, the DEM source, scale and resolution (Bolstad and Stowe, 1994; Florinsky, 1998; Hengl, 2006; Erskine, et al. 2007; Liu, 2008). Several studies highlight the importance of using appropriate resolution and scale for topographic attribute derivation (Curran and Atkinson, 1999; Florinsky and Kuryakova, 2000; Erskine et al., 2007), and it appears that optimal attribute derivation to obtain the most accurate relationships between terrain and biophysical variables varies from landscape to landscape (Dobermann et al., 1997; Smith et al., 2006; Wu et al., 2008; Roecker and Thompson, 2010). For instance, Wu et al. (2008) showed how topographic attributes differed across

resolutions, finding that the best model to get correlations between soil properties and topography was not always the one with the highest resolution. Since there is no unique resolution and scale that would yield the best attribute derivation for all geographical scenarios and applications, it is important to develop an approach for identifying such derivations of topographic attributes that would provide optimal results for different specific landscapes and application scenarios.

Even though the terms DEM scale and DEM resolution are often used interchangeably, they are different concepts (Gallant and Hutchinson, 1997). Resolution is typically defined as the pixel size (Curran and Atkinson, 1999; Hengl, 2006), also often called grid size (Florinsky and Kuryakova, 2000) or cell size (Erskine et al., 2007). On the other hand, scale is defined as the spatial extent at which terrain attributes vary (Gallant and Hutchinson, 1997; Smith et al., 2006). A proper digital representation of a terrain attribute should match the scale at which it varies; for instance the large scale of a watershed or the small scale of micro-relief observed within a field. Fine resolution DEMs derived from sensor-based data (e.g. LiDAR) frequently require pre-processing steps before their use in most environmental and geomorphological applications (Milledge et al., 2009). However, the typical practice of increasing pixel size (i.e. decreasing resolution) could mask topographic features at smaller scales, for instance micro-relief variation.

The effect of changing scale is known as scaling effect (Wu et al., 2008; Roecker and Thompson, 2010). Since scale is dependent on the resolution in the DEM source, changing the pixel size will affect the scale. According to Gallant and Hutchinson (1997), if we wish to study the effect of scale rather than the effect of pixel size, we need to use a method that permits changes in scales without changing the pixel size. Changing the neighborhood size in attribute derivation is one of the ways of studying the effect of scale (Bian and Walsh, 1993; Roecker and

Thompson, 2010). For example, Smith et al. (2006) observed that the neighborhood sizes that produce the most accurate results in digital soil survey may vary depending on the landscape, thus soils on a gently rolling landscape were most accurately mapped using neighborhood sizes of 33–48 m, whereas soils on short, steep backslope positions were most accurately mapped using neighborhood sizes of 24–36 m. Roecker and Thompson (2010) suggested that the optimal neighborhood size might be specific to the landscape scenarios and that the relative size of landforms could serve as a guide for choosing an optimal neighborhood size. To get the greatest benefit from the topographic data for predicting biophysical attributes we need to use the topographic data with the scale at which IT has the highest influence on the studied attributes.

Direct measurement of cover crop biomass is a time- and labor-consuming task and even more so when the aim is to characterize their spatial variability. Remote sensing tools, e.g., GreenSeekerTM, a ground-based optical sensor that measures normalized difference vegetation index (*NDVI*), can be used as a proxy for plant biomass estimation (Muñoz et al., 2010). However, similar to the DEM data, the remote sensing *NDVI* data can also be collected with different resolutions and processed at different neighborhood sizes. For example, the resolution of *NDVI* obtained from the GreenSeeker device can be set to <1 m. However, such high resolution may include noise due to small patches of covertures. Thus it is likely that *NDVI* filtered at a greater neighborhood size might be better for predicting field-scale cover crop patterns. For example, Flynn et al. (2008) filtered out the GreenSeeker *NDVI* data using a spatial buffer of 11.43 m in radius, finding better correlations between *NDVI* and biomass of tall fescue [*Schenodorus arundinaceus* (Schreb.) Dumort]. Determining optimal parameters for processing cover crop remote sensing biomass data would improve both the efficiency of data collection and the accuracy of cover crop mapping and landscape scale assessments.

The objectives of this study are 1) to examine the relationship between cover crop prediction using topographic data and the neighborhood size of topographic attribute derivation and *NDVI* measurements; 2) to identify the optimal neighborhood size for correlating them; and 3) to find criteria that could be used to facilitate optimal neighborhood size.

3.2. Materials and methods

3.2.1. Study site

The data were collected from eight agricultural fields of the scale-up experiment at the Long Term Ecological Research site of the Kellogg Biological Station located in southwest Michigan, USA ($42^{\circ} 24' N$, $85^{\circ} 24' W$) (Fig. 3.1). The dominant soil series are the Kalamazoo (fine-loamy, mixed, mesic Typic Hapludalfs) and Oshtemo (coarse-loamy, mixed, mesic Typic Hapludalfs) (Crum and Collins, 2012). All fields are in a corn-soybean-wheat rotation with a winter leguminous cover crop, red clover (*Trifolium pratense* L.). In each field red clover is frost seeded in winter wheat (i.e. March) and then ploughed prior to corn planting in spring the following year (i.e. May). Four of the studied fields are in a certified organic management system (T4) and receive no chemical inputs, whereas the other four are in a reduced chemical input system (T3) and receive banded herbicide and starter nitrogen (N) at planting. Both T3 and T4 receive additional post-planting cultivation and T4 is rotary-hoed to control weeds (Simmons, 2012). Table 3.1 shows the list of the studied fields, their sizes and the years when red clover data were collected in them. The fields selected for this study represent diverse topographic conditions, ranging from almost flat, e.g. fields 52 and 791 where the difference between maximum and minimum elevations is just 3.7 m, to more complex topographic shapes, e.g. field

30 where the difference is 10.9 m (Table 3.1). Field 38 has a clear concave “bowl-shaped” depression, while gently rolling slopes dominate in field 87. Fig. 3.2 shows the 3D representation of relative elevation of the studied fields overlaid with the *NDVI* map. The map is a representation of the spatial variability of cover crop biomass.

Figure 3.1. Location of the Kellogg Biological Station (KBS) in Southwest Michigan. Studied fields are marked by red outlines; field ID labels are shown within each field.

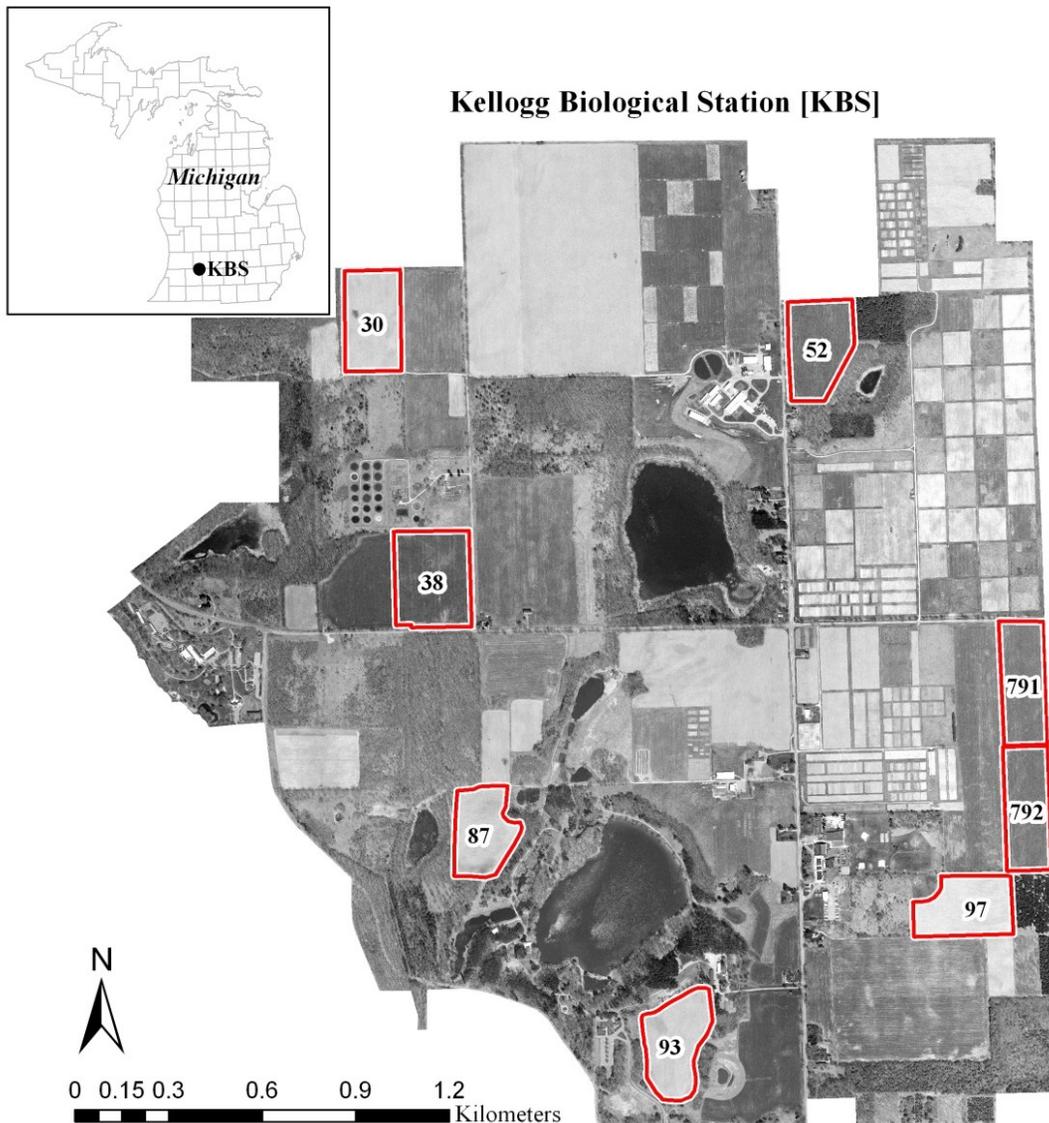


Figure 3.2. NDVI values of the studied fields plotted over fields' digital elevation models (DEMs). The DEMs show the landscape configuration of each field, whereas NDVI represents the spatial variability of cover crop biomass. The map resolution is 1×1 m.

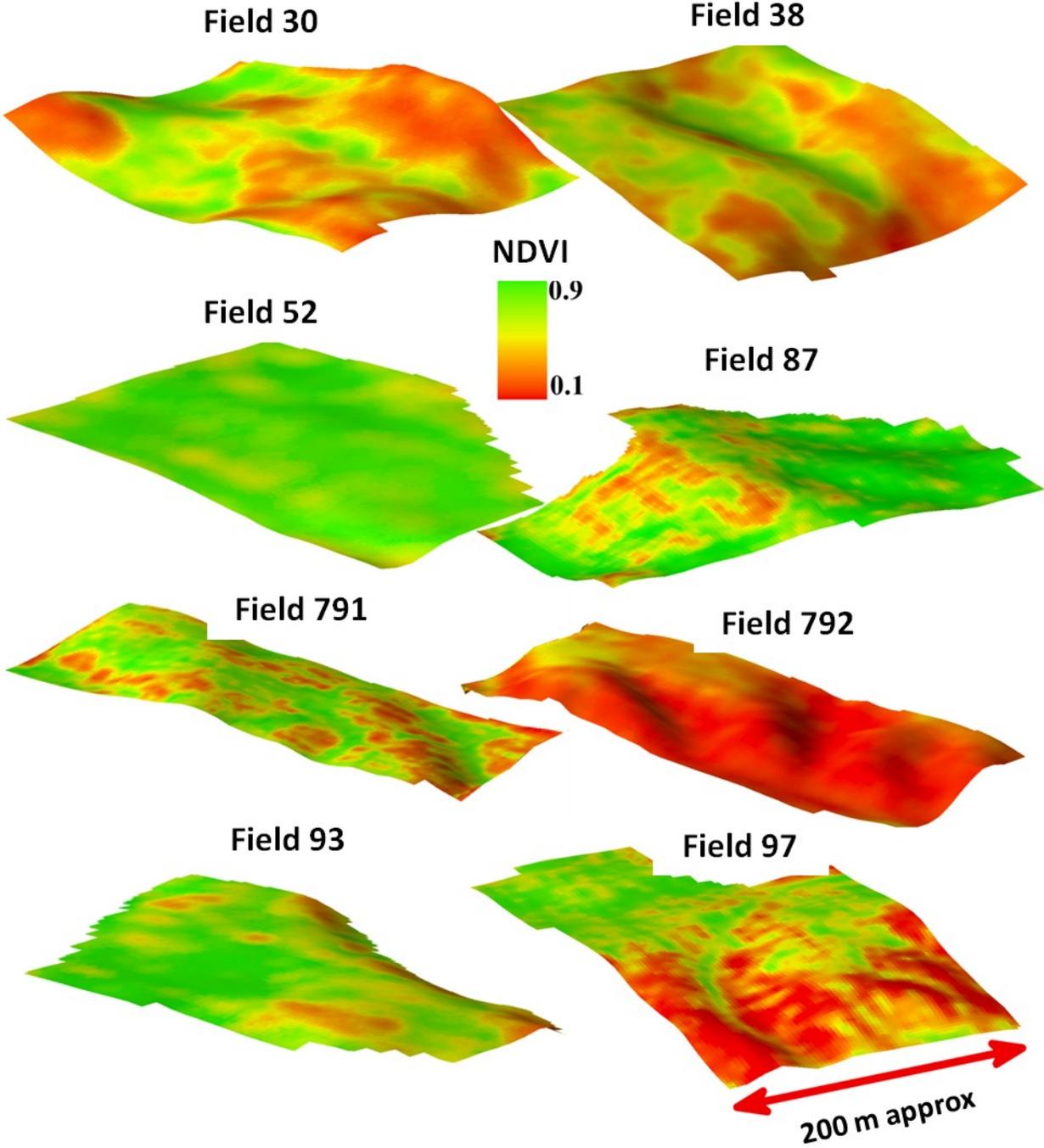


Table 3.1. List of the fields used in the study along with the year when *NDVI* data were collected from the field, field size, and the range of elevation values.

Field Name	Year	Area (ha)	Range of elevation (m)
30	2007	5.8	10.9
38	2007	7.4	6.1
792	2007	5.7	7.0
97	2007	5.4	6.3
93	2008	5.6	3.8
87	2008	4.9	8.3
52	2008	5.9	3.7
791	2008	5.7	3.7

3.2.2. Data collection and pre-processing

Biomass of red clover was indirectly assessed using measurements of *NDVI* taken just before red clover plowing, i.e. the first week of May each year. The data were taken with a portable optical sensor device (GreenSeeker model RT200; NTech industries, Inc., Ukia, CA, USA) installed on a cart. As the cart was driven through the field, *NDVI* readings were taken approximately every 3.5 m along cart passes with 10 m distance between the passes. The cart was driven through the field twice in north–south and east–west directions. This produced approximately 570 observations per ha. The height of the sensor (1 m above the plant) ensured an approximate field of view of 1×61 cm for each reading. The optical detector measured reflectance at two wavelengths (red and near infrared) originally emitted by the same device.

Ordinary kriging was used to interpolate *NDVI* to a continuous map for each field using the Geostatistical Analyst in ArcGIS 9.2 (ESRI, 2003). Semivariances were computed with 1 m lag for a total of 100 lags. For all fields, the semivariance for the first lag (1 m) was computed based on more than 800 data pairs. Raster files were exported to 1 m resolution (Fig. 3.2).

A LiDAR dataset with altimetry values was used to generate a DEM in each field, with vertical accuracy < 15 cm and horizontal accuracy < 1 m. LiDAR-based elevation data points were available for the study area with a density of ca. 0.8 points m⁻². A DEM was generated from these LiDAR points using inverse distance weighting interpolation and a raster file was exported with 1 m resolution. This file is hereafter referred to as “original DEM”. Topographic attributes derived from the DEM included relative elevation, slope, flow accumulation, flow length, curvature, and the potential solar radiation index. The detailed descriptions of these topographic attributes have been reported in a number of previous studies (Moore et al, 1993; Townsend and Walsh, 1996) and I do not repeat them here. Values of the potential solar radiation index were derived using the hemispherical viewshed algorithm (Fu and Rich, 2002); it is computed using spatial location parameters (latitude and altitude), time configuration (time period) for specific days or seasons (i.e. summer) and terrain attributes (slope and aspect).

The topographic attributes were used to quantify the influence of topography on red clover cover crop, specifically on its biomass. As a proxy for the biomass, I use *NDVI* collected at a fine resolution across each field, thus permitting detailed spatial analysis. The use of *NDVI* as a predictor of red clover biomass was previously tested using non-linear hierarchical models (Munoz et al., 2010). Thus, the relationship between topography and *NDVI* will be used in this study as a representation of the relationship between topography and red clover biomass. To

study the relationship at different scales, the original DEM was filtered at different neighborhood sizes. A circle window was drawn around the center of each pixel and the original data value of each pixel was replaced by the average value of the neighboring pixels whose centers fell inside the circle (i.e. the focal circle filter of ArcGIS 9.2). The window was moved to the next pixel and the average was computed again, thus part of the filtering windows for two contiguous pixels overlapped. Six sizes of the moving window radius, i.e., 1, 3, 5, 10, 20, 40 m, were applied. These sizes are hereafter referred to as *NHS* (neighborhood size). The topographic attributes were derived first from the original DEM and then from DEMs of different *NHS* values. The same approach was followed for *NDVI* maps.

DEM and *NDVI* manipulation and processing were performed using ArcGIS 9.2, and 3D representations of terrain surfaces were generated under ArcGIS-ArcScene 9.2 (ESRI, 2003).

3.2.3. Statistical analysis

Descriptive statistics were obtained for all studied attributes at each *NHS*. Probability density functions were estimated for each topographic attribute and *NDVI* using the kernel density estimation, which is a non-parametric approach that optimizes the representation of a histogram (Rizzo, 2008). A kernel weight function was defined based on the sample data, and subsequently the optimum bandwidth was selected based on the method of Sheather and Jones (1991). Probability densities were estimated for the original layer and 5 and 40 m *NHS* values.

The relationships between original topographic attributes, *NDVI* data, and their filtered counterparts at different *NHS* values were assessed using Pearson's correlation coefficients (r). Multiple regression analyses were performed using *NDVI* as a response variable and the entire

set of topographic attributes as predictors. Significance of the contribution of individual variables to the model was evaluated with a t-test ($\alpha = 0.05$). Residuals were checked for normality using normal probability plots. An optimal *NHS* was defined as the one that produced the highest adjusted- R^2 value of multiple regression. To identify the optimal *NHS* value I performed the analysis of variance with the adjusted- R^2 as the response variable, and topographic *NHS* and *NDVI NHS* as fixed factors with 7 levels in each factor (i.e. original layer and 1, 3, 5, 10, 20 and 40 m *NHS*). Then, the mean adjusted- R^2 values among different combinations of the factor levels were compared using Tukey's HSD test for all pair-wise comparisons. Differences among levels of each factor were considered statistically significantly at $\alpha = 0.05$. In order to increase computational efficiency and to avoid possible autocorrelation in the datasets, I have extracted values of the topographic parameters and *NDVI* from every 30th pixel in a regular grid (i.e. 30×30 m). This reduced the number of data points from >60,000 to about 72 per field. Selection of every 30th pixel was sufficient to avoid spatial autocorrelation among the data; for example the semivariograms of *NDVI* and topographic attributes showed no evident autocorrelation beyond a lag distance of 30 m (data not shown).

Principal component analysis was performed on the topographic attributes. The loading values for each attribute were extracted from the principal components (PCs) with eigenvalues larger or equal to 1 (Johnson and Wichern, 2002). The scores in each observation were extracted from each PC and the average score in each field was computed. I used these scores as potential indicators of the optimal *NHS* size by regressing them against the mean adjusted- R^2 from the multiple linear regressions between topographic predictors and *NDVI*. Average PC scores were also regressed against the change in adjusted- R^2 values after filtering with each *NHS* as an

indication of model improvement. The standardized range of topographic variables in each field was computed as

$$S_i = (T_i - X) / D \quad (1)$$

where S_i is the standardized value in field i , T_i is the original value in field i , and X and D are the mean and the standard deviation of the topographic attribute across all fields, respectively. Then the standardized ranges were regressed against the mean adjusted- R^2 from the multiple linear regressions.

Semivariogram models were fitted to the studied topographic attributes and $NDVI$ by ordinary least squares using GeoR library (Ribeiro and Diggle, 2001) of R (R Development Core Team, 2009). The lag size used for variography was equal to 2 m and the maximum distance was equal to 100 m. Variogram model selection was based on the lowest Root Mean Square Error ($RMSE$) obtained during cross-validation. I then explored the correlation between semivariogram parameters and mean adjusted- R^2 from multiple linear regressions. These statistical analyses were performed using R and proc MIXED, SAS/STAT software version 9.2 (SAS Institute, 2008).

3.3. Results and discussion

3.3.1. Effect of scaling on the topographic factors

Increasing NHS decreased the ranges of all topographic attributes and changed the means of some attributes. Fig. 3.3 shows the histograms of the topographic attributes from the original

and filtered DEMs. When *NHS* increased to 40 m the range of relative elevation was reduced to 82.4% of the original data (Fig. 3.3A). Slope has become substantially smoother, with maximum reduced from 9.3% to 4.9% and the mean from 2.2% to 1.4% (Fig. 3.3B). The mean values of curvature did not change; however, the range drastically decreased to 2.5% of the original. Already at 5 m radius, the smoothing eliminated most extreme curvature values (Fig. 3.3E). Similar decreases in the ranges of slope and curvature with increasing *NHS* were reported by Roecker and Thompson (2010). Flow accumulation and flow length increased considerably as *NHS* increased. For example, the mean of flow accumulation changed from 27 to 59 pixels (i.e. 1.5 to 3.6 in log scale), indicating that the area contributing to water flow increased up to 218% in the 40-m filtered map (Fig. 3.3G). Fig. 3.3H shows the original, 5-m filtered and 40-m filtered maps of flow accumulation in a portion of field 792, confirming such an effect of filtering. Natural barriers in the surface were filtered out; hence water flow paths increased in length and on average, the depressed areas received more water as compared to the original flow accumulation map (Fig. 3.3H). In addition, the range of flow accumulation decreased with increasing *NHS*. The average flow length of the original DEM was 65 pixels, whereas that of *NHS* = 40m smoothed version increased to 128 pixels (respectively, 7.3 and 40.9 in square root scale), indicating longer travel paths (Fig. 3.3C). The range of potential solar radiation somewhat decreased, but the change in mean value was relatively minor (Fig. 3.3D). Overall, the relative elevation and potential solar radiation were the least sensitive to DEM filtering, followed by slope, and then by flow accumulation, flow length and curvature that experienced the greatest changes.

Figure 3.3. Histograms of the original topographic features and NDVI values overlaid with their probability density functions at three studied neighborhood sizes (original, 5 m and 40 m). A) Relative elevation. B) Slope. C) Square root of the flow length. D) Potential solar radiation index. E) Curvature. F) NDVI. G) Log of flow accumulation. H) Illustration of the effect of the neighborhood size on the flow accumulation maps for original (ORI), 5 and 40 m neighborhood sizes in a portion of field 792. Values in the map are the numbers of contributing pixels.

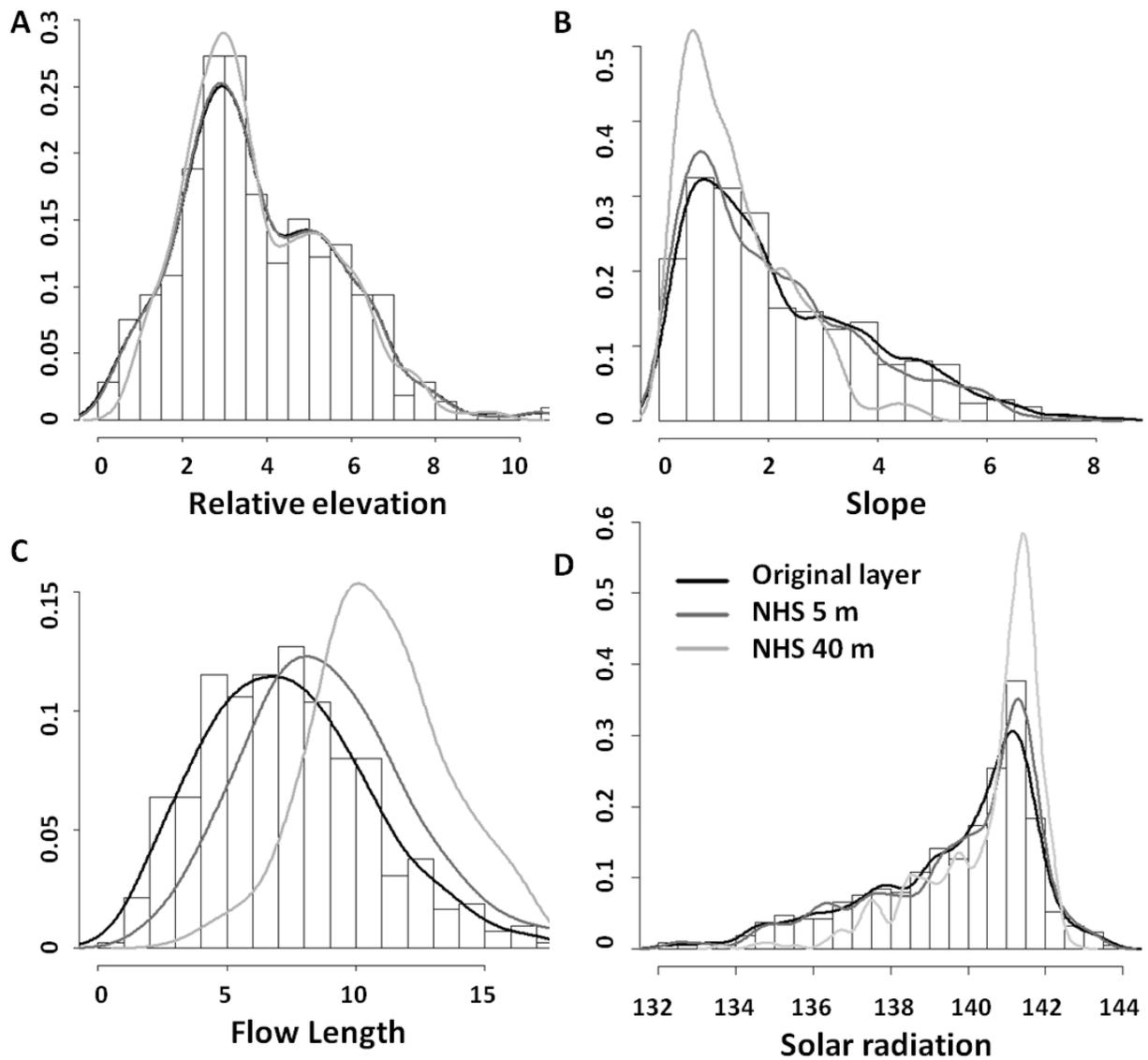
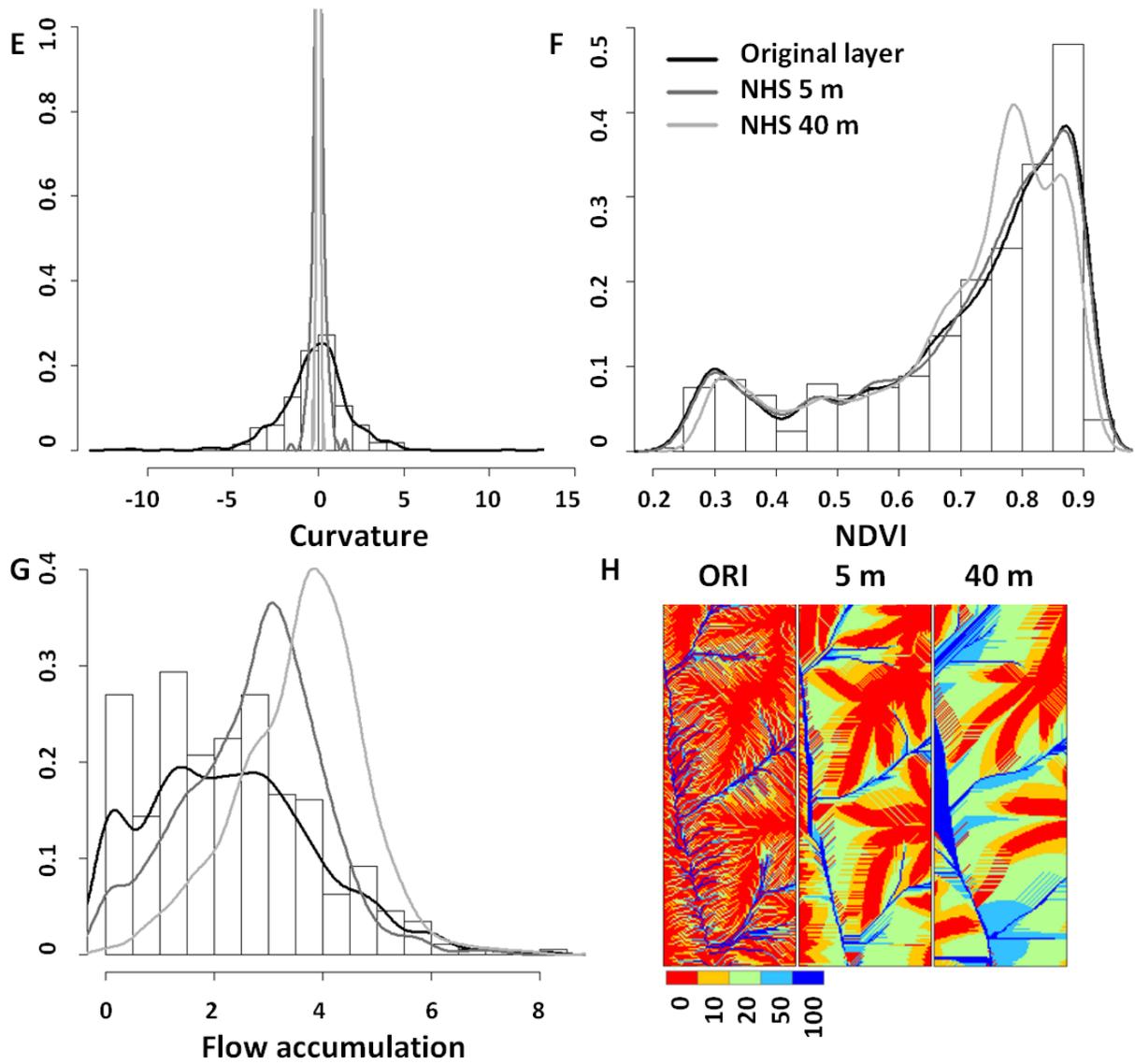


Figure 3.3 (cont'd)



To obtain a general quantitative overview of the changes in topographic attributes as a function of DEM filtering, correlation coefficients between the original layer and the filtered layers (i.e. 1, 3, 5, 10, 20 and 40 m) were computed. Fig. 3.4 shows the average correlation coefficients between the original map and the filtered versions from all eight fields. As expected the correlations with the original layer were higher at smaller *NHS* and decreased with increasing *NHS*. The topographic attributes for which the correlation with the original layer decreased the least included relative elevation, slope and potential solar radiation (Fig. 3.4A). The correlations for relative elevation remained high across the whole range of filtered layers ($r = 0.95$), indicating little change in elevation pattern across studied *NHS* values. Both slope and potential solar radiation were more sensitive to DEM smoothing than relative elevation. They were closely correlated to the original DEM up to $NHS = 10$ m ($r = 0.95$), after which the correlation decreased to 0.68 in slope and to 0.72 in potential solar radiation (i.e. $NHS = 40$ m). Conversely, for flow accumulation, flow length, curvature and the wetness index, the correlation coefficients dropped to values lower than 0.4 already when comparing the original DEM layer to the next filtered layer (i.e. map derived from 1 m *NHS*) (Fig. 3.4B). The drop was less drastic for flow length. Filtered versions of flow length remained closely correlated to the original map up to 3 m *NHS* ($r = 0.78$); however, the correlation dropped substantially at 20 and 40 m *NHS* ($r = 0.5$). These results indicate that substantial care should be taken when using the four DEM derivatives that are sensitive to smoothing.

It is obvious that differences in their values and patterns induced by different smoothing settings will lead to different outcomes in terms of prediction and estimation of water and material redistribution through the studied fields and to differences in relating topography to plant growth parameters. It is possible that original maps might provide the most accurate

representation of the fluxes and redistribution patterns on a small scale, while smoother maps may provide a better representation of the fluxes and patterns on a bigger scale. It is also likely that optimal smoothing settings might vary by the response variable of interest specific to the within-field fluxes and patterns. For example the maps that are less smooth might be necessary for accurate prediction of vegetation patterns, while a smoother map might be better in predicting the soil moisture within a field.

3.3.2. Effect of scaling on the sensor-based variable (NDVI)

The changes in *NDVI* with increasing *NHS* were relatively minor, and became noticeable only for the largest neighborhood size, i.e. 40 m. Filtering *NDVI* using a 40 m radius of neighborhood reduced the range of values to 96% of the original range (Fig. 3.3F). The correlation between the original layer and filtered layers were generally high, except for the largest *NHS* (40 m) where correlation dropped to ca. 0.7 (Fig. 3.4A). There were substantial variations among the fields, for example, filtering affected *NDVI* maps the most in fields 97 and 791, and was less pronounced in fields 52 and 87 (Data not shown). This could be related to field specific spatial patterns in *NDVI* distributions. For example, relatively noisy random patches were observed on original maps of *NDVI* values in fields 97 and 791 (Fig. 3.2), while *NDVI* varied more gradually through fields 52 and 87.

Figure 3.4. Mean correlation coefficients between the studied variable values obtained at different neighborhood sizes and the original layer. Error bars indicate the standard deviation across the fields. A) Variables with minor decrease in the correlation coefficient (i.e. relative elevation, slope, solar radiation and NDVI). B) Variables with major decrease in the correlation coefficient (i.e. flow accumulation, flow length, curvature and wetness index).

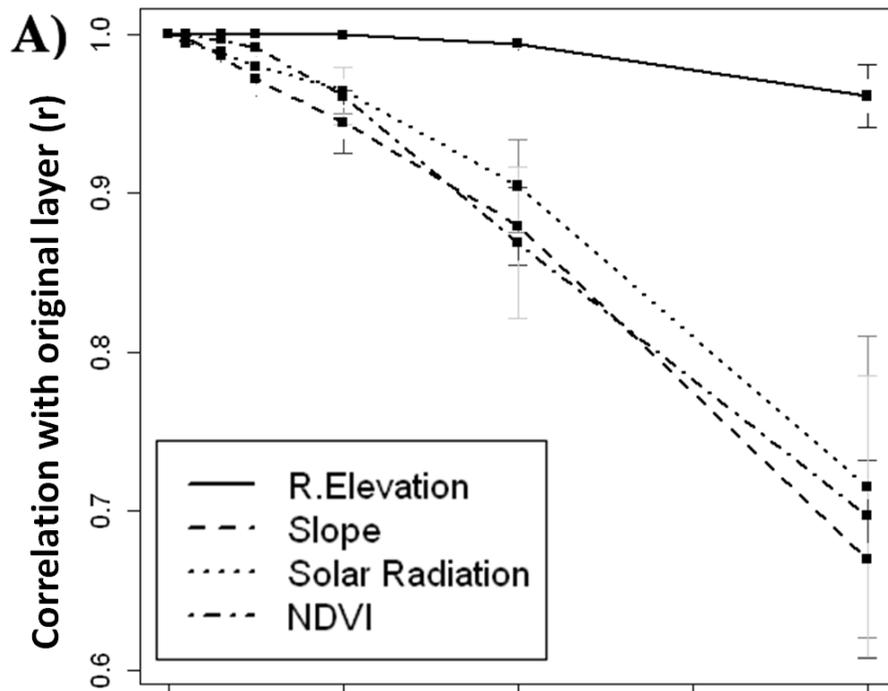
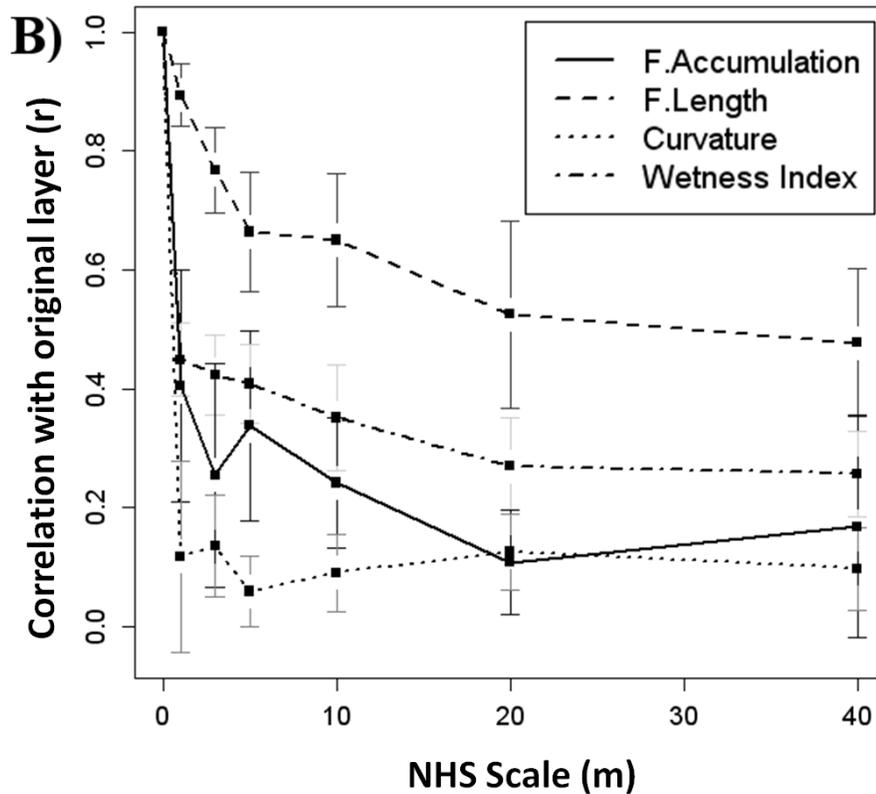


Figure 3.4. (cont'd)



3.3.3. Relationship between topographic factors and NDVI at different scales

Table 3.2 shows a list of the topographic attributes that were selected as statistically significant predictors for *NDVI* in the multiple regression analysis for each *NHS* value in the studied fields. Relative elevation, slope and the potential solar radiation index were among the significant predictors most often, with relative elevation being a significant predictor in at least one of the *NHS* values in all eight fields, slope in seven fields and the potential solar radiation index in five fields. Flow length, curvature and the wetness index tended to become significant predictors at larger *NHS* values (i.e. 5 to 40 m), indicating that importance of different topographic variables for cover crop biomass prediction varies with scale. For example, flow

length was a significant predictor for $NHS > 20$ m in six of the studied fields, while in only one of the fields it was a significant predictor at the original and $NHS = 1$ m levels. In field 30, relative elevation and slope captured the patchiness in $NDVI$ in the original layer, however the prediction performance increased as I included flow length, potential solar radiation and flow accumulation (Table 3.2), i.e., the variables that seemed to explain $NDVI$ variability better at larger NHS values.

Table 3.2. List of topographical variables that were statistically significant predictors ($p < 0.05$) for *NDVI* in multiple regression analysis for each neighborhood size and each studied field.

Layers	FIELD 30							FIELD 38							FIELD 52							FIELD 87								
	<i>RE</i>	<i>SL</i>	<i>FL</i>	<i>FA</i>	<i>CU</i>	<i>WI</i>	<i>SR</i>	<i>RE</i>	<i>SL</i>	<i>FL</i>	<i>FA</i>	<i>CU</i>	<i>WI</i>	<i>SR</i>	<i>RE</i>	<i>SL</i>	<i>FL</i>	<i>FA</i>	<i>CU</i>	<i>WI</i>	<i>SR</i>	<i>RE</i>	<i>SL</i>	<i>FL</i>	<i>FA</i>	<i>CU</i>	<i>WI</i>	<i>SR</i>		
Original	x	x							x					x		x								x						x
NHS 1	x	x						x	x					x		x				x				x						x
NHS 3		x	x					x	x				x	x		x				x				x						x
NHS 5		x	x					x	x					x		x								x						x
NHS 10		x	x	x					x					x							x			x					x	x
NHS 20		x	x	x	x			x	x					x	x		x							x					x	x
NHS 40	x		x	x				x	x	x				x	x									x				x		x
	FIELD 93							FIELD 97							FIELD 791							FIELD 792								
	<i>RE</i>	<i>SL</i>	<i>FL</i>	<i>FA</i>	<i>CU</i>	<i>WI</i>	<i>SR</i>	<i>RE</i>	<i>SL</i>	<i>FL</i>	<i>FA</i>	<i>CU</i>	<i>WI</i>	<i>SR</i>	<i>RE</i>	<i>SL</i>	<i>FL</i>	<i>FA</i>	<i>CU</i>	<i>WI</i>	<i>SR</i>	<i>RE</i>	<i>SL</i>	<i>FL</i>	<i>FA</i>	<i>CU</i>	<i>WI</i>	<i>SR</i>		
Original	x	x	x					x	x							x						x	x	x						
NHS 1	x	x	x	x			x		x			x	x		x						x	x	x						x	
NHS 3	x	x	x			x		x	x			x	x		x							x	x							
NHS 5	x	x	x			x		x	x			x										x	x	x			x			
NHS 10	x	x	x			x		x	x			x			x						x	x	x	x			x			
NHS 20		x				x	x	x	x			x			x						x	x	x	x			x		x	
NHS 40	x	x	x	x			x	x	x	x					x						x	x	x	x			x		x	

RE = Relative elevation. *SL* = Slope. *FA* = Flow accumulation. *FL* = Flow length. *CU* = Curvature. *SR* = Solar radiation. *WI* = Wetness index.

Original layers are raw-data derived layers at 1 m resolution without any filtering.

NHS (Neighborhood size): Layers derived from a filtered DEM and *NDVI*. Filters used are 1, 3, 5, 10, 20, and 40 m in radius.

Fig. 3.5A shows the prediction performance of multiple regression models using only significant predictors; the adjusted- R^2 using the original layer ranged from 0.10 (field 52) to 0.56 (field 97). The performance of topographic attributes as a set of predictors for *NDVI* varied considerably with the field; the best performance was observed in fields 97 and 792 (adjusted- $R^2 > 0.45$) and the worst performance was observed in fields 52 and 791 (adjusted- $R^2 < 0.15$). However, the prediction performance in field 52 (i.e. the flattest field in my study, Fig. 3.2) was considerably lower across all *NHS* values, indicating that topography was not an important factor of *NDVI* distribution throughout this field (Fig. 3.5A). Adjusted- R^2 values increased with increasing *NHS* in all fields (Fig. 3.5A), except for field 52. The increase was relatively small as *NHS* increased from original, to 1–3 m level, becoming greater at larger *NHS* values (i.e. 10 to 40 m). At larger *NHS* the performance of topography as a predictor of *NDVI* could be regarded as acceptable, i.e. adjusted- $R^2 > 0.5$, in all fields, except 38 and 791.

Changes in adjusted- R^2 values appeared to be related to the changes in significant predictors. For example, in the field 87, adjusted- R^2 increases considerably at larger *NHS* values as curvature and the wetness index become significant predictors. In field 792, adjusted- R^2 increased considerably as flow length, curvature and potential solar radiation were included in the model at *NHS* = 10, 20 and 40 m (Fig. 3.5A, Table 2). A similar result was reported by Bian and Walsh (1993) in the relationship between topographic attributes and a vegetative index in a regional study. Using *NHS* from 30 to 3300 m, they showed that elevation was a significant predictor for the vegetative index at all *NHS* values; however, the individual R^2 for slope

increased from 0.17 at $NHS = 30$ m to 0.53 at $NHS = 3300$ m, indicating that slope becomes an important predictor at larger NHS values. This observation points to the effect of the neighborhood size when evaluating contributions of different topographic attributes as predictors of plant characteristics, since different outcomes and conclusions could be obtained at different neighborhood sizes. The effect also should be considered while comparing the results from different studies.

Figure 3.5. Adjusted-R2 for multiple regression models. (A) Adjusted-R2 as a function of neighborhood size in each individual field. B) Mean adjusted-R2 for multiple regression from all studied fields as a function of neighborhood size. Different uppercase letters mark significant differences between topographic scales ($p < 0.05$). Different lowercase letters mark significant differences between NDVI scales ($p < 0.05$). Values presented in (A) are for the cases where NDVI and topography were filtered at the same neighborhood sizes. Only significant predictors were used in the multiple regression models ($\alpha = 0.05$).

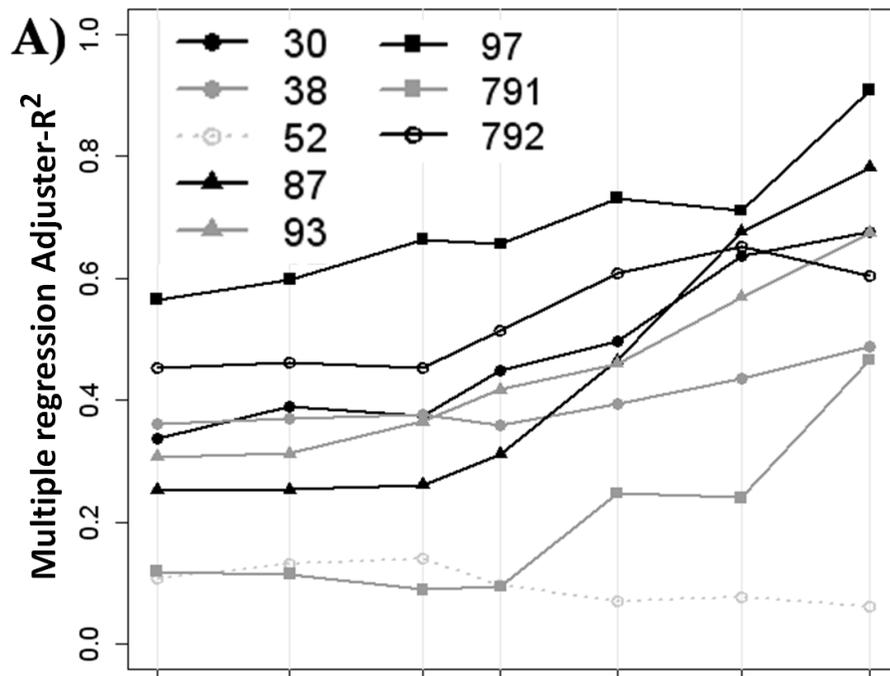
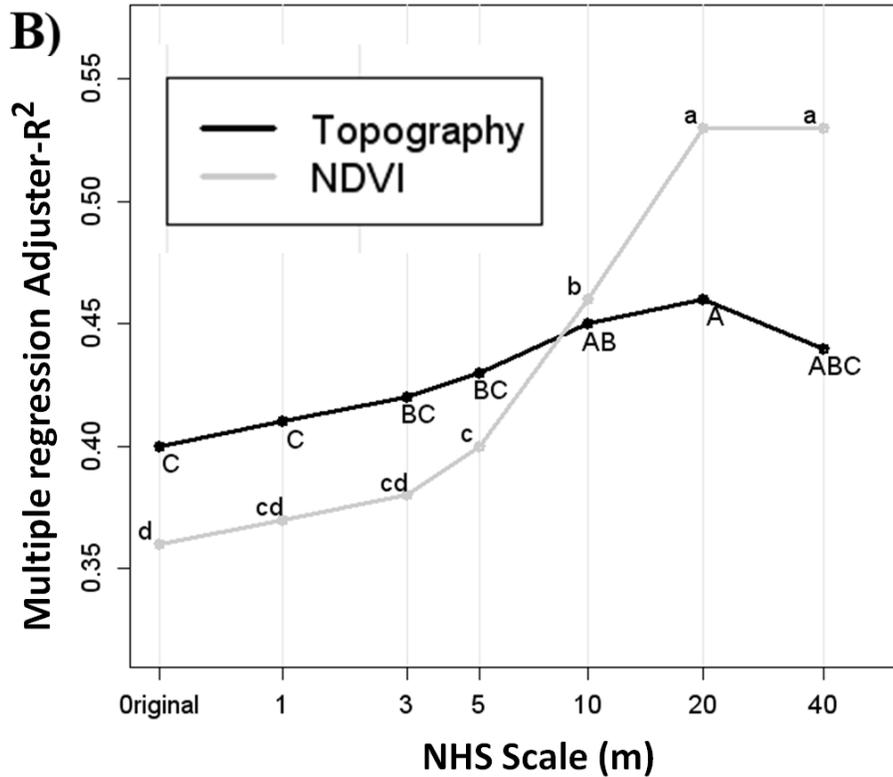


Figure 3.5. (cont'd)



In order to find the *NHS* values of topographic attributes and *NDVI* (hereafter called topographic *NHS* and *NDVI-NHS*) that maximize the prediction of *NDVI* from topography, I conducted analysis of variance with topographic *NHS* and *NDVI-NHS* as two studied factors and adjusted- R^2 as the response variable. The interaction between topographic *NHS* and *NDVI-NHS* was not significant, indicating that the effect of topographic *NHS* did not depend on *NDVI-NHS*. Fig. 3.5B shows the mean values of adjusted- R^2 for topography across all *NDVI-NHS* values and for *NDVI* across all topographic *NHS* values. The highest adjusted- R^2 was observed at *NHS* of 20 m, though it was not significantly different from *NHS* values of 10 and 40 m. The highest

adjusted- R^2 values in *NDVI* were observed at *NHS* of 20 and 40 m. Thus using these *NHS* values for both topography and *NDVI* yielded better predictions across all the fields. Bian and Walsh (1993) also reported increases in R^2 from multiple regression models with increasing *NHS*, reaching an optimal R^2 value at *NHS* = 2250 m. They defined the concept of “characteristic” scale as the *NHS* value at which the spatial dependencies of topography and the vegetative index converge. In my study the “characteristic” scale between topography and *NDVI* ranged from 20 to 40 m.

3.3.4. Indicators of optimal topographic scale

The topographic features of each field (i.e. flat, rolling or steep) can potentially serve as an indicator of the usefulness of topographic attributes as explanatory variables for cover crop biomass prediction. For example, fields 30, 87 and 792 had diverse topographic patterns, particularly depression areas and undulating terrain (Fig. 3.2), with wider ranges of relative elevation, slope and potential solar radiation. In these fields the choice of topographic *NHS* most significantly affected the performance of the multiple regression models. Conversely, fields 38 and 52 had lower topographic diversity and performance of regression models was not substantially affected by changes in *NHS*. Overall, I observed positive correlations between the mean adjusted- R^2 and the range of relative elevation, slope and potential solar radiation in each field (Fig. 3.6A).

I explored the use of principle components of the topographic attributes as a potential indicator of the fields where topography can be expected to play a significant role as a cover crop biomass predictor. Three principal components, PC1, PC2, and PC3, were extracted from

topographic attributes. PC1 explained 31.7% of the total variability, PC2 explained 23.5%, whereas PC3 explained 20.0%. The major contributors in PC1 were potential solar radiation and slope. The major contributors in PC2 were relative elevation and flow accumulation. The major contributor in PC3 was curvature. The average value of the scores in each field for PC1 showed a larger range as compared to the scores in PC2 and PC3, indicating that topographic differences between fields were mainly explained by PC1. I present the PC1 scores as a measure of topographic “complexity”. The mean adjusted- R^2 values from regression analyses in each field and the averaged PC1 score were significantly correlated indicating that PC1 was a good indicator of the strength in the relationship between topography and *NDVI* (Fig. 3.7). The fields with lower adjusted- R^2 in the multiple regressions tended to have higher PC1 values; those were the fields with flat topography and high values of the potential solar radiation index. Fig. 3.7 also shows the change in the mean adjusted- R^2 of regression models after filtering the original layers. For all the fields there was a positive change in the adjusted- R^2 which was also significantly correlated to the PC1 scores, indicating that a better performance in regression models can be obtained when filtering topographic attributes in fields with high slope values. On the other hand, filtering fields with flat topography and high potential solar radiation will not increase the performance of regression models.

Among the variogram characteristics the spatial autocorrelation range and the nugget/sill ratio appeared to be most promising indicators of the optimal neighborhood size. The range parameter reflects the maximum distance of spatial autocorrelation. The nugget/sill ratio is a measure of the proportion of micro-variance and/or experimental error in respect to the total variance. Small values of this ratio indicate that a significant proportion of the variance can be

explained by the spatial dependence, whereas only a minor proportion is attributed to fine-scale variation and experimental error. After examining the variograms of all studied topographic attributes (results not shown) I decided to use variograms of slope to summarize the spatial autocorrelation in topography as it was the most significant attribute across scales and fields. Slope variogram models for all fields showed a nested structure that could be typically described by two spherical models with the first range parameter varying between 30 and 49 m; and the second range parameter varying between 77 and 140 m. Variograms for *NDVI* also showed spherical nested structures in all fields with the first range parameter between 9 and 63 m and the second range parameter between 98 and 169 m.

Figure 3.6. Indexes for selecting optimum neighborhood size. A) Linear regression between multiple regression adjusted-R² and the standardized ranges of solar radiation, relative elevation and slope in the studied fields. B) Linear regression between the change in adjusted-R² after filtering and semivariogram nugget/sill ratios from NDVI maps of the studied fields. C) Linear regression between the change in adjusted-R² after filtering and the semivariogram nugget/sill ratio from slope maps of the studied fields. D) Optimal neighborhood size for topography plotted versus the partial sill value from semivariograms of the slope maps.

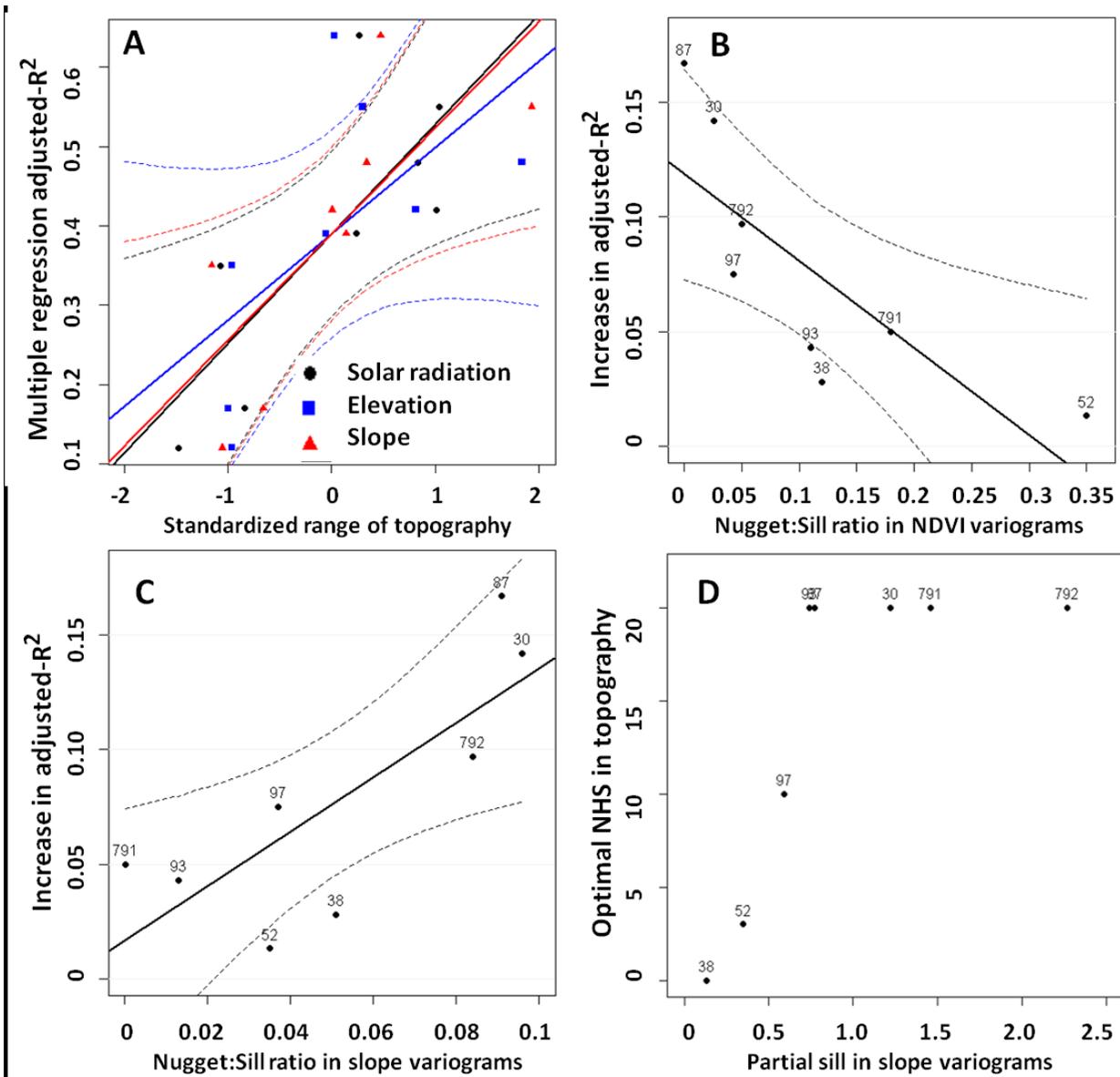
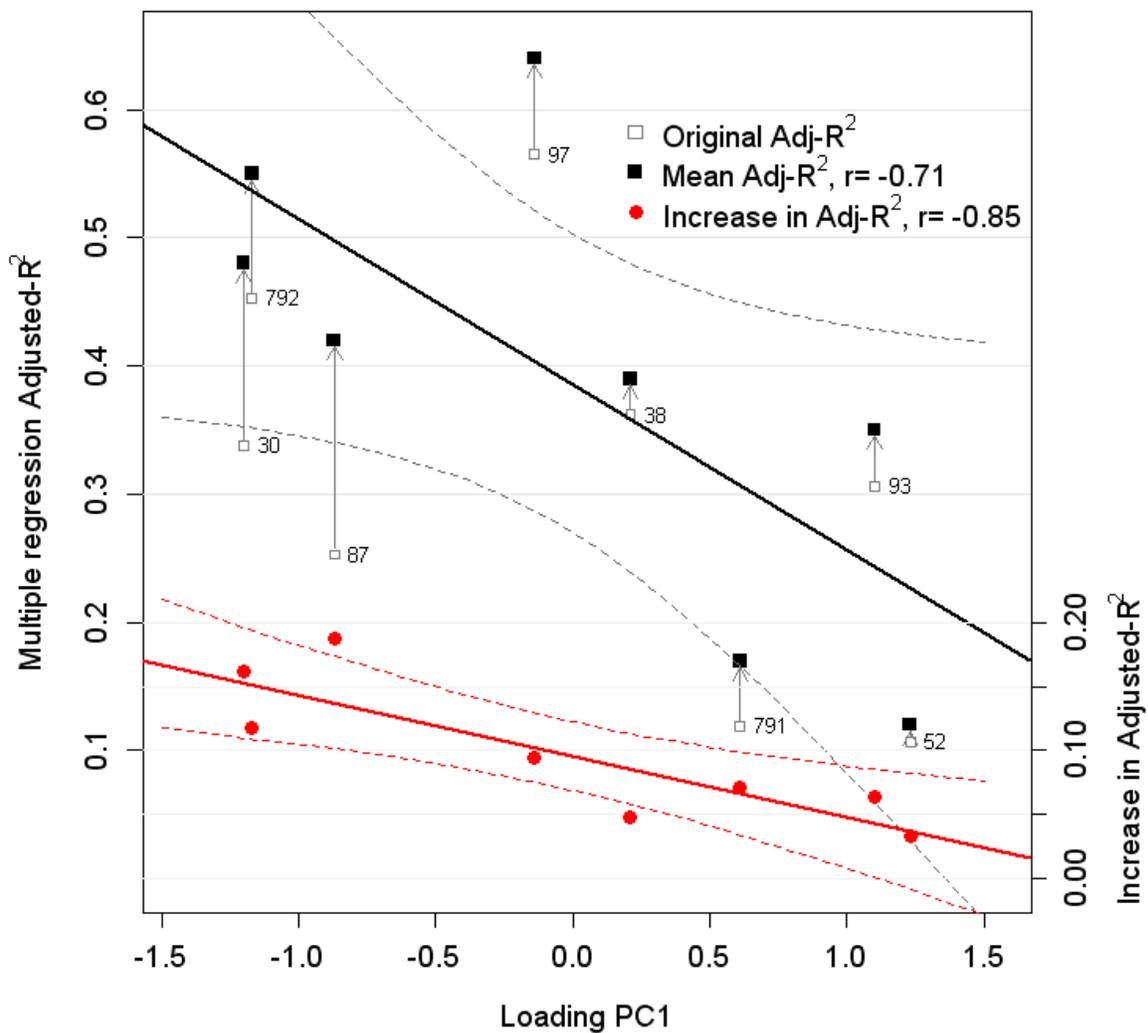


Figure 3.7. Correlation between score values of the first principal component (PC1) and the mean adjusted-R² values from multiple regression analyses. Black squares: Before filtering. Fields with relatively flat topography had positive values of PC1, while fields with more diverse topography had negative values of PC1. Red circles: After filtering. Adjusted-R² values from multiple regressions in the original layers were plotted as a reference with their respective change values (White squares). Numbers indicate the field labels.



Changes in adjusted- R^2 values were positively related to the nugget/sill ratios of both slope and *NDVI* variograms (Fig. 3.6B,C). Larger values in the ratio for the slope variogram were associated with larger gain in the adjusted- R^2 from the multiple regression models after filtering (Fig. 3.6C). Large values of this ratio indicate a greater topographic “complexity” which can be expected to have a greater effect on spatial patterns in *NDVI*. On the other hand, the ratio for *NDVI* was negatively related to the change in adjusted- R^2 (Fig. 3.6B), indicating that *NDVI* could be better explained by topography when the ratio was small, i.e., in the cases with overall smooth spatial patterns in *NDVI*. Nugget/sill ratios in both slope and *NDVI* could serve as a potential indicator of those cases in which filtering can improve the performance of the predictive models.

Although across all fields the best adjusted- R^2 values were found in *NHS* values ranging from 20 to 40 m, the optimal *NHS* value in each individual field varied considerably. The best possible combination of topographic *NHS* and *NDVI-NHS* was different for each field, influenced by the field specific topographic “complexity”, e.g. reflected in the PC1 scores (Fig. 3.7). These results suggest that there is no single optimal scale for all fields, thus the selection of *NHS* might be field specific. Similar results were also reported by Smith et al., (2006) in a study that used topographic attributes in soil mapping. They showed that the neighborhood sizes for topographic derivation varied from landscape to landscape.

The optimal *NHS* value for filtering was also reflected in the partial sill parameter in the first spherical structure of slope variograms (Fig. 3.6D). This parameter shows the proportion of

the total variability explained by autocorrelation at ranges lower than 49 m. Fields with larger proportions of variance explained by slope autocorrelation required larger *NHS* values to match the spatial variation in *NDVI*. On the other hand, filtering did not have a noticeable impact in fields with small proportion explained by the autocorrelation in slope at distances <49 m.

3.4. Conclusions

I studied the effect of the neighborhood size derivations of topographic data and *NDVI* for red clover cover crop on the strength of the relationship between them. Curvature, flow accumulation and flow length were more affected by the scale of derivation than relative elevation, slope and the potential solar radiation index. Relative elevation, slope and potential solar radiation were significant predictors for *NDVI* at most studied scales. Conversely, flow accumulation, curvature, flow length and the wetness index were significant predictors of *NDVI* mainly at large scales. The strength of the relationship between topographic attributes and *NDVI* readings varied with the scale. Increasing the neighborhood size in topographic and *NDVI* data derivations tended to increase the strength of the relationship. Among the tested scales, i.e., 1, 3, 5, 10, 20 and 40 m, the scales 20 and 40 m led to the regression models with better prediction performance. I conclude that using the original DEM for analyzing the relationships between topography and biophysical variables, e.g., cover crop biomass, may not always be appropriate, since not all possible associations with topographic attributes are depicted at the small scale of the original DEM. Thus, prior to developing predictive regression models I recommend exploring the relationships between topography and biophysical variables of interest at a range of scales.

Topographic “complexity” of the studied terrain expressed as the range of the topographic variable values, the loading values of principal components, and the semivariogram parameters of the topographic and studied biophysical variables can serve as aids in such exploration. I found that slope and the potential solar radiation index were the most relevant indicators of the topographic “complexity” in the studied agricultural fields, thus these attributes could be used to identify the fields where topography could be successfully used as a predictor of cover crop biomass. Partial sill and nugget/sill ratios in terrain slope and *NDVI* semivariograms were good indicators of the optimal derivation scale. Larger nugget/sill ratios for slope and smaller ratios for *NDVI* variograms were associated with high regression prediction performance, while smaller ratios for slope and larger ratios for *NDVI* yielded poor predictions.

I showed that topographic effects on growth and biomass production of cover crops are most pronounced at certain spatial scales, thus predictive models that use topographic attributes will be most accurate when used at the optimal scales. In this study the effect of scale of derivation was demonstrated for the relationship between topography and cover crop biomass; however, a similar approach can be utilized when relating topography with hydrological, soil and other biophysical variables.

REFERENCES

REFERENCES

- Bian, L., Walsh, S., 1993. Scale dependencies of vegetation and topography in a mountainous environment of Montana. *Prof. Geogr.* 45, 1-11.
- Bolstad, P., Stowe, T., 1994. An evaluation of DEM accuracy: Elevation, slope and aspect. *Photogram. Eng. and Remote Sen.* 60, 1327-1332.
- Crum, J., Collins, H., 2012. KBS Soils. KBS LTER site description, Available: http://lter.kbs.msu.edu/about/site_description/soils.php (Accessed: 2012, Jun 20)
- Curran, P., Atkinson, P., 1999. Issues of scale and optimal pixel size. In: Stein, A., van Der Meer, F., Gorte, B., (Eds.), *Spatial Statistics for Remote Sensing*. Kluwer Academic Publish., Dordrecht, pp. 115-134.
- Dobermann, A., Goovaerts, P., Neue, H., 1997. Scale-dependent correlations among soil properties in two tropical lowland rice fields. *Soil Sci. Soc. Am. J.* 61, 1483-1496.
- ESRI, 2003. ArcGIS-Desktop v9.2. USA [software], <http://www.esri.com/software/arcgis/arcgis-for-desktop/index.html> (Accessed: 2011, Jan 16).
- Erskine, R., Green, T., Ramirez, J., MacDonald, L., 2007. Digital elevation accuracy and grid cell size effects on estimated terrain attributes. *Soil Sci. Soc. Am. J.* 71, 1371-1380.
- Fraisse, C., Sudduth, K., Kitchen, N., 2001. Delineation of site-specific management zones by unsupervised classification of topographic attributes and soil electrical conductivity. *Trans. ASAE.* 44, 155-166.
- Florinsky, I., Kuryakova, G., 1996. Influence of topography on some vegetation cover properties. *Catena* 27, 123-141.
- Florinsky, I., 1998. Accuracy of local topographic variables derived from digital elevation models. *Int. J. Geo. Inf. Sci.* 12, 47-61.
- Florinsky, I., Kuryakova, G., 2000. Determination of grid size for digital terrain modelling in landscape investigations-exemplified by soil moisture distribution at a micro-scale. *Int. J. Geogr. Inf. Sci.* 14, 815-832.
- Flynn, E., Dougherty, C., Wendroth, O., 2008. Assessment of pasture biomass with the Normalized Difference Vegetation Index from active ground-based sensors. *Agron. J.* 100, 114-121.
- Fu, P., Rich, P., 2002. A geometric solar radiation model with applications in agriculture and forestry. *Comp. Elec. Agri.* 37, 25-35.

- Gallant, J., Hutchinson, M., 1997. Scale dependence in terrain analysis. *Math. Comput. Simulat.* 43, 313-321
- Goovaerts, P., Webster, R., 1994. Scale-dependent correlation between topsoil copper and cobalt concentrations in Scotland. *Eur. J. Soil Sci.* 45, 79-95.
- Green, T., Erskine, R., 2004. Measurement, scaling, and topographic analyses of spatial crop yield and soil water content. *Hydrol. Proc.* 18, 1447-1465.
- Hengl, T., 2006. Finding the right pixel size. *Comput. Geosci.* 32, 1283-1298.
- Huang, X., SenthilKumar, S., Kravchenko, A., Thelen, K., Qi, J., 2007. Total carbon mapping in glacial till soils under near-infrared spectroscopy Landsat imagery and topographical information. *Geoderma* 141, 34-42.
- Huang, X., Wang, L., Yang, L., Kravchenko, A., 2008. Management effects on relationships of crop yields with topography represented by wetness index and precipitation. *Agron. J.* 100, 1463-1471.
- Johnson, R., Wichern, D., 2002. *Applied Multivariate Statistical Analysis*, Ed. 5, Prentice Hall. Englewood Cliffs, NJ. 767 p.
- Kravchenko, A., Bullock, D., 2000. Correlation of corn and soybean grain yield with topography and soil properties. *Agron. J.* 92, 75-83.
- Kravchenko, A., Robertson, G., Thelen, K., Harwood, R., 2005. Management, topographical, and weather effects on spatial variability of crop grain yields. *Agron. J.* 97, 514-523.
- Lark, R., 1998. Forming spatially coherent regions by classification of multivariate data: An example from the analysis of maps of crop yield. *Int. J. Geogr. Inf. Sci.* 12, 83-98.
- Liu, X., 2008. Airborne LiDAR for DEM generation: some critical issues. *Prog. Phys. Geog.* 32, 31-49.
- McBratney, A., Mendoca-Santos, M., Minasny, B., 2003. On digital soil mapping. *Geoderma* 117, 3-52.
- McMillan, R., Pettapiece, W., Watson, L., Goddard, T., 1998. A landform segmentation model for precision farming. In: Robert, P., Rust, R., Larson, W. (Eds.), *Proc. 4th Int. Conf. Precision agriculture*. ASA, CSSA, and SSSA, Madison, USA. pp. 1335-1346.
- McCann, B., Pennock, D., van Kessel, C., Walley, F., 1996. The development of management units for site specific farming. In: Robert, P., Rust, R., Larson, W. (Eds.), *Proc. 3rd Int. Conf. Precision Agriculture*. ASA, CSSA, and SSSA, Madison, USA. pp. 295-302.
- Milledge, D., Lane, S., Warburton, J., 2009. The potential of digital filtering of generic topographic data for geomorphological research digital filtering of generic topographic data for geomorphological research. *Earth Surf. Proc. Land.* 34, 63-74

- Moore, I., Gessler, P., Nielsen, G., Peterson, G., 1993. Soil attribute prediction using terrain analysis. *Soil Sci. Soc. Am. J.* 57, 443–452.
- Muñoz, J., Finley, A., Gehl, R., Kravchenko, S., 2010. Nonlinear hierarchical models for predicting cover crop biomass using Normalized Difference Vegetation Index. *Remote Sens. Environ.* 114, 2833-2840.
- Roecker, S., Thompson, J., 2010. Scale effects on terrain attribute calculation and their use as environmental covariates for digital soil mapping. In: Boettinger, J., Howell, D., Moore, A., Hartemink, A., Kienast-Brown, S. (Eds.), *Digital Soil Mapping: Bridging Research, Environmental Application and Operation*, Springer, Dordrecht, pp. 55-66.
- R Development Core Team. 2009. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org>.
- Ribeiro Jr. P., Diggle, P. 2001. geoR: A package for geostatistical analysis. *R-News* 1, No 2. <http://cran.r-project.org/doc/Rnews>
- Rizzo, M. 2008. *Statistical Computing with R*. Computer Science and Data Analysis Series. Chapman and Hall. Boca Raton, 399 p.
- SAS Institute. 2008. SAS user's guide. Version 9.2. SAS Inst., Cary, NC. USA
- Sheather, S., Jones, M., 1991. A reliable data-based bandwidth selection method for kernel density estimation. *J. Royal Stat. Soc. Series B* 53, 683–690.
- Simmons, J. 2012. KBS004: Agronomic Protocol: Main Cropping System Experiment. LTER research protocols, Available: <http://lter.kbs.msu.edu/protocols/104> (Accessed: 2012, Jun 20).
- Smith, M., Zhu, A., Burt, J., Stiles, C., 2006. The effects of DEM resolution and neighborhood size on digital soil survey. *Geoderma* 137, 58–69.
- Snapp, S., Swinton, S., Labarta, R., Mutch, D., Black, R., Leep, R., Nyiraneza, J., O'Neil, K., 2005. Evaluating cover crops for benefits, costs and performance within cropping system niches. *Agron. J.* 97, 322-332.
- Ticehurst, J., Cresswell, H., McKenzie, N., Glover, M., 2007. Interpreting soil and topographic properties to conceptualise hillslope hydrology. *Geoderma* 137, 279-292.
- Townsend, P., Walsh, S., 1996. Estimation of soil parameters for assessing potential wetness: Comparison of model responses through GIS. *Earth Surf. Proc. Land.* 21, 307-326
- Wu, W., Fan, Y., Wang, Z., Liu, H., 2008. Assessing effects of digital elevation model resolutions on soil–landscape correlations in a hilly area. *Agr. Ecosyst. Environ.* 126, 209–216.

Chapter 4: Cover crop effect on corn growth and yield in agricultural fields with diverse terrain

Abstract

The use of cover crops is reported to enhance agro-ecological services in rotational crop systems, however their adoption by farmers has remained limited. A challenge to farmer uptake is high spatial and temporal variability in cover crop establishment and growth. The spatial and temporal performance uncertainty reduces the potential for benefits to be derived from cover crops. Since topography plays an important role in spatial processes that ultimately affect plant performance, it could be used to quantify and predict cover crop spatial variability and cover crop contribution to a subsequent cash crop. In this 3-year study conducted in southwest Michigan I assessed the effects of topography and cover crop (red clover) biomass on corn yields. I used path analysis within a Bayesian framework to identify direct and indirect relationships among topography, red clover biomass, and corn yield, while taking into account the effects of agricultural management practices, multiple years, and multiple experimental fields. I observed that topographic attributes, e.g. terrain slope and flow accumulation, contribute significantly to explaining the variability in both red clover biomass and corn yields.

However, the cover crop was affected by topography differently than the main crop, corn. Higher red clover biomass was produced in flat locations, whereas higher corn yield was produced in areas with high curvature. Red clover biomass positively influenced corn yield, however, the magnitude of that effect varied both temporally and spatially. Notably, the positive

effect of red clover on corn yields was significant only in the years with average precipitation; and the magnitude of the effect was more pronounced at summit and slope topographical positions, than in topographical depressions. My results indicate that efforts to ensure a good cover crop stand will be most beneficial to subsequent corn crop at summit and slope positions. Accounting for variability in fields and years has significantly improved analysis of the interactive relationships between topography, red clover, and corn yields.

Keywords: cover crops, agro-ecological services, topography, hierarchical models, Bayesian path analysis

4.1. Introduction

The use of cover crops in row crop systems has been reported to improve agro-ecological services provided by the ecosystems including enhanced carbon (C) sequestration, reduced erosion, increased nutrient supply, increased weed suppression, and reduced losses of agrochemicals (Snapp et al., 2005; Bhardwaj et al., 2011). It is assumed that these benefits mainly depend on the amount of cover crop biomass that is produced and incorporated into the soil (Kuo et al., 1997; Queen et al., 2009; Fageria, et al. 2011). For example, it has been demonstrated that cover crops, when incorporated into the soil, can increase or maintain the concentration of soil C and nitrogen (N) (McVay et al., 1989; Kuo and Sainju, 1998; Sainju et al., 2002; Mazzoncini et al., 2011), and improve nutrient use efficiency in agricultural systems (Fageria et al., 2011; Dabney et al., 2012). Cover crop use has also been reported to influence soil physical conditions and water retention (Drury et al., 2003; Papadopoulus et al., 2006). For example, cover crops reduce bulk density and increase porosity (Ess et al., 1998); enhance root activity and C inputs, which improve soil aggregation (Miller and Dick, 1995); improve soil

hydraulic conductivity and infiltration by modifying soil structure and aggregate stability (Murphy et al., 1993); and preserve soil moisture (Gallager, 1977; Smith et al., 1987).

Use of leguminous cover crops can be particularly advantageous. Benefits of using legumes include increased N supply and reduced amount of N fertilizer required for the next crop (Meisinger et al. 1991), increased microbial biomass (Jokela, et al., 2009), and increased N mineralization rates (Dinesh et al., 2001). In addition, legumes possess the ability to absorb scarce nutrients in the soil profile. Therefore, legume biomass incorporation will increase nutrient concentration in the surface layer of the soil (Fageria et al., 2005). A meta-analysis of the response of corn yield to cover crops was provided by Miguez and Bollero (2005), who analyzed the results from 36 peer-reviewed publications and concluded that legume cover crops increase average corn yield by 37%. Red clover (*Trifolium pratense L.*) is among the most commonly used legume cover crops in the northeastern United States (Singer and Cox, 1998), and has been shown to provide up to 85 kg N ha⁻¹ for a subsequent crop (Vyn et al., 1999). Hively and Cox (2001) found that the use of red clover resulted in higher corn yield compared to a no-cover crop control. Sarrantonio and Molloy (2003) reported that red clover used in rotational systems on sandy loam soils improved corn yield parameters, i.e., plant height, plant biomass, and yield.

Despite all their benefits, adoption of cover crops by farmers remains low, and only a very small percentage of the U.S. cropland is planted with cover crops (Dabney et al., 2012). A possible explanation for the low adoption rate is that cover crop establishment, growth, and subsequent nutrient supply to a cash crop are spatially and temporally variable. High spatial and temporal variability increases management challenges of cover crop-based systems.

High spatial variability of legume and non-legume cover crop biomass within a single

field has been reported in several studies (Boyer et al., 1996; Harmony et al., 2001; Guretzky et al., 2004; Guretzky et al., 2005). Legumes in particular are known to have high spatial variability in N-fixation patterns in addition to variable biomass production (Boyer et al., 1996; Nykanen et al., 2008; Hauggaard-Nielsen et al., 2010). Soil properties may contribute to this variability, Boyer et al. (1996) found a positive moderate correlation ($r=0.25$) between biomass production and soil pH. Hauggaard-Nielsen et al. (2010) found that biomass production was correlated to soil humus content and total N in the topsoil. They also found that N-fixation at flowering stage was negatively correlated to clay content (25-75 cm); however, the authors indicate that it was not possible to build a satisfactory regression model to explain biomass and N-fixation due to the large spatial variability observed in both variables.

Topography plays an important role in the spatial distribution of soil particles (erosion/deposition), organic matter, nutrients, and hydrologic conditions throughout the landscape and has been shown to be correlated with a variety of soil properties. For instance, topographic attributes have been used to explain the spatial distribution of soil water content (Kang et al., 2003; Green and Erskine, 2004; Ticehurst et al., 2007; Zhu and Lin, 2011), hydraulic conductivity (Jiang et al. 2007), soil nutrient content (Aandahl, 1948; Moore et al., 1993; Wang et al., 2009; Zhang et al., 2011), soil carbon content (Gregorich et al., 1998; Terra et al. 2005; Ritchie et al., 2007), soil respiration (Kang, et al., 2003), and soil temperature and microclimate conditions (Kang et al., 2000; Kang et al., 2003; Bennie et al., 2008).

The effect of topography on performance of row crops has been studied extensively and found to be substantial across diverse cropping systems in the U.S. (Simmons et al., 1989; Stone et al., 1985; Halvorson and Doll, 1991; Kravchenko and Bullock, 2000; Kravchenko and Bullock, 2002; Jiang and Thelen, 2004). In row crops in the Midwestern U.S., studies have

shown a significant effect of topography on the spatial distribution of several crop characteristics, in particular crop yield (Kravchenko and Bullock, 2000; Kravchenko and Bullock, 2002; Jiang and Thelen, 2004; Kravchenko et al., 2005; Huang et al., 2008). Considerably less attention has been devoted to the effects of topography on performance of cover crops (Harmony et al, 2001; Guretzky et al., 2004). Since spatial and temporal variability in cover crop performance is one of the factors that limit cover crop use by farmers, a better understanding of the effect of topography would contribute to more effective cover crop implementation, allowing producers to tailor recommendations to site-specific features of their fields.

Field sites used for agricultural experiments have been traditionally placed on flat land with homogeneous soil properties. This allows researchers to minimize variability in crop growth and performance conditions external to the study and to maximize detection of the effects of the agricultural management treatments considered in the study. However, such experimental settings also limit the ability of traditional field research experiments to address the role of topography and edaphic factors on variability in crop performance as well as the role of interactions between the studied agricultural management practices and other factors affecting crop growth. Such interactions can potentially change the practical outcomes of the research findings, e.g., when one management practice can outperform another under some but not other soil/terrain settings. This is particularly relevant to commercial agriculture that uses large fields with diverse topographic and soil conditions which make crop production highly spatially variable. The concern about the role of management vs. terrain interactions is especially valid for managements involving cover crop use. In commercial cropping systems it may be difficult to see the positive effect of cover crops on row-crop yields because variations in soil/terrain settings

could mask the cover crop effects. I hypothesize that in the systems with cover crops, the main row crop yield will be affected by topography both directly, through topography's role in water redistribution and spatial patterns of soil properties, and indirectly, through the contribution that topography makes to spatial patterns in biomass inputs by cover crops. While general effects of topography on plant growth cannot be doubted, the magnitudes, temporal patterns, and synchronies between these direct and indirect topographical influences on main row crops in cover crop based systems have not been addressed before.

To assess direct and indirect effects of topography on row crop yields I will use Path Analysis (PA), a procedure that allows testing direct and indirect relationships (“paths”) between topography, red clover and corn yield (Hoyle, 1995). With PA I can study the contribution of cover crop biomass to the main crop while accounting for confounding effects of topographic attributes (Gajewski, et al. 2006). I propose the implementation of PA under the Bayesian framework as it is particularly advantageous for inferential purposes in studies with multiple factors within a hierarchical structure. For example, implementation of linear models could be challenging in agricultural studies with several topographical positions within fields, and then diverse fields within management systems and years. In addition under this framework, I can generate estimates and standard errors that are not based on asymptotic theory, and therefore do not rely on the normality assumption (Gelman et al. 2004; Gajewski, et al. 2006). This is a desirable property because some variables, e.g. topographical features such as slope, flow accumulation, flow length, and curvature, are commonly not normally distributed; and because I want to estimate total effects (the sum of direct and indirect effects) and its associated standard errors which are particularly challenging to obtain for non-normal distributions (Sobel, 1982; Gajewski, et al. 2006; Preacher and Hayes, 2004).

4.2. Objectives

The main goal of my study is to examine the influence of cover crops on growth and yield of the following row-crop, i.e., corn, on a scale of a typical agricultural field. Since topography could affect growth and production of both cover and main crop, it becomes a potentially confounding factor that may mask the contribution of the cover crop, thus I aim at separating the partial effect of topography from the effect of the cover crop. The specific objectives of the study are 1) to determine the relative significance of direct and indirect effects of red clover biomass and topographic attributes on corn biomass and yield using hierarchical PA under a Bayesian framework; and 2) to quantify the effect of red clover biomass on corn biomass and yield at different landscape positions.

4.3. Materials and methods

4.3.1. Study site

The study was carried out at Kellogg Biological Station (KBS) located in southwest Michigan at 42° 24' N, 85° 24' W. Annual rainfall averages 890 mm/y and mean annual temperature is 9.7 °C. The dominant soil series are the Kalamazoo (fine-loamy, mixed, mesic Typic Hapludalfs) and Oshtemo (coarse-loamy, mixed, mesic Typic Hapludalfs). The data were collected from a total of ten experimental fields in three different years. A list of the fields along with field sizes, years when each field was sampled, and the numbers of samples collected from each field in each year are shown in Table 4.1. Five fields were under a certified organic management system (T4) that receives no chemical inputs at any time, whereas the other five were under a reduced chemical input system (T3) that receives banded herbicide and starter N

fertilizer only at planting time for wheat and corn, for a 50% reduction in chemical inputs compared to the conventional management. Both T3 and T4 fields receive additional post-planting cultivation and T4 is rotary-hoed to control weeds (Simmons, 2012). All fields were in corn-soybean-wheat rotation with red clover (*Trifolium pratense L.*) used as a cover crop. The rotational system is completed in three years: corn is planted in May and harvested by September of year one; soybean is planted in May and harvested by October of year two; winter wheat is planted in November of year two and harvested by July of year three; and red clover is frost-seeded in existing winter wheat in February of year three and mowed by May of the following year before planting corn again. In my experiment, red clover was frost-seeded in February of 2007, 2008, and 2010, and mowed in May of the following year. Corn was planted in May and harvested in September 2008, 2009, and 2011. Four fields were sampled in 2008, four fields in 2009, and five fields in 2011 (Table 4.1).

4.3.2. Data collection

Red clover biomass was sampled on three dates: immediately after the previous winter wheat was harvested (August); in fall (late September); and the following year just before the red clover was plowed down and corn was planted (May). The total dataset consisted of 392 sampling locations randomly selected within each field; each sampling location was georeferenced using a global positioning system (GPS) (Fig. 4.1). At each sampling location, red clover was cut at ground level from a 0.5 x 0.5 m quadrat. Dry weight biomass values were obtained after drying the samples at 60° C for 48 h (Corbin and VanderWulp, 2010).

Prior to red clover biomass removal, normalized difference vegetative index (NDVI) measurements were taken at each sampling location using a portable optical sensor device (The

GreenSeeker™ optical sensor unit, model RT200; NTech industries, Inc., Ukia, CA, USA)

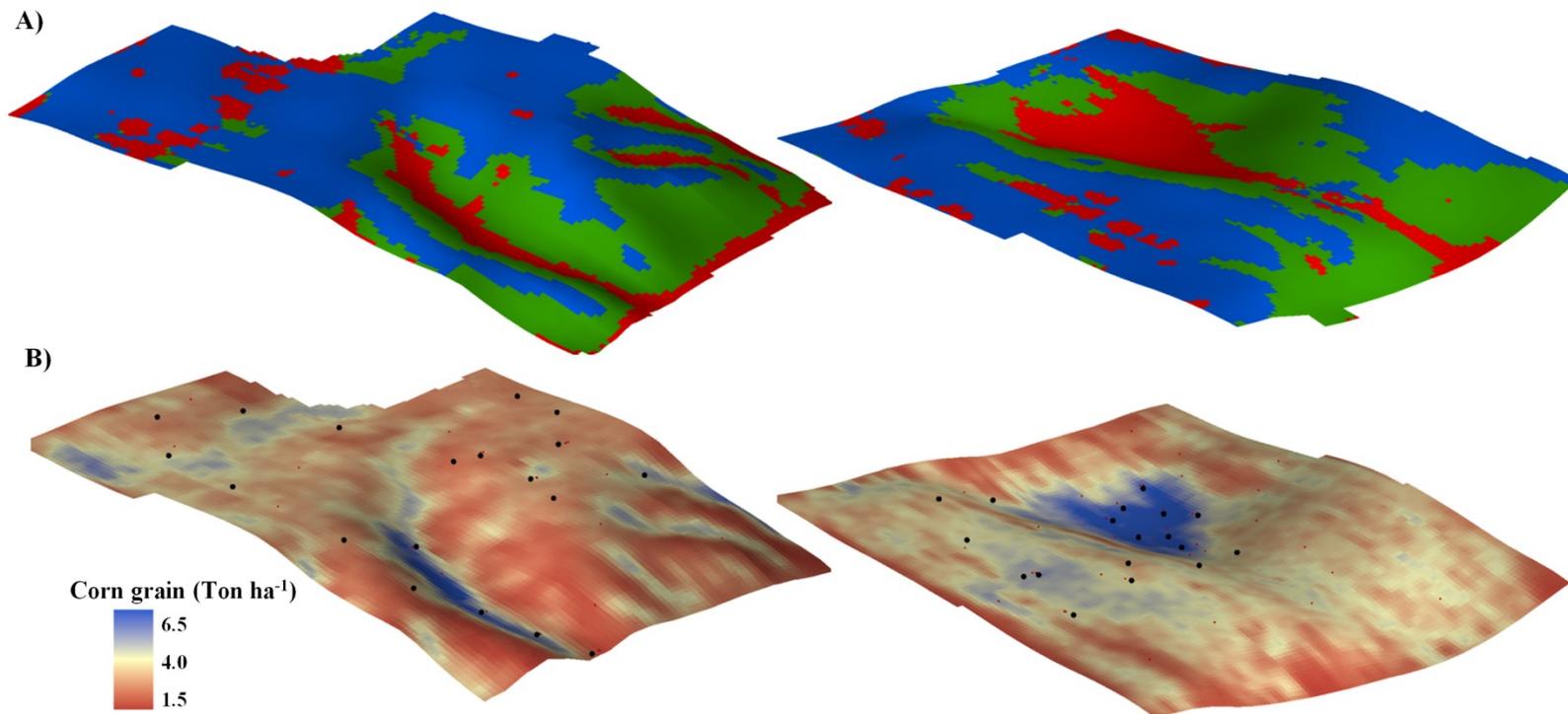
installed on a cart (Munoz et al., 2010). NDVI measurements were also taken for the entire field approximately every 3.5 m along Green Seeker™ cart passes with 10 m distance between the cart passes. This dataset was used to build continuous maps of red clover biomass for each field.

Table 4.1. List of the fields used in the study, along with years when the fields were sampled, their agricultural management treatments (T3=Low input and T4=Organic), field sizes, and the numbers of samples collected per field.

Field	Years	Treatment	Area (ha)	Samples
301	2008 and 2011	T3	5.8	20* / 6 [†]
38	2008 and 2011	T4	7.4	56* / 6 [†]
79-S	2008 and 2011	T3	5.7	12* / 6 [†]
97	2008	T4	5.4	42
93	2009	T4	5.6	56
87	2009	T3	4.9	57
52	2009	T3	5.9	61
79-N	2009	T4	5.7	64
91	2011	T4	3.3	3
822	2011	T3	1.4	3

* Number of samples collected in 2008. [†]Number of samples collected in 2011.

Figure 4.1. Example of topographical and yield information available at two of the studied fields (i.e. field-97 on the left and field-38 on the right). A) Map showing field topography classified as depression (red), slope (green), summit (blue) positions; vertical gradient corresponds to field elevation. B) Sampling points where the cover crop biomass was collected before plowing down in along with map of corn yield in 2008.



Corn biomass was collected at another set of random locations within each field at phenological stage V4. All corn plants were cut at ground level along a 1 m transect in two adjacent rows. Corn dry biomass was recorded after drying at 60° C for 72 h. A total of 150 samples were collected across all three sampling years. Corn was harvested using a combine equipped with precision agriculture software to allow yield measurements with coincident GPS latitude and longitude data (Robertson et al. 2012). Grain flow rate (lb. sec^{-1}) was measured across each field with a density of approx. 870 readings per hectare. Yield data were processed by removing errors using yield editor software (Sudduth and Drummond, 2007). Point observations were interpolated using ordinary kriging, and a raster yield map with 1-m resolution was generated.

Topographic data were derived from a Digital Elevation Model (DEM). High resolution elevation data was collected by airborne Light Detection and Ranging (LiDAR) from all fields on April 2008 (Kucera International Inc.), with vertical accuracy <15 cm and horizontal accuracy <1 m. Point elevation observations were processed for outliers and anomaly data before interpolation using ordinary kriging to generate a 1-m resolution elevation maps. Elevation maps were filtered for fine-spatial variation by using a neighborhood radius of 20 m (Munoz and Kravchenko, 2012). Subsequently, topographic attributes were derived from the filtered Digital Elevation Model (DEM), including relative elevation, slope, aspect, wetness index, flow accumulation, solar radiation, flow length, and curvature. For descriptions of the above mentioned terrain attributes see Moore et al. (1993) and Townsend and Walsh (1996).

I have used the absolute value of curvature as it was linearly associated to biomass and yield. I also prefer to use absolute curvature because it serves as an indicator of topographic “complexity” in my agricultural fields. Other metrics derived from topographical features has

been shown to be field specific (Timlin et al., 1998; Munoz et al., 2013). Values for each topographical attribute and corn yield were extracted for each sample location of red clover. Values were also extracted from red clover biomass map for each location of corn biomass at stage V4. ArcMap (ESRI, 2003) was used for all topographic data processing.

4.3.3. Landscape classification

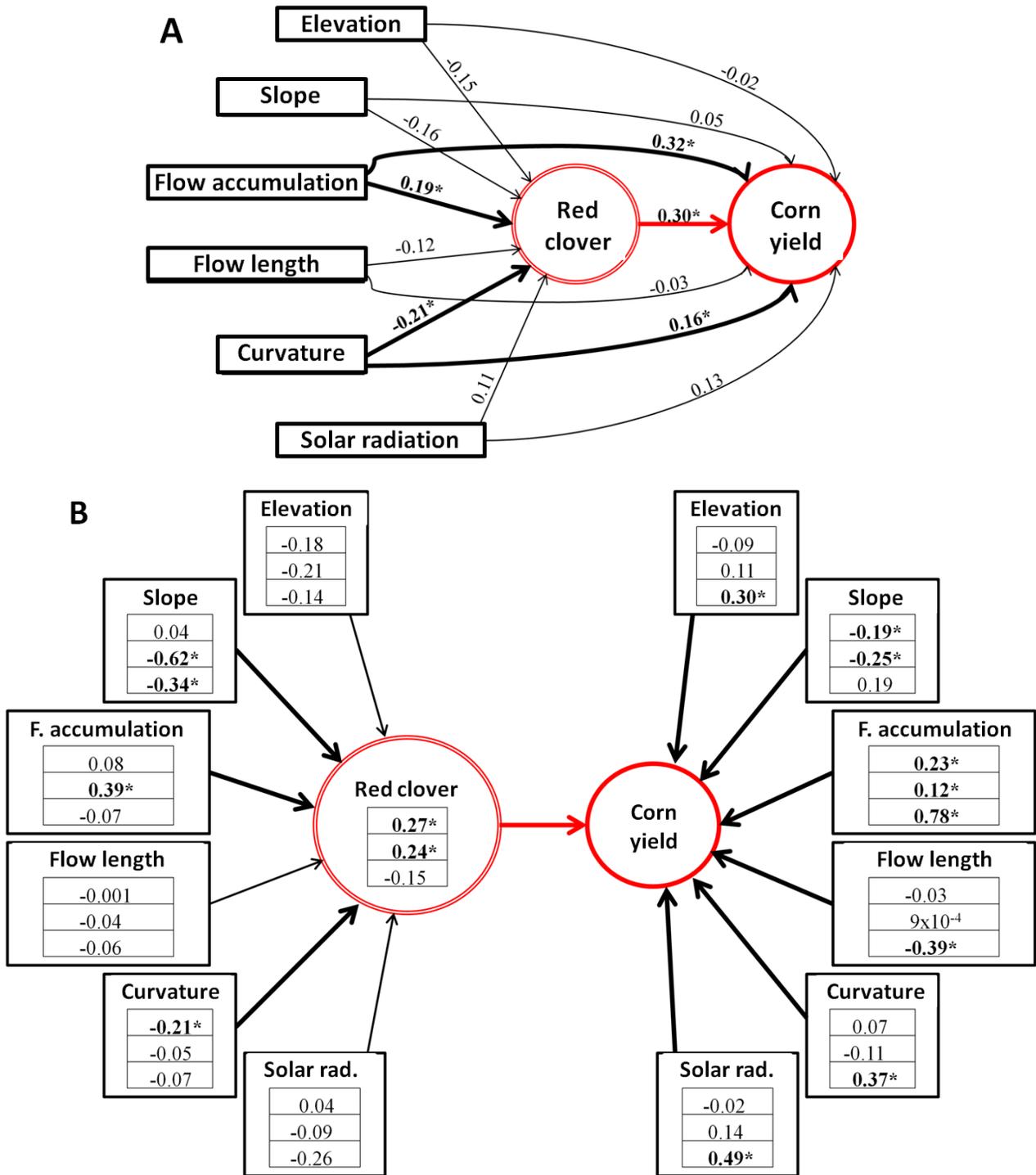
In order to classify the terrain by landscape position I used a total of 193 geo-referenced survey points recorded from 27 experimental fields at KBS in early May, 2010. The location of each survey point, randomly selected within each field, was categorized as one of three landscape positions, namely depression, slope, or summit. A total of 65 locations were recorded in depression positions, 72 in slope positions, and 56 in summit positions. For each survey point, the values of the topographic attributes were extracted and used as predictor variables in a linear discriminant analysis (Venables and Ripley, 2002), using MASS package in R (<http://www.r-project.org>). The original dataset was divided into a training subset (130 locations) and a testing subset (63 locations). Discriminant functions were computed for each landscape position based on the training subset and then used to predict the classification membership in the testing subset. Classification results, expressed as the percentage of the samples in the testing subset that was correctly classified were: 91.0% for depression, 100% for summit, and 88.0% for slope positions, indicating high classification accuracy. The total percentage of correctly classified positions using leave-one-out cross-validation was 90.4%. The discriminant functions obtained by linear discriminant analysis of the entire dataset were used to predict the classification membership in the regular sampling grid (2x2 m) for each of the ten fields. Finally, a continuous map of landscape categories was built using the classification membership (Fig. 4.1).

4.3.4. Statistical Analysis-Hierarchical path analysis

In order to examine direct and indirect effects of red clover biomass and topographic attributes on corn yield I used Path Analysis (PA), which provides estimates of the magnitude and significance of hypothesized causal relationships between sets of variables (Gajewski, et al. 2006). PA requires a formal specification of a theory-based model. Schematic representation of the model used in this study is shown on Fig. 4.2-A. It includes effects of topographical variables on red clover and on corn yield and the effect of red clover on corn. Specifically, I hypothesize that topography affects growth and performance of both red clover and corn, while red clover biomass provides an additional influence on the performance of the following corn crop.

The magnitudes and directions of the effects are defined by path coefficients. Path coefficients are standardized regression coefficients which can be interpreted as the expected change in the standard deviation of the response associated with a unit change in the predictor (Pedhazur, 1997). They can be also interpreted as the fractions of the standard deviation of the dependent variable (with the appropriate sign) for which specific factors are directly responsible (Wright, 1934). The magnitude of the path coefficient serves as an indicator of the relative importance of each predictor (Wright, 1960). PA allows errors or unexplained variance to be specified; therefore enabling estimations of the predictive power in red clover biomass and corn yield independently.

Figure 4.2. Path diagram for the relationship between topographical attributes, red clover biomass and corn yield. A) Basic model structure (M-1). B) Hierarchical model structure with varying slopes per year (M-4).



Modeling ecological processes using hierarchical statistical model approach has been presented as a way to integrate experimental and observational data (Cressie et al. 2009; Ogle, 2009), facilitating model-based inference. As described by Ogle (2009), this approach provides a flexible environment that allows us to acknowledge a richer understanding of the ecological systems, and thus permitting multiple hypotheses to be considered. I now present the models considered in this study and its supporting hypothesis.

The main model M-1 (Fig. 4.2-A), which is a representation of my main hypothesis, considers two response variables, red clover biomass and corn yield, potentially explained by a set of topographic attributes. In addition, red clover biomass is a mediating variable of the possible indirect effect of topography on corn yield. M-1 does not consider the effect of experimental factors such as years, fields, or agricultural management (Table 4.2). In practical terms, M-1 assumes that the relationships between topography, red clover and corn yield (magnitude and direction) do not depend on the experimental factors and the path coefficients in M-1 are the same across all years, fields, or agricultural management. For example, for the inference on corn yield, one general intercept (α) and seven slope coefficients (β_t , one for each topographical predictor plus red clover) are computed (Fig. 4.2-A).

Table 4.2. Summary of the statistical models used to analyze the relationships between corn, red clover cover crops and topographical variables.

Model	Structure	Experimental factors included		
		Fields	Years	Management
M-1	All samples pooled together	No	No	No
M-2	Hierarchical-data pooled by fields	varying intercepts	No	No
M-3	Hierarchical-data pooled by fields	varying intercepts and varying slopes	No	No
M-4	Hierarchical-data pooled by fields and years	varying intercepts	varying slopes	No
M-5	Hierarchical-data pooled by fields and managements	varying intercepts	No	varying slopes

The advantage of my dataset is that the data represent a diverse set of experimental factors: different agricultural fields, management systems and years. Analysis of the contribution of these factors is of great interest and importance, first, because it contributes to explaining the variance in the response variables and thus reduces model errors, and second, because it could serve to identify the potential effects associated with these factors and thus estimate differences between treatments (e.g. difference between years). Hierarchical models have been used with promising results in similar studies, that is when estimating regression coefficients was conducted along with accounting for individual- and group-level sources of variation (Gelman

and Hill, 2007; Cressie et al. 2009). However, implementation of hierarchical structures is not readily available in PA software packages, and including multiple factors in the model often induces lack of identification of the latent variable model. In other words, the number of parameters in the model can quickly become too large to be supported by the data. An unsatisfactory solution is to ignore the factors and pool all the observations without considering effects of field, management system and year (as in M-1). Implementation of PA under the Bayesian framework addresses this problem as the inference in each parameter is based individually on its posterior probability distribution, which describes the probability of the parameter given the data (Schneider et al., 2006).

Models M-2 through M-5 (Table 4.2), are hierarchical models that consider the effect of experimental factors. These models are a generalization of the model M-1, where intercepts, and possibly slopes, are allowed to vary by group (i.e. by factor's level). Including one experimental factor implies the estimation of regression coefficients for every level of this factor, thus increasing model complexity. These hierarchical models could have varying intercepts, varying slopes, or varying intercepts and slopes (Gelman and Hill, 2007). For example, the most complex model representing the effect of individual fields (ten fields used in this study) will require the estimation of ten intercepts and seventy slopes. Agricultural managements (T3 and T4) and years of data collection (2008, 2009 and 2011) can be also included in a similar fashion. I explored several candidate models that include the effect of each experimental factor and combinations of two factors (as varying intercepts and/or slopes). However, I only present here the four most relevant candidate models as they serve to highlight important points in my discussion. A model that considers all three experimental factors was not fitted because of insufficient number of observations in all possible groups. The systems of equations for all five studied models are

summarized in the Appendix.

Model M-2, is a hierarchical model with varying intercepts per field, thus data is pooled by fields (Table 4.2). M-2 allows each field to have an independent regression line for the relationships between topography, red clover and corn yield (lines with different intercepts); however M-2 does not allow a different regression slope (magnitude) in the effects for each field (all lines have the same slope). In practical terms, M-2 is a model that recognizes that every field could have a different effect on the response variables (for example due to differences in fertility), but at the same time assumes that the effects of topography on red clover and corn yield do not depend on the field.

Model M-3, is a hierarchical model with varying intercepts and varying slopes per field, thus as in M-2 data is pooled by field (Table 4.2). M-3 is similar to M-2 in that independent intercepts can be fitted for each field; however they differ in that M-3 allows different magnitude in the effects for each field (varying slopes). In practical terms, M-3 is a model that recognizes that each field could have an effect on the response variables, and also the magnitude of the effects, between topography, red clover and corn yield, could vary from field to field. In other words, the effects of topography on red clover and corn yield now can depend on the field.

Model M-4, is a hierarchical model with varying intercepts per field and varying slopes per year, thus data is pooled by field and year (Table 4.2). M-4 is similar to M-2 in that independent intercepts can be fitted for each field; however, M-4 allows different magnitude in the effects for each year. In practical terms, M-4 is a model that recognizes the possible variation from field to field, and also assumes that the magnitude of the effects depends on the year. In other words, the effect of topography on red clover and corn yield could be different from year to year. Regression lines in M-4 could have different intercepts for each field; however, fields

coming from the same year follow the same regression slope.

Model M-5, is a hierarchical model with varying intercepts per field and varying slopes per management system, thus data is pooled by field and management (Table 4.2). M-5 allows independent intercepts to be fitted for each field, and allows different magnitude in the effects for each agricultural management system. In practical terms, M-5 is a model that recognizes variation from field to field, but also assumes that effects depend on the management system. Regression lines in M-5 could have different intercepts for each field; however, fields coming from the same management system will have the same regression slope.

Given the hierarchical structure of these models I chose to use a Bayesian paradigm for model fitting and prediction (Carlin and Louis, 2000; Gelman et al., 2004). I assigned non-informative prior distributions to each model parameter. For both the population and field random effect parameters I assumed α_i and β_i to follow $\text{Normal}(0, 10^5)$ distribution. Variances are estimated based on the parameter τ^2 , which it is assumed to follow a $\text{Gamma}(0.001; 0.001)$ distribution. Inference, within a Bayesian paradigm, is based on Markov Chain Monte Carlo (MCMC) samples (post burn-in) from the parameters' posterior distribution (Gelman et al., 2004). Given posterior samples for a generic parameter θ , I can calculate any desired summary statistic, thus point predictions and associated measures of uncertainty can be calculated from the posterior predictive distribution. To compare all candidate models, I used the Deviance Information Criterion (DIC) (Spiegelhalter et al., 2002). DIC compares models based on the trade-off between goodness of fit and model complexity. Goodness of fit is expressed as the deviance, whereas complexity is measured by an estimate of the effective number of parameters. DIC is calculated from the MCMC samples. The expected posterior deviance ($D^*(\Omega)$) of the collection of parameters (Ω) is computed as $E_{\Omega|y}\{-2\log L(\text{Data}|\Omega)\}$. The effective number of

parameters, which is used as a penalty, is computed as $P_D = D^*(\Omega) - D(\Omega^*)$, where Ω^* is the posterior mean of the model parameters. Finally, DIC is given by $D^*(\Omega) + P_D$; therefore, lower DIC values indicate better models. The candidate models were fit using the software WinBUGS (Lunn, et al., 2000). Convergence diagnostic and posterior inferences were implemented in the package ‘coda’ in R (R Development Core Team, 2008).

4.3.5. Statistical Analysis-Analysis of covariance

Hierarchical path analysis allows us to identify the relative importance of red clover biomass on corn yield; however it does not help to quantify the magnitude of this effect. In order to quantify how much red clover biomass contributes to the increase in corn biomass and yield at different landscape positions I used analysis of covariance (ANCOVA), which is a general linear model that combines analysis of variance and regression analysis. Corn yield is the response variable, red cover biomass is the continuous predictor (covariate), and topographical position is the categorical factor with three levels (depression, summit or slope). The effect of field, year and management system were considered in ANCOVA, and model selection was based on DIC. The best goodness of fit was achieved with the model that included the factors landscape position and field, thus only results from this model are reported further.

4.4. Results and discussion

4.4.1. Relationship between topography, red clover biomass and corn yield

Model selection was achieved by comparing goodness of fit statistics among all candidate models (Table 4.3). Model M-1, which does not consider effects of experimental factors, had the largest MSE (0.79) and Deviance Information Criteria (DIC) (971.1). Goodness of fit increased

considerably as I included experimental factors, as evidenced by the lower total MSE and total DIC of the hierarchical models (M-2 to M-5). Among experimental factors, field was the factor that improved model fit the most. Including the effect of field as a factor reduced MSE to 0.40 and 0.28, and DIC to 818.5 and 833.2 (models M-2 and M-3 respectively) (Table 4.3); no other single factor reduced MSE and DIC as much as the field (data not shown). This is an indication that field-to-field variation explains an important portion of the total variance. Model M-3 showed the lowest MSE (0.28) among all tested models, but this model also showed the largest P_D (124.4), indicating that the model could be over-specified because it included excessively large number of parameters. For example, for the relationship between topography and red clover biomass, at least six regression intercepts and six regression slopes were estimated individually for each field. Based on the DIC, model M-2 is preferable to M-3 because it is more parsimonious, that is, model M-2 conveys equivalent information with fewer assumptions than model M-3.

Presence of substantial field-to-field differences is well documented in the literature. However, in most of the studies the authors tend to fit an individual model to each individual field instead of considering between-field variation as a factor in a global statistical model that applies to the complete dataset. For example, Green et al. (2007) studied the relationship between winter wheat yield and topographic attributes by using multiple linear regressions. They reported individual statistical models for each field with substantial differences in the number of predictors and predictive power. Kravchenko and Bullock (2000) studied the association between yield and topography across eight different fields, and found that correlations between topographic attributes and corn yield were field-specific. Similarly, Jiang and Thelen (2004) studied the effect of topography on corn yield and reported notable differences between the

predictive models for two fields, even though the fields were adjacent.

In addition to the field effect, inclusion of the year and management system effects also improved model fit (when combined with field as in M-4 and M-5). MSE was reduced once the effect of either year or management system was included. However, goodness of fit only improved when the year was included as evidenced by the DIC values (Table 4.3), indicating that the effect of year was relatively more important than the effect of the management system. MSE dropped from 0.40 (M-2) to 0.33 and to 0.39 as year (M-4) and management system (M-5) factors were included respectively (Table 4.3). The lowest DIC observed in model M-4 (798.9) compared to models M-2 (818.5) and M-5 (826.0) indicates that including independent intercepts per field and independent regression slopes per year produces the most parsimonious model. This result does not mean that M-2 and M-5 are incorrect options; it only suggests that model M-4 produces the best balance between the amount of information and the number of assumptions.

This result indicates that the relationships between topography, red clover biomass, and corn yields may change substantially from year to year. Variability from year to year is a common feature of agronomical research. The relationships between topography and crop performance has been found to vary substantially between years. For example, in a 4-year study, Kravchenko and Bullock (2000) reported substantial changes in the magnitude and significance of correlation coefficients between yield and topographic features. In a 6-year study, Jiang and Thelen (2004) reported that slope terrain was correlated to crop yield only in four years. However, no significant correlation was found for the other two years. In a 2-year study, Iqbal et al. (2005) showed that elevation was a significant predictor of cotton yield only in the first year of the study; and that the predictive power of the regression model changed substantially from 0.40 to 0.21 in the second year. Differences between years were also reported for the distribution

of legumes (Guretzky et al., 2004), where substantial changes in the significant topographical predictors were observed from year to year.

Table 4.3. Fit statistics for each studied model. MSE: mean squared error. pD: penalized deviance. DIC: deviance information criteria.

Model	Variable	MSE	pD	DIC	R ²
M-1. Basic structure	Corn yield		9.0	494.7	0.27
	Cover crop		8.1	476.4	0.33
	Total	0.79	17.1	971.1	
M-2. Hierarchical structure [varying intercepts per field]	Corn yield		17.2	377.5	0.62
	Cover crop		15.1	441.0	0.46
	Total	0.40	32.3	818.5	
M-3. Hierarchical structure [varying intercepts and slopes per field]	Corn yield		67.8	360.9	0.90
	Cover crop		56.5	472.3	0.78
	Total	0.28	124.4	833.2	
M-4. Hierarchical structure [varying intercepts per field and varying slopes per year]	Corn yield		30.7	351.5	0.75
	Cover crop		26.1	447.3	0.57
	Total	0.33	56.8	798.9	
M-5. Hierarchical structure [varying intercepts per field and varying slopes per treatment]	Corn yield		23.8	375.0	0.67
	Cover crop		21.1	451.0	0.49
	Total	0.39	44.8	826.0	

Since model M-4 showed the best goodness of fit of all tested models, I decided to use it in the further analysis. Once the effect of the year was included, relationships (paths) among topography, red clover biomass, and corn yields became more evident in path analysis (Fig. 4.2-B). Parameter estimates from M-4 along with their credible intervals are shown in Table 4. The

R^2 -values for the prediction of red clover (0.57) and corn yield (0.75) increased considerably, indicating a better predictive power of model M-4 (Table 4.3). Significant topographic predictors for red clover now included not only flow accumulation and curvature, but slope as well; interestingly, all topographic attributes became significant predictors for corn yield. This shows how pooling the dataset in a model (M-1) that does not consider the effect of experimental factors (field, year, or management system) could lead to incomplete interpretations of how the system functions. Model M-4 highlighted some important relationships between topographic attributes, red clover biomass, and corn yield. For instance, from M-4 terrain slope appeared to be an important attribute that affects not only corn yield but also red clover biomass production; in flatter areas the production of biomass was higher, whereas steep slope areas tended to produce lower biomass. The contribution of year-to-year variations was particularly noticeable in the portion of the path analysis that predicted corn yield; I noticed that the same parameter was not significant in all years. For instance, terrain slope was negatively correlated to corn yield in 2008 and 2009, indicating that flatter areas produced higher yields than sloped areas in those years. On the other hand, elevation, flow length, curvature and solar radiation were significant attributes only in 2011, indicating that areas with higher elevation, faster drainage, higher curvature and more exposure to solar radiation produced higher corn yields in that year (Table 4.4). Differences between years can be attributed largely to differences in weather. Precipitation for the year 2011 was higher compared to years 2008 and 2009 (Fig. 4.3), leading to visible water accumulation on the surface in depression areas of some fields in 2011. Overall, red clover biomass and soil water played an important role in corn production during the two years with average precipitation (2008 and 2009); however, water re-distribution was more important than the effect of red clover in the wet year (2011) (Table 4.4).

Figure 4.3. Cumulative precipitation in KBS for the years 2008, 2009 and 2011. The black line is the 24-year cumulative precipitation computed as the daily average from 1988 to 2011. Vertical gray lines are approximate day of planting, end of vegetative stage and harvesting of corn. Bars in the bottom of the figure represent values of daily precipitation.

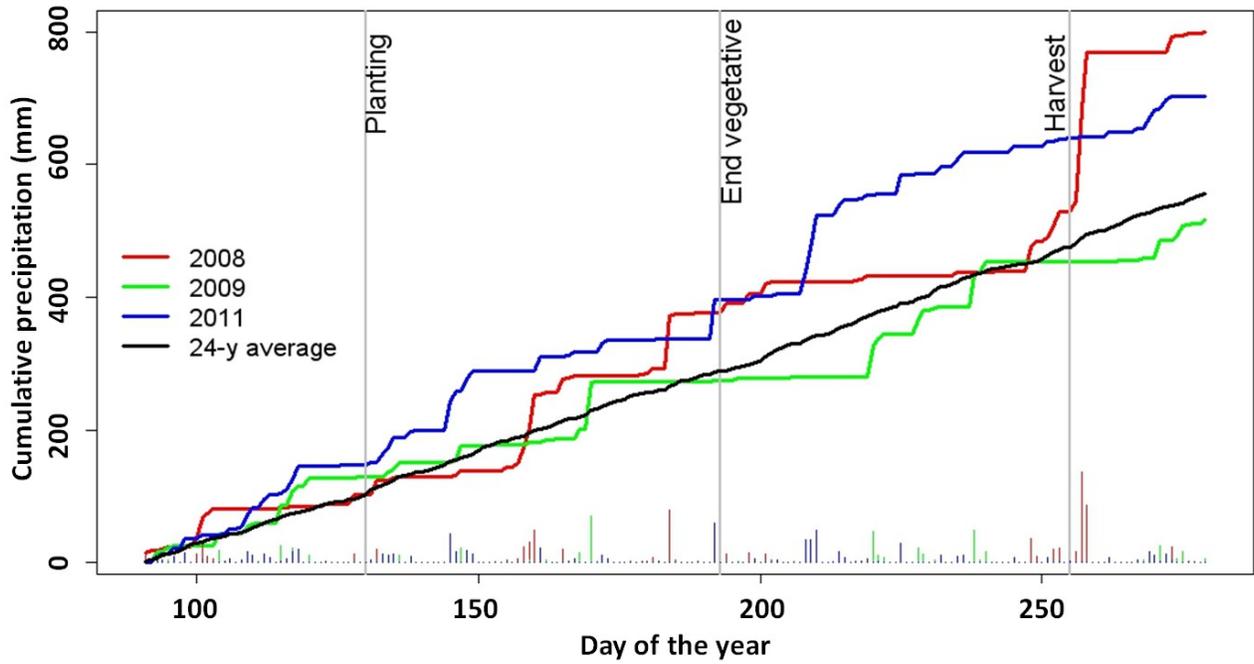


Table 4.4. Parameter estimates and credible intervals for the model with varying intercepts per field and varying slopes per year (M-4). Estimates expressed as 2.5%50%97.5% percentiles.

Node	Year 1	Year 2	Year 3
Inference on Red clover			
Relative elevation	-0.48 -0.18 _{0.13}	-0.55 -0.21 _{0.14}	-0.45 -0.14 _{0.17}
Slope	-0.26 0.04 _{0.35}	-1.00 -0.62 _{-0.25}	-0.64 -0.34 _{-0.04}
Flow accumulation	-0.17 0.08 _{0.33}	0.16 0.39 _{0.61}	-0.47 -0.07 _{0.33}
Flow length	-0.23 -0.001 _{0.23}	-0.25 -0.04 _{0.16}	-0.48 -0.06 _{0.36}
Curvature	-0.39 -0.21 _{-0.02}	-0.35 -0.05 _{0.24}	-0.32 -0.07 _{0.19}
Solar radiation	-0.24 0.04 _{0.33}	-0.41 -0.09 _{0.23}	-0.74 -0.26 _{0.23}
Inference on Corn yield (Direct)			
Relative elevation	-0.32 -0.09 _{0.16}	-0.20 0.11 _{0.41}	0.05 0.30 _{0.55}
Slope	-0.40 -0.19 _{-0.05}	-0.53 -0.25 _{-0.05}	-0.15 0.19 _{0.54}
Flow accumulation	0.04 0.23 _{0.43}	0.01 0.12 _{0.31}	0.44 0.78 _{1.12}
Flow length	-0.21 -0.03 _{0.14}	-0.16 9x10⁻⁴ _{0.16}	-0.77 -0.39 _{-0.008}
Curvature	-0.09 0.07 _{0.22}	-0.33 -0.11 _{0.13}	0.15 0.37 _{0.59}
Solar radiation	-0.24 -0.02 _{0.19}	-0.10 0.14 _{0.39}	0.03 0.49 _{0.95}
Red clover biomass	-0.02 0.27 _{0.57}	0.11 0.24 _{0.37}	-0.81 -0.15 _{0.49}
Inference on Corn yield (Total)			
Relative elevation	-0.39 -0.14 _{0.12}	-0.27 0.06 _{0.37}	0.04 0.32 _{0.61}
Slope	-0.41 -0.17 _{-0.01}	-0.70 -0.40 _{-0.09}	-0.03 0.24 _{0.52}
Flow accumulation	0.04 0.26 _{0.47}	0.03 0.21 _{0.40}	0.42 0.79 _{1.15}
Flow length	-0.23 -0.03 _{0.16}	-0.18 -0.01 _{0.15}	-0.79 -0.38 _{-0.01}
Curvature	-0.15 0.008 _{0.16}	-0.36 -0.12 _{0.12}	0.15 0.38 _{0.60}
Solar radiation	-0.25 -0.01 _{0.22}	-0.14 0.12 _{0.38}	0.06 0.53 _{0.99}

In my experiment, I noticed a major influence of topography in the wet year and a combined influence of red clover and topography in average precipitation years. Similarly, Kravchenko and Bullock (2000) concluded that weather conditions influenced the relationship topography to crop yields. They reported negative correlations between yield and curvature associated with low precipitation periods, and positive correlations in wet periods, which is in accordance with my findings. Further, Kravchenko and Bullock (2000) reported negative correlations with slope in dry environments, and Kaspar et al. (2004) reported negative relationships between corn yields and slope and curvature in dry years, whereas positive relationships were observed with elevation and slope in wet years. My results agree with those reported by Chi et al. (2009), who studied the effect of topography on wheat and found that grain yields were positively correlated to elevation and negatively correlated to flow length in a wet year. They reported an opposite trend in a dry year: crop yields were negatively correlated to elevation and positively correlated to flow length. Similar to my findings, Halvorson and Doll (1991) reported a major influence of topography in wet years. Clearly, there is strong evidence that weather conditions can play an important role in the relationship of topography to crop yield.

The relationship of weather and topography is not always consistent, as Simons et al. (1989) and Chi et al, (2009) reported limited influence of topography in wet years. Overall, I agree with Kravchenko and Bullock (2000) that the impact of topography is accentuated in extreme weather conditions; conversely, on average weather conditions the effect of topography is relatively modest as observed in the particular soil and topographical conditions of my study.

The effect of red clover on corn yield was significant in 2008 and 2009, however not in the wet year 2011. Note that the path of red clover biomass on corn yield was significant even after considering the partial paths of topography. This positive association was a common feature

observed in all tested models (data not shown). Using path analysis I found a positive effect of cover crop biomass on subsequent corn yield, an effect that was independent of topographical effects. In the studied agricultural landscapes, topography had a profound effect on all factors affecting corn yield, including water redistribution, soil properties, cover crop growth, thus until I applied path analysis it was not possible to isolate the specific contribution of cover crops. It has been difficult to unequivocally demonstrate that performance of cover crops can have a positive impact on the corn yields in fields with diverse topography and soil characteristics. My results indicate that site-specific cover crop management is worth exploring – such as high density planting of cover crops in field locations where they produce large amounts of biomass – as this could increase yield of a subsequent cash crop.

The effect of topography on corn yield differs from those of red clover. As indicated by their paths, red clover biomass and corn yield were affected differently by curvature (Fig. 4.2-A, 4.2-B). For example, the path of curvature on red clover was negative, indicating that red clover biomass decrease when curvature index increased. Conversely, the path of curvature on corn yield was positive, indicating that corn yield increased with curvature. These relationships indicate that higher red clover biomass was produced in flat locations; whereas higher corn yield was produced in areas with high curvature. Different magnitude and direction (sign) in these paths can be explained by the fact that the two crops are grown in different seasons. For instance, red clover was grown in fall-winter-spring, whereas corn was grown in summer right after mowing red clover. Lower red clover biomass in areas with higher curvature can be possibly explained by faster water movement and lower water storage. On the other hand, corn yield gains in areas with high curvature could be explained by faster water movement during wet summer months. Absolute values of curvature also indicate “topographic complexity” within a field. For

instance, in my experiment, fields with large range in curvature values were associated with complex terrain features. Conversely, flat fields tend to show lower range in values of curvature. Therefore, the variation of soil properties could be similar for those fields with similar absolute curvature. Hattar et al. (2010) studied the effect of curvature on the variation of soil properties along a toposequence. They found that the effect of curvature was more pronounced at the summit and shoulder positions. They also found that most soil properties showed similar values in the shoulder and the backslopes, which are the areas of highest and lowest values of curvature. The authors conclude that toposequence dynamics influence the distribution of soil properties. For example, the variation observed in some properties in the backslope position is related to the variation in the shoulder position at the same transect. Other differences can be noted using model M-4, indicating an advantage of this model over the simple model M-1. For example, though elevation had a positive effect on corn yield in the wet year, it had no significant effect on red clover biomass. Similarly, solar radiation and flow length had a significant effect on corn yield that is not present on red clover. Note that only two topographical features (flow accumulation and curvature) showed significant paths on red clover and corn yield in the model M-1. Though model M-1 could be useful to highlight some relationships between topography, red clover, and corn yield, the interpretation is incomplete compared to the best model (M-4).

4.4.2. Effect of red clover biomass on corn yield at different topographical positions

The performance of red clover differs significantly at different landscape positions. Depression areas tended to produce higher red clover biomass compared to summits and slopes through the entire growing season (Fig. 4.4-A). On average, the total production of biomass on summits and slopes by the time of plowing (May) was 85% and 72% of the biomass produced in

depressions, respectively. The final biomass produced on the summits was higher than on the slopes, although this difference was not statistically significant (Fig. 4.4-A). In addition, the performance of red clover also differs significantly across years. The final red clover biomass produced in 2009 was 75% higher than the final biomass in 2008 and 2011 (Fig. 4.4-B). Biomass data for August and September were only available for 2008 and 2009, and no differences were observed between these two years. To assess the magnitude of the sole effect of red clover biomass on corn biomass and yield, I use an ANCOVA model that considered the effects of landscape position and field. Figure 5 shows the effect of red clover biomass on both corn biomass (at stage V4) and corn yield, by landscape position. Red clover biomass had a significant effect on corn biomass in the stage V4 in all three landscape positions; the global R^2 -value in this ANCOVA model was 0.67. The regression slopes for all three positions, depression (0.05), summit (0.08), and slope (0.15), were significantly greater than zero, indicating that increases in red clover biomass had a positive effect in the early stages of corn growth (Fig. 4.5-A). Comparisons between landscape positions were performed at the lower, median, and upper quartile of the covariate, which corresponds to 1.0, 1.8 and 2.6 t ha^{-1} of red clover biomass, respectively. Corn biomass differ significantly among the three landscape positions at lower values of red clover biomass (1.0 t ha^{-1}), with depression areas showing the highest values followed by summit and then slope (Fig. 4.5-A). However, corn biomass did not differ between summit and slope positions at values of red clover biomass larger than 1.8 t ha^{-1} . These results indicate that maintaining high biomass of red clover can be especially beneficial on slope landscape positions where it can make a strong positive impact on early growth of a subsequent corn crop.

Figure 4.4. Red clover biomass during the growing season. Biomass values at 3 landscape positions combined across studied years (A). Biomass values in different years combined across topographical positions (B). Error bars represent standard errors of the mean. Bars with different letter within the same sampling date indicate significant differences at $p < 0.05$.

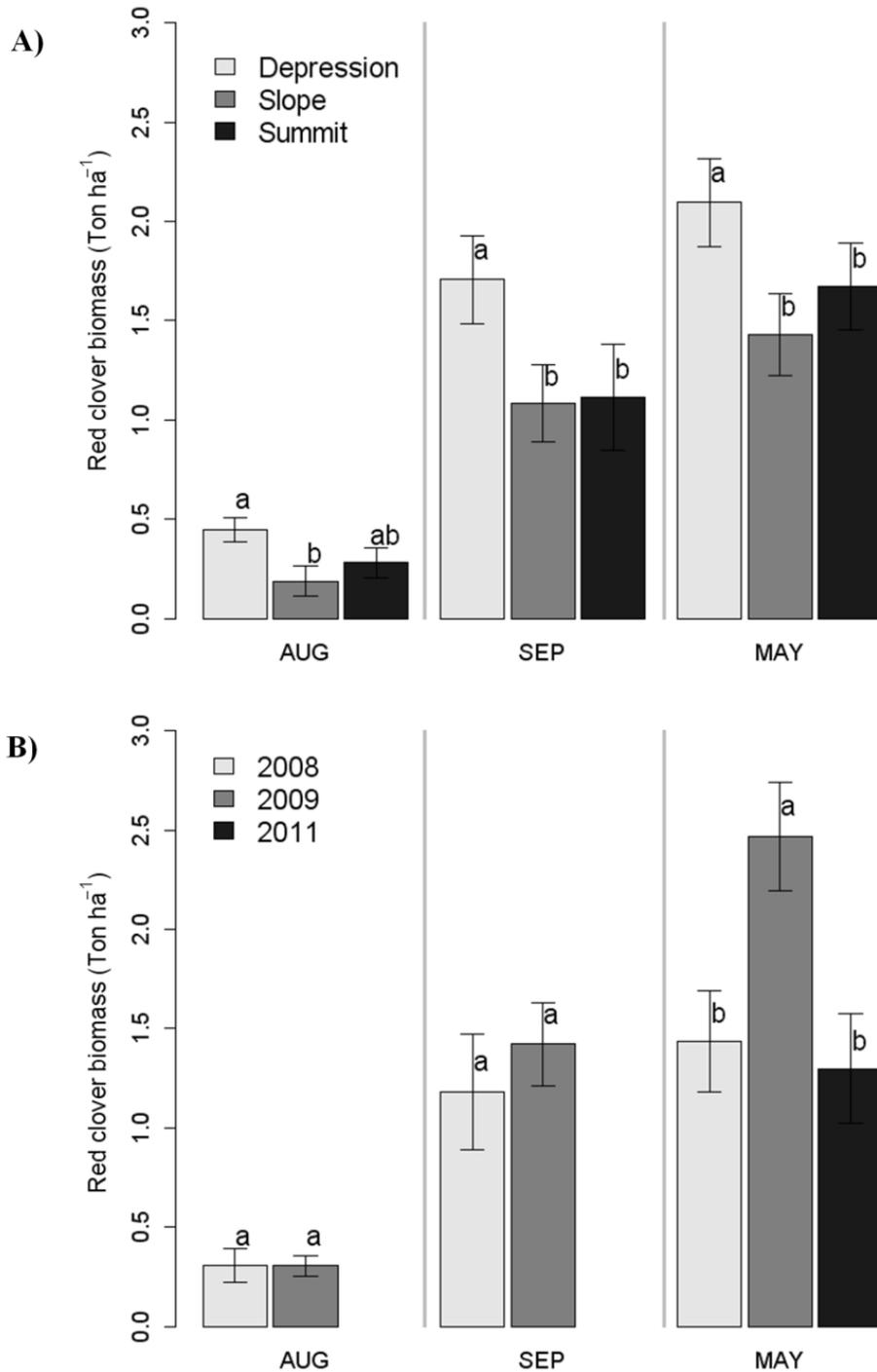


Figure 4.5. Effect of red clover biomass on corn traits at three different landscape positions. Relationship between corn biomass at stage V4 and red clover biomass (A). Relationship between corn yield and red clover biomass (B). Comparisons between landscape positions were performed at the lower quartile, median, and upper quartile of the covariate (Vertical gray lines). Different letters within the same value of red clover biomass indicate statistically significant difference between topographical positions at $p < 0.05$

* All regression slopes were > 0 , except the one for depression in the estimation of corn yield.

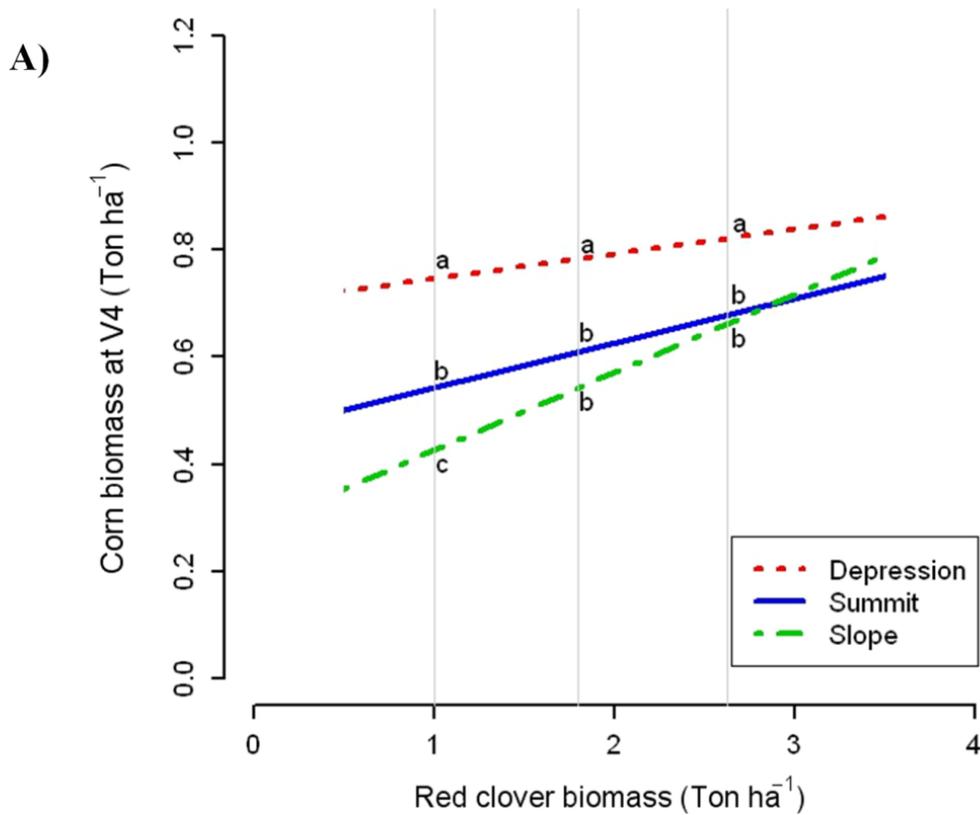
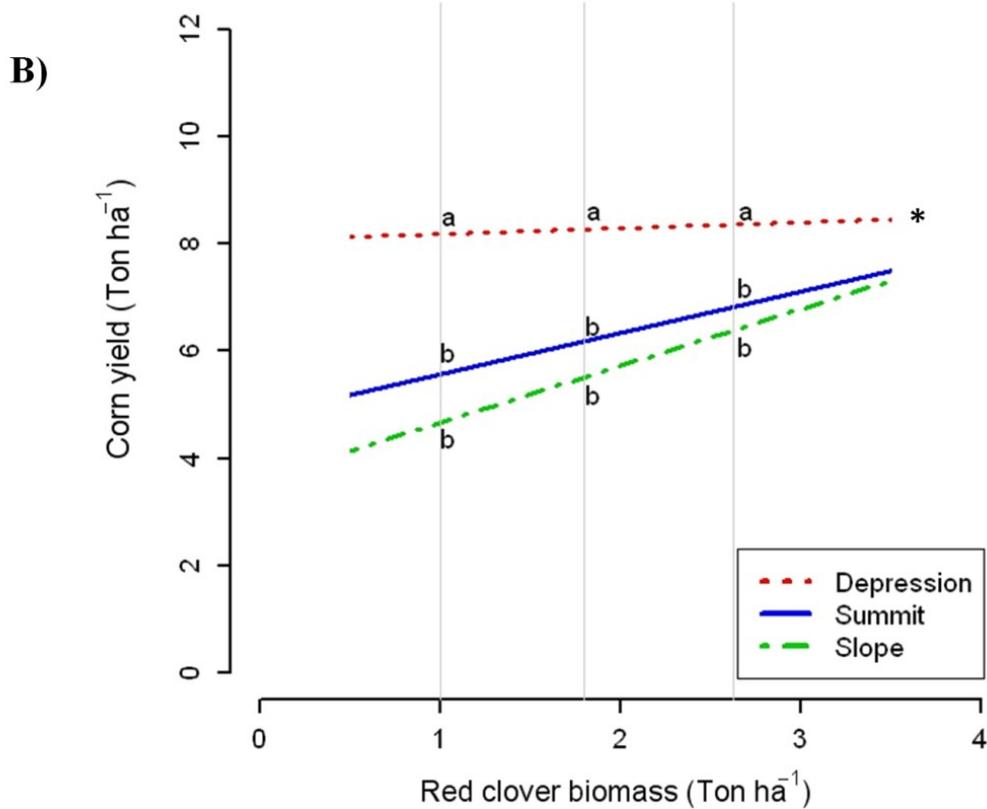


Figure 4.5. (cont'd)



Red clover biomass also had a marked effect on corn yields, particularly in slope and summit positions (global R^2 -value was 0.63). The regression slopes for summit (0.77) and slope (1.06) positions were significantly greater than zero, whereas the regression slope for depression areas (0.11) was not different from zero (Fig. 4.5-B). This result indicates that increases in red clover biomass in summit and slope areas had a positive effect on corn yield, but no effect in depression areas. For example, increasing red clover biomass by 2 t ha^{-1} (from 1 to 3 t ha^{-1}) increased average corn yield by about 2.1 t ha^{-1} in slope positions and 1.5 t ha^{-1} in summit positions, while such an increase did not have a substantial effect on yields in depression.

Depression areas produced the highest corn biomass and yields regardless of the amounts

of red clover biomass incorporated. In contrast, corn biomass production and yield in slope and summit positions depended on the amount of red clover biomass provided. These results again point to benefits of site-specific cover crop management. For example, efforts to increase red clover biomass in summit and slope positions can contribute considerably to corn yield increase, therefore helping to reduce the gap in corn yields between unfavorable (summit and slopes) and favorable (depression) areas.

4.5. Conclusions

Topographic attributes have significant explanatory power with respect to the variability of both red clover biomass and corn yields; in particular, terrain slope and flow accumulation were the most common significant predictors. However, the cover crop and the main crop were affected differently by topography. Higher red clover biomass was produced in flat locations, whereas higher corn yield was produced in areas with high curvature, which is explained by the different growing season in each crop. Attributes associated with water accumulation and water availability played an important role in corn production under dry conditions, whereas attributes related to drainage and water re-distribution played a more important role under wet conditions.

Red clover biomass had a positive effect on corn yield, but, the magnitude of that effect varied both temporally and spatially. For instance, the positive effect of red clover was significant only in the years with average precipitation, and the magnitude of this effect was clearly more pronounced in summit and slope than in depression positions. Since the independent effect of red clover biomass on corn performance was more pronounced on slopes and summits, I conclude that it is worth the effort to ensure a good cover crop stand on summit and slope areas, because this is where it can be most beneficial to the subsequent cash crop. Efforts to increase red clover

biomass in summit and slope positions supported gains in corn yield, helping to reduce the gap in corn yields between unfavorable (summit and slopes) and favorable (depression) areas.

Although the positive effect of cover crop biomass on corn yields was observed only in summit and slope positions, the positive effect of cover crops at the early stage of corn growth occurred at all landscape positions, consistent with integration of cover crops into field crop systems as a means to support improved stand establishment and early growth. Overall, I found that accounting for experimental management effects significantly improved model fit; therefore, when analyzing the relationships between topography, red clover, and corn yield, it is important to consider different sources of variability, in particular the effects of field and year.

APPENDIX

APPENDIX

The following are the systems of equations used in the five candidate statistical models:

M1: The basic structure model in which all samples are pooled together

$$z_i = f(X_i; \alpha) + e_i^{RC} \quad ; \text{ and } \quad y_i = f(X_i, z_i; \beta) + e_i^C$$

M2: A hierarchical structure with random intercepts per field, data pooled by fields

$$z_{ij} = f(X_i, F_j; \alpha) + e_{ij}^{RC} \quad ; \text{ and } \quad y_{ij} = f(X_i, z_i, F_j; \beta) + e_{ij}^C$$

M3: A hierarchical structure with random intercepts and slopes per field, data pooled by field

$$z_{ij} = f(X_i, F_j; \alpha_j) + e_{ij}^{RC} \quad ; \text{ and } \quad y_{ij} = f(X_i, z_i, F_j; \beta_j) + e_{ij}^C$$

M4: A hierarchical structure with random intercepts per field and fixed slopes per year, data pooled by field and year

$$z_{ijk} = f(X_i, F_j, Y_k; \alpha_k) + e_{ijk}^{RC} \quad ; \text{ and } \quad y_{ijk} = f(X_i, z_i, F_j, Y_k; \beta_k) + e_{ijk}^C$$

M5: A hierarchical structure with random intercepts per field and fixed slopes per management system, data pooled by field and management.

$$z_{ijl} = f(X_i, F_j, M_l; \alpha_l) + e_{ijl}^{RC} \quad ; \text{ and } \quad y_{ijl} = f(X_i, z_i, F_j, M_l; \beta_l) + e_{ijl}^C$$

where z is a response variable (red clover biomass); X is the set of predictors (topographic attributes); y is a response variable (corn yield), potentially explained by X and z (as a mediating variable). The index i correspond to each observation with $i:1 \dots n$; j is the index for each field with range $j:1 \dots 10$; k is the index for year with $k:1, 2, 3$; and finally l is the index for management system with $l:1, 2$. F is the field effect; Y is the effect of year; and M is the effect of management system. The vector α contain parameters associated with the effect of each topographic attribute in the prediction of z ; whereas β is a vector with parameters associated with topographic attributes in the prediction of y . There are two residual terms, the first one on the prediction of red clover biomass (e^{RC}), the second on the prediction of corn yield (e^C).

REFERENCES

REFERENCES

- Bhardwaj, A.K., Jasrotia, P., Hamilton, S.K., Robertson, G.P. 2011. Ecological management of intensively cropped agro-ecosystems improves soil quality with sustained productivity. *Agri. Ecos. Environ.*, 140: 419-429.
- Bennie, J., Huntley, B., Wiltshire, A., Hill, M.O., Baxter, R. 2008. Slope, aspect and climate: Spatially explicit and implicit models of topographic microclimate in chalk grassland. *Ecol. Modelling*, 216: 47–59
- Boyer, D.G., Wright, R.J., Feldhake, C.M., Bligh, D.P. 1996. Soil spatial variability relationships in a steeply sloping acid soil environment. *Soil Sci.*, 161(5): 278-287
- Carlin, B.P., Louis, T.A. 2000. Bayes and empirical Bayes methods for data analysis (2nd ed.) Chapman and Hall/CRC Press, Boca Raton. 399 pp.
- Chi, B.L., Bing, C.S., Walley, F., Yates, T. 2009. Topographic indices and yield variability in a rolling landscape of western Canada. *Pedosphere*. 19(3): 362–370.
- Corbin, D., VanderWulp, S. 2012. KBS019: Aboveground net primary productivity protocol. LTER research protocols Available: <http://lter.kbs.msu.edu/protocols/111> (Accessed: 2012, Nov 20).
- Cressie, N.A.C. 1993. *Statistics for Spatial Data*, 2nd edition. Wiley, New York
- Cressie, N.A., Calder, C.A., Clark, J.S., Hoef, J.M.V., Wikle, C.K. 2009. Accounting for uncertainty in ecological analysis: the strengths and limitations of hierarchical statistical modeling. *Ecol. Applications*. 19: 3, 553-570
- Crum, J.R., Collins, H.P. 1995. KBS soils. Kellogg Biological Station Long-Term Ecological Research. Available at <http://lter.kbs.msu.edu/research/site-description-and-maps/soil-description> (Accessed: 2013, Aug 2)
- Dabney, S.M., Delgado, J.A., Reeves, D.W. 2001. Using winter cover crops to improve soil and water quality. *Comm. Soil Sci. Plant Anal.* 32: 1221-1250.
- Dinesh, R., Suryanayarana, M.A., Nair, A.K., Ghoshal, S. 2001. Leguminous cover crop effects on nitrogen mineralization rates and kinetics in soils. *J. Agron. Crop Sci.* 187: 161-166.
- Drury, C.F., Tan, C.S., Reynolds, W.D., Welacky, T.W., Weaver, S.E., Hamill, A.S., Vyn, T.J. 2003. Impacts of one tillage and red clover on corn performance and soil physical quality. *Soil Sci. Soc. Am. J.* 67: 867-877.
- ESRI, 2004. *ArcGIS version 9. Using ArcGIS Spatial analyst*. 232 pp. Redlands, California.

- Ess, D.R., Vaughan, D.H., Perumpral, J.V. 1998. Crop residue and root effects on soil compaction. *Trans. ASAE*. 41: 1271-1275.
- Fageria, N.K., Baligar, V.C., Bailey, B.A. 2005. Role of cover crops in improving soil and row crop productivity. *Comm. Soil Sci. Plant Anal.*, 36: 19, 2733-2757
- Gajewski, B.J., Lee, R., Thompson, S., Dunton, N., Becker, A., Wells, V. 2006. Non-normal path analysis in the presence of measurement error and missing data: a Bayesian analysis of nursing homes' structure and outcomes. *Stat. Med.* 25:3632-3647.
- Gallaher, R. 1977. Soil moisture conservation and yield of crops no-till planted in rye. *Soil Sci. Soc. Amer.*, 41: 145–147.
- Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B. 2004. *Bayesian data analysis* (2nd ed.), Chapman and Hall/CRC Press, Boca Raton.
- Gelman, A., Hill, J., 2007. *Data analysis using regression and multilevel/hierarchical models*. Cambridge Univ. Press, Cambridge, 648 pp.
- Green, T.R., Erskine, R.H. 2004. Measurement, scaling, and topographic analyses of spatial crop yield and soil water content. *Hydrol. Processes*, 18: 1447-1465.
- Green, T.R., Salas, J.D., Martinez, A., Erskine, R.H. 2007. Relating crop yield to topographic attributes using Spatial Analysis Neural Networks and regression. *Geoderma* 139, 23–37.
- Gregorich, E.G., Greer, K.J., Anderson, D.W., Liang, B.C. 1998. Carbon distribution and losses: erosion and deposition effects. *Soil Till. Res.*, 47, 291–302.
- Guretzky, J.A., Moore, K.J., Brummer, E.C., Burras, C.L. 2005. Species diversity and functional composition of pastures that vary in landscape position and grazing management. *Crop Sci.*, 45: 282-289.
- Guretzky, J.A., Moore, K.J., Burras, C.L., Brummer, E.C. 2004. Distribution of legumes along gradients of slope and soil electrical conductivity in pastures. *Agron. J.*, 96: 547-555.
- Halvorson, G.A., Doll, E.C. 1991. Topographic effects on spring wheat yield and water use. *Soil Sci. Soc. Am. J.* 55: 1 680–1 685.
- Harmoney, K.R., Moore, K.J., Brummer, E.C., Burras, C.L., George, J.R. 2001. Spatial legume composition and diversity across seeded landscapes. *Agron. J.*, 93: 992-1000.
- Hattar, B.L., Taimeh, A.Y., Ziadat, F.M. 2010. Variation in soil chemical properties along toposequences in an arid region of the Levant. *Catena*, 83: 34-45.
- Hauggaard-Nielsen, H., Holdensen, L., Wulfsohn, D., Jensen, E.S. 2010. Spatial variation of N₂-fixation in field pea (*Pisum sativum* L.) at the field scale determined by the 15N natural abundance method. *Plant Soil*. 327:167-184

- Hoyle, R. 1995. The structural equation modeling approach: Basic concepts and fundamental issues. In *Structural equation modeling: Concepts, issues, and applications*, R. Hoyle (Ed). Thousand Oaks, CA: Sage Publications, Inc., pp. 1-15.
- Huang, X., Wang, L., Yang, L., Kravchenko, A.N. 2008. Management effects on relationships of crop yields with topography represented by wetness index and precipitation. *Agron. J.* 100:1463–1471.
- Iqbal, J., Read, J.J., Thomasson, A.J., Jenkins, J.N. 2005. Relationships between soil–landscape and dryland cotton lint yield. *Soil Sci. Soc. Am. J.* 69:872–882.
- Jokela, W.E., Grabber, J.H., Karlen, D.L., Balsler, T.C., Palmquist, D.E. 2009. Cover crop and liquid manure effects on soil quality indicators in a corn silage system. *Agron. J.* 101:727-737.
- Jiang, P., Thelen, K.D. 2004. Effect of soil and topographic properties on crop yield in a North-Central corn–soybean cropping system. *Agron. J.* 96:252–258
- Kang, S., Doh, S., Lee, D., Lee, D., Jin, V.L., Kimball, J.S. 2003. Topographic and climatic controls on soil respiration in six temperate mixed-hardwood forest slopes, Korea. *Global Change Biol.* 9: 1427–1437.
- Kaspar, T.C., Pulido, D.J., Fenton, T.E., Colvin, T.S., Karlen, D.L., Jaynes, D.B., Meek, D.V. 2004. Relationships of corn and soybean yield to soil and terrain properties. *Agron. J.* 96: 700–709.
- Kravchenko, A.N., Bullock, D.G. 2000. Correlation of corn and soybean grain yield with topography and soil properties. *Agron. J.*, 92: 75-83.
- Kravchenko, A.N., Bullock, D.G. 2002. Spatial variability of soybean quality data as a function of field topography: I. Spatial data analysis. *Crop Sci.* 42: 804–815.
- Kravchenko, A.N., Robertson, G.P., Thelen, K.D., Harwood, R. 2005. Management, topographical, and weather effects on spatial variability of crop grain yields. *Agron. J.*, 97: 514-523.
- Kuo, S., Sainju, U.M., Jellum, E.J. 1997. Winter cover cropping influence on nitrogen in soil. *Soil Sci. Soc. Am. J.* 61: 1392-1399.
- Liang, S. 2004. *Quantitative remote sensing of land surfaces*. Wiley-Interscience, USA. 534 pp.
- Lunn, D.J., Thomas, A., Best, N., Spiegelhalter, D. 2000. WinBUGS: A Bayesian modeling framework: concepts, structure, and extensibility. *Stat. Comp.* 10:325-337.
- Mazzoncini, M., Sapkota, T.B., Barberi, P., Antichi, D., Risaliti, R. 2011. Long-term effect of tillage, nitrogen fertilization and cover crops on soil organic carbon and total nitrogen content. *Soil Till. Res.* 114: 165-174.
- Meisinger, J., Hargrove, W., Mikkelsen, R., Williams, J., Benson, V. 1991. Effects of cover

- crops on groundwater quality. pp. 57–68. In: Hargrove, W. (Ed.) *Cover Crops for Clean Water*. SWCS. Ankeny, IA.
- Miguez, F.E., Bollero, G.A. 2005. Review of corn yield response under winter cover cropping systems using meta-analytic methods. *Crop Sci.* 45:2318-2329,
- Miller, M. Dick, R.P. 1995. Dynamics of soil C and microbial biomass in whole soil and aggregates in two cropping systems. *Applied Soil Ecology*, 2: 253–261.
- Moore, I.D., Gessler, P.E., Nielsen, G.A., Peterson, G.A. 1993. Soil attribute prediction using terrain analysis. *Soil Sci. Soc. Am. J.* 57, 443–452.
- Munoz, J.D., Finley, A.O., Gehl, R., Kravchenko S. 2010. Nonlinear hierarchical models for predicting cover crop biomass using Normalized Difference Vegetation Index. *Remote Sen. of Environ.* 114: 2833-2840
- Munoz, J.D., Kravchenko, A. 2012. Deriving the optimal scale for relating topographical attributes and cover crop plant biomass. *Geomorphology*, 179: 197-207
- Murphy, B.W., Koen, T.B., Jones, B.A., Huxedrup, L.M. 1993. Temporal variation of hydraulic properties for some soils with fragile structure. *Aust. J. Soil Res.*, 31: 179–197.
- Nykanen, A., Jauhiainen, L., Kempainen, J., Lindström, K. 2008. Field-scale spatial variation in yields and nitrogen fixation of clover-grass leys and in soil nutrients. *Agric. Food Sci.* 17: 376-393
- Ogle, K. 2009. Hierarchical Bayesian statistics: merging experimental and modeling approaches in ecology. *Ecol. Applications*. 19:3, 577-581
- Papadopoulos, A., Mooney, S.J., Bird, N.R.A. 2006. Quantification of the effects of contrasting crops in the development of soil structure: an organic conversion. *Soil Use Manag.*, 22: 172-179.
- Pedhazur, E. 1997. *Multiple regression in behavioral research* (3rd ed.). Wadsworth, Thomson Learning, USA. 1058 pp.
- Preacher, K., Hayes, A. 2004. SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Res. Met. Instr. Comp.*, 36, 717-731.
- Queen, A., Earl, H., Deen, W. 2009. Light and moisture competition effects on biomass of red clover underseeded to winter wheat. *Agron. J.* 101:1511-1521
- R Development Core Team (2008). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org>.
- Ritchie J.C., McCarty, G.W., Venteris, E.R., Kaspar, T.C. 2007. Soil and soil organic carbon redistribution on the landscape. *Geomorphology* 89: 163–171
- Robertson, G.P., Kravchenko, S., Sippel, S. 2012. KBS056: Raw Geo-referenced Agronomic

Yields (grain flow rate). LTER research protocols. Available: <http://lter.kbs.msu.edu/datatables/297> (Accessed: 2012, Nov 20).

Sarrantonio, M., Molloy, T. 2003. Response of sweet corn to red clover under two tillage methods. *J. Sust. Agric.* 23:91–109.

Schneider, M.K., Law, R., Illian, J.B. 2006. Quantification of neighbourhood-dependent plant growth by Bayesian hierarchical modelling. *J. Ecology*, 94: 310–321.

Simmons, J. 2012. KBS004: Agronomic Protocol: Main Cropping System Experiment. LTER research protocols. Available: <http://lter.kbs.msu.edu/protocols/104> (Accessed: 2012, Jun 20).

Simmons, F.W., Cassel, D.K., Daniels, R.B. 1989. Landscape and soil property effects on corn grain yield response to tillage. *Soil Sci. Soc. Am. J.* 53:534–539.

Singer, J.W., Cox, W.J. 1998. Agronomics of corn production under different crop rotations. *J. Prod. Agri.*, 11: 462–468.

Smith, M., Frye, W., Varco, J. 1987. Legume winter cover crops. *Advances Soil Sci.*, 7: 95–139.

Snapp, S.S., Swinton, S.M., Labarta, R., Mutch, D., Black, J.R., Leep, L., Nyiraneza, J., O'Neil, K. 2005. Evaluating cover crops for benefits, costs and performance within cropping system niches. *Agron. J.*, 97: 322-332.

Sobel, M. 1982. Asymptotic intervals for indirect effects in structural equations models. In Leinhardt, S. (Ed), *Sociological Methodology* 290-312. San Francisco: Jossey-Bass.

Spiegelhalter, D.J., Best, N.G., Carlin, B.R., van der Linde, A. 2002. Bayesian measures of model complexity and fit (with discussion). *J. Royal Stat. Soc., Series B*, 64, 583–639

Stone, J.R., Gilliam, J.W., Cassel, D.K., Daniels, R.B., Nelson, L.A., Kleiss, H.J. 1985. Effect of erosion and landscape position on the productivity of piedmont soils. *Soil Sci. Soc. Am. J.* 49:987–991.

Sudduth, K., Drummond, S. 2007. Yield editor: Software for removing errors from crop yield maps. *Agron. J.* 99:1471-1482.

Terra, J.A., Reeves, D.W., Shaw, J.N., Raper, R.L. 2005. Impacts of landscape attributes on carbon sequestration during the transition. *J. Soil and Water Conservation.* 60: 438-446

Ticehurst, J.L., Cresswell, H.P., McKenzie, N.J., Glover, M.R. 2007. Interpreting soil and topographic properties to conceptualise hillslope hydrology. *Geoderma*, 137: 279-292.

Timlin, D.J., Pachepsky, Y., Snyder, V.A., Bryant, R.B. 1998. Spatial and temporal variability of corn grain yield on a hillslope. *Soil Sci. Soc. Am. J.* 62:764-773

Townsend, P.A., Walsh, S.J. 1996. Estimation of soil parameters for assessing potential wetness: comparison of model responses through GIS. *Earth Surf. Proc. Landforms* 21: 307-326.

Venables, W., Ripley, B. 2002. *Modern Applied Statistics with S* (4th ed.). Springer, USA. 495 pp.

Vyn, T.J., Janovicek, K.J., Miller, M.H., Beauchamp, E.G. 1999. Soil nitrate accumulation and corn response to preceding small-grain fertilization and cover crops. *Agron. J.*, 91: 17–24.

Wright, S. 1934. The method of path coefficients. *Annals of Mathematical Statistics*, 5: 161-215.

Wright, S. 1960. Path coefficients and path regressions: alternative or complementary concepts?. *Biometrics*, 16: 189-202.

Zhang, S., Zhang, X., Huffman, T., Liu, X., Yang, J. 2011. Influence of topography and land management on soil nutrients variability in Northeast China. *Nutr Cycl Agroecosyst.* 89:427–438

Zhu, Q., Lin, H. 2011. Influences of soil, terrain, and crop growth on soil moisture variation from transect to farm scales. *Geoderma.* 163: 45-54.

Chapter 5: Modeling performance of row crop systems with and without cover crops across topographically diverse terrain in future climate

Abstract

Climate change has become a source of concern due to its potential impact on future global food production. The simulated responses in crop performance to climate change reported in hundreds of studies have produced contrasting results. The final impact of changing climate seems to depend not only on the particular region and crop species, but also on a multitude of factors associated with the cropping systems. Topography is one of the factors playing an important role in Midwest cropping systems. However, the magnitude of topographical influences can vary in different management systems. Crop simulation models have been shown to be a good tool to investigate the impact of climate change on crop production because they integrate our knowledge about biophysical processes of crop growth and development. I studied the impact of climate change on a set of management practices while considering the effect of topography. I hypothesized that under future climate scenarios topography will play an important role in the dynamics of water and nutrient re-distribution that depends on the particular management system. I validated crop simulation model SALUS using data from a 6-year experiment that includes 10 fields, 3 crops, 3 management systems, and 3 topographical positions. Projections of crop performance under 100 years of future climate scenarios showed a significant decline in corn yields, in particular for the organic-based management; whereas

soybean and wheat yields slightly increased. Corn yield decline is explained by increased temperatures and higher nitrogen stress. Increased atmospheric CO₂ explained the yield increase in the C₃ crops (soybean and wheat). However, drought stress in the last third of the century would be negatively impacting soybean yields in slope positions; and nitrogen stress would be impacting negatively wheat yields in the organic-based treatments. Mitigation strategies for corn will require development of new genetic materials matching the new short growing season and improved strategies of nitrogen management.

5.1. Introduction

Increasing atmospheric CO₂ has been a source of concern in the last decades due to its potential effects on global future climate (IPCC, 1995). These changes in future climate characterized by increased temperature and decreased soil moisture suggest a negative impact in food production (Long et al., 2006). However, the increasing CO₂ could also impact positively global food production because of the stimulatory effect of CO₂ on photosynthesis (Long et al., 2006; Schimel, 2006). The responses of crop performance to climate change reported in hundreds of studies have produced diverse and variable results (Schimel, 2006; Hatfield et al., 2011). For example, global assessment of the potential impact of climate change presented by Rosenzweig and Parry (1994), showed a decrease in global crop production, although reductions were less at the middle and higher latitudes compare to the tropics. Conversely, projections from the Intergovernmental Panel on Climate Change (IPCC) showed that cereal agriculture in mid- to high-latitude regions will increase as a result of moderate increases in temperature (IPCC, 2007). The positive or negative impact of changing climate seems to depend not only on the particular

region and crop species, but also on a multitude of factors associated with the cropping systems that interact with the climatic variables (Hatfield et al., 2011). Understanding the effect of climate change on cropping systems is still a challenge that requires improvements in the accuracy of the projections. The accuracy in these projections depends largely on the CO₂ fertilization effect under the particular growing conditions (Long et al., 2006). However, the accuracy of projections also depend on the quality of climate scenarios and crop models used because future crop yields are typically assessed using simulations from a numerical climate model as an input to a crop simulation model.

A climate scenario is a plausible representation of the future climate (Carter and La Rovere, 2001). These scenarios are developed from coherent and consistent ‘storylines’ that assume different social, economic, technological, and environmental developments (IPCC, 1995). Therefore, each scenario represents a projection in greenhouse gas emissions as a function of the assumptions in global development. Projections of future climate data is then derived as the climate system’s response to varying greenhouse gas emissions obtained through different Global Climate Models (GCMs) (Winkler, et al., 2011). Climate projections derived for its use in crop models should represent climatic trends at a reasonable time scale resolution to match with the biophysical processes of crop growth and development (i.e. daily basis). In addition, the climate projection should be uninterrupted in time if we want to assess the sequential performance of agricultural systems across the years.

Crop simulation models have been proved to be a good tool to investigate the impact of climate change on crop production (Rosenzweig and Parry, 1994; Parry et al., 2004; Hatfield et al., 2011). The System Approach to Land Use Sustainability (SALUS) has been used with success to simulate topographic effect, management systems, and rotational systems (Basso et

al., 2001; Basso et al., 2006). SALUS is a model designed to simulate continuous crop, soil, water, and nutrient conditions under different management strategies (Basso et al., 2006). The program simulates plant growth and soil conditions every day for any time period with available weather sequences. SALUS is also able to model crop rotational systems without interruption which is an important capability to study the sequential synergistic performance of cropping systems in time.

Crop models have been used extensively to produce point estimations of yields in diverse crop species and under multiple management scenarios (Tsuji et al., 1998; Ahuja et al., 2002). In most of these studies the authors typically used different measures of error to test the validity of their predictions, however no model is perfect in the prediction and an off-set between measured and observed data is always present. To my best knowledge there are no studies that using crop models, consider the different sources of variability involved in crop production. For example, the variability in the responses obtained from different fields under a particular management system. If we want to use simulated data to formulate conclusions about management systems (e.g. differences between two management treatments) we should be able to demonstrate that differences obtained from simulated experiments represent closely differences observed in real experiments. Comparing the measures of variability observed in real experiments with simulated experiments under the same management systems will help to validate crop models as a tool to detect effects. This will be of great importance when investigating the potential of mitigation strategies using simulation models. In order to claim significant differences between mitigation strategies or practices we must assure that the simulated data represent appropriately the mean pattern and the variability associated with a mitigation strategy.

Topography is one of the factors playing an important role in Midwest rotational cropping systems (Munoz et al., 2013). In addition, the role of topography could interact with other factors innate of each management system. Unfortunately, field sites used for agricultural experiments have been traditionally placed on flat land with homogeneous soil properties. Studies used to calibrate crop models were not the exception. These experimental settings limit our ability to address the role of topography and edaphic factors on variability in crop performance as well as the role of interactions between the studied agricultural management practices and other factors affecting crop growth. This is particularly relevant to commercial agriculture that uses large fields with diverse topographic and soil conditions. The concern about the role of management and topographic interactions is especially valid for managements involving cover crop use. For example, cereal crop yields were shown to respond to the influence of cover crop biomass across diverse terrain (Munoz et al., 2013), indicating that the final impact of management systems under climate change may vary across landscapes.

I want to study the impact of climate change on a set of management practices and still considering the effect of topography. I hypothesize that topography will play an important role in the dynamics of water and nutrient re-distribution within a toposequence, that depends on the particular management system.

5.2. Objectives

The objectives in this study are: 1) validate a simulation model that considers the effect of management system and topographical position in Michigan row-crop agriculture using crop yield data; 2) evaluate the potential use of simulated data to detect effects in agricultural

practices under future climate scenarios; and 3) simulate the effect of future climate (i.e. next 100 years) on three management systems across three contrasting landscape positions.

5.3. Materials and Methods

5.3.1. Study site and management systems

The study was carried out at Kellogg Biological Station (KBS) located in southwest Michigan at 42° 24' N, 85° 24' W. The historical annual rainfall averages 890 mm/y and the mean annual temperature is 9.7 °C. The dominant soil series are the Kalamazoo (fine-loamy, mixed, mesic Typic Hapludalfs) and Oshtemo (coarse-loamy, mixed, mesic Typic Hapludalfs) (Crum and Collins, 1995). I used three annual crop treatments (T1, T3 and T4) from the Long Term Ecological Research (LTER) experiment at KBS. The treatment T1 is a conventional system, the treatment T3 is an organic system with reduced chemical input, and the treatment T4 is a certified organic system. The data were collected from a total of ten experimental fields (Fig. 5.1). Four fields were under the conventional management (T1) that receives standard levels of chemical inputs and is chisel plowed; three fields were under the certified organic management system (T4) that receives no chemical inputs at any time; and three were under the reduced chemical input system (T3) that receives banded herbicide and starter N fertilizer only at planting time for wheat and corn, for a 50% reduction in chemical inputs compared to the conventional management. Both T3 and T4 fields receive additional post-planting cultivation and T4 is rotary-hoed to control weeds. Specific dates of field activities, i.e. tillage practices and chemical input application, are available in the agronomic protocol of the LTER experiment (Simmons, 2012). All fields were in a three-year rotation that includes corn (*Zea mays* L.), soybean (*Glycine max*), and wheat (*Triticum aestivum* L.). Red clover (*Trifolium pratense* L.) was used as a cover crop in

the treatments T3 and T4. The rotational system is completed in three years: corn is planted in May and harvested by September of year one; soybean is planted in May and harvested by October of year two; winter wheat is sowed in November of year two and harvested by July of year three. For the treatments T3 and T4, red clover is seeded in existing wheat in February of year three when the ground was still frozen and finally ploughed by May of the following year before planting corn again.

5.3.2. Landscape classification and experimental plots

Field continuous landscape was classified into three topographical positions: depression, summit and slope. The classification was performed first in a dataset composed of 193 georeferenced survey points recorded from 27 experimental fields at KBS (Munoz et al., 2013). Topographic attributes derived from a Digital Elevation Model (Munoz and Kravchenko, 2012), were used as predictors in a linear discriminant analysis (Venables and Ripley, 2002). The total percentage of correctly classified positions in the testing dataset using leave-one-out cross-validation was 90.4%. The discriminant functions obtained by linear discriminant analysis were used to predict the classification membership in the regular sampling grid (2x2 m) for each of the ten fields; and a continuous map of landscape positions was derived (Fig. 5.1).

Experimental plots of 10 x 10 m were located at each topographical position within a toposequence. A total of 23 toposequences were identified in the 10 fields, therefore a total of 69 experimental plots (Fig. 5.1). Deep soil cores (7.6 cm dia.) to 1 m depth were collected from the center of each experimental plot. Soil cores were taken with the hydraulic probe, Geoprobe model 540MT (Geoprobe Systems; Salina, Kansas) in April 2010 (LTER, 2013). From the original 69 experimental plots, a reduced number (n=53) were used to monitor crop growth and

performance. Fifteen plots were located in depression position, 19 in summit position, and 19 in slope position (Table 5.1).

Figure 5.1. Location of the experimental fields at KBS (Top). Example of the topographical classification for 9 fields (Bottom). Black dots indicate the location of experimental plots for soil and yield data collection.

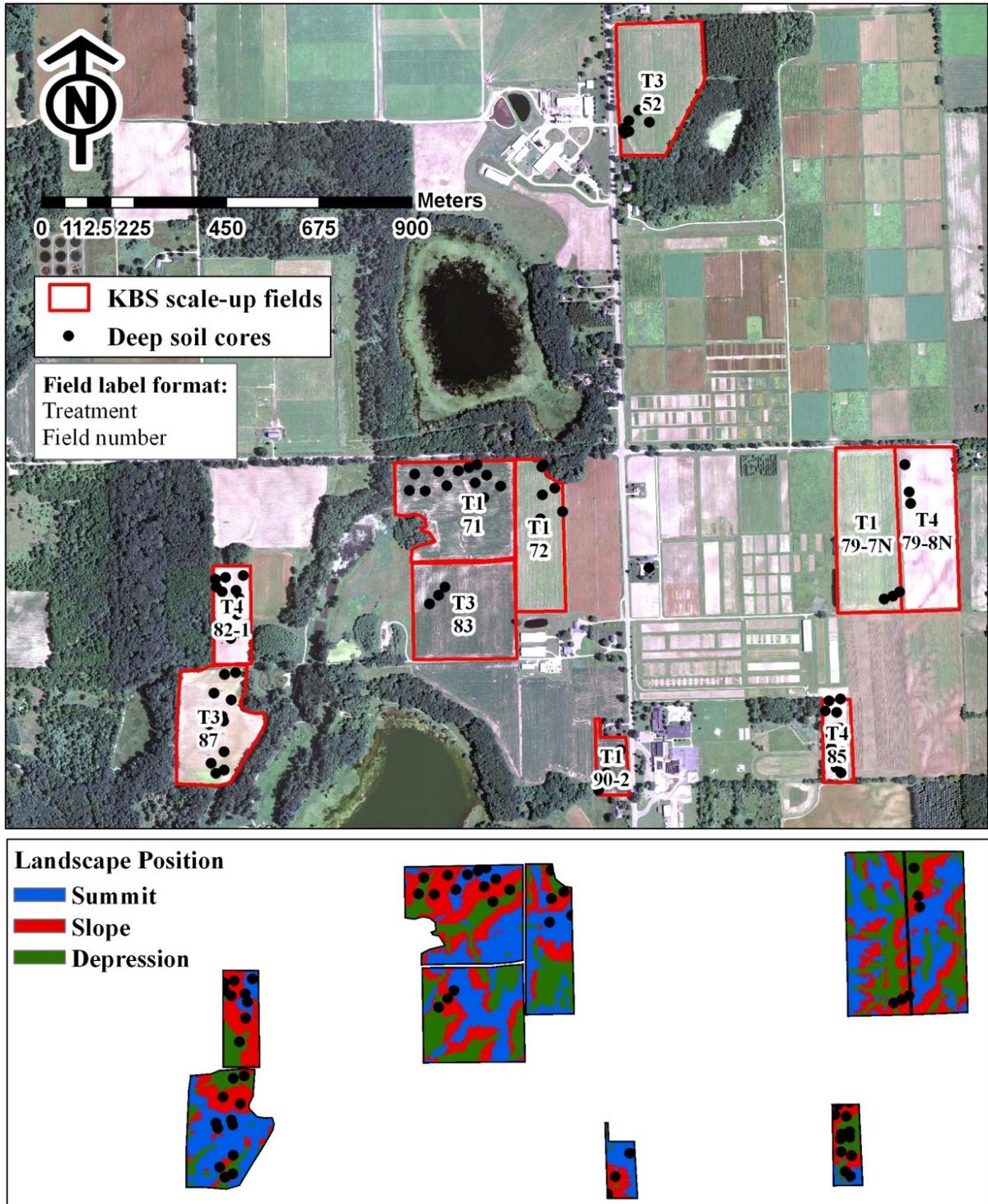


Table 5.1. Number of fields and experimental plots used in the validation of the crop model.

Numbers of plots are presented per agricultural management system and topographical position.

Agricultural management system	Number of fields	Experimental plots per position		
		Depression	Slope	Summit
Conventional (T1)	4	7	8	8
Reduced input (T3)	3	4	5	5
Organic (T4)	3	4	6	6

5.3.3. Field data for model validation

Field data was collected from 2007 to 2012 to validate the crop simulation model. Soil data, crop data, weather data, and field operation activities were the main input for the crop model. Soil properties were measured from the deep soil cores, which include soil texture, bulk density, carbon, and water limits. Field operations (i.e. planting, tillage, fertilization, and harvesting) were monitored and dates were recorded for each field in each year. A detailed list of field operations from 2007 to 2012 is available in the research protocols of KBS (Robertson and Bronson, 2011).

5.3.3.1. Soil data

Soil core analyses include bulk density, total carbon, total nitrogen, and particle-size analysis. Soil analyses were performed by specific depth intervals (0-20, 20-35, 35-50, 50-70, and 70-100 cm). Core sections were separated and dried at 60 °C for 48 h. Bulk density was

computed based on the total dry weight and the volume of each core section. Subsequently, soil samples were ground and sieved using 2-mm mesh sieve, and then cleaned from visible plant residues. Soil carbon and total nitrogen was analyzed by dry combustion method using Costech ECS 4010 Carbon-Nitrogen analyzer. Particle-size analysis was performed by the pipette-extraction method (Gee and Bauder, 1986). Figure 5.2 shows the distribution of textural classification for each topographical position in the first 20 cm. Soil water limits were estimated based on textural, bulk density, and soil carbon data using the equations presented by Ritchie et al., (1999). Volumetric upper limit was estimated from sand and clay content, and bulk density. Volumetric plant extractable water was estimated from sand content. The lower limit was estimated as the difference between the upper limit and the plant extractable water. Estimated values were adjusted for coarse fragments and soil carbon (Ritchie et al., 1999). Figure 5.3 shows mean values of soil properties in depth for each topographical position. Notice the larger range of variation in soil properties for depression positions.

5.3.3.2. Crop data

Crop population density and phenological stage were assessed in each topographical position during two years (2010 and 2011). Mean values of population density at each agricultural system and topographical position were used as input in the crop model. Corn biomass was collected at each experimental plot at phenological stage V4 during July to August 2011. All corn plants were cut at ground level along a 1 m transect in two adjacent rows. Corn dry biomass was recorded after drying at 60 °C for 72 h. Phenological stage and corn biomass values were used to check the simulated values of plant growth and performance during the growing season. Crop grain was harvested using a combine equipped with precision agriculture

software to allow yield measurements with coincident GPS latitude and longitude data (Robertson et al. 2012). Grain flow rate (lb. sec^{-1}) was measured across each field with a density of approx. 870 readings per hectare. Yield data were processed by removing errors using yield editor software (Sudduth and Drummond, 2007). Point observations were interpolated using ordinary kriging, and a raster yield map with 1 m resolution was generated. Interpolated yield data was extracted in each experimental plot (10 x 10 m) for each year. Crop yields at each experimental plot for each year were used to validate the crop model.

5.3.3.3. Weather data

Historical weather data of solar radiation, rainfall, and air temperature were available from the LTER Meteorological Station located at $42^{\circ} 24' \text{ N}$, $85^{\circ} 22' \text{ W}$ (Bergsma, 2013).

Precipitation is measured with a weighing bucket rain gauge. Air temperature and solar radiation are measured with electronic sensors. Campbell Scientific sensors are connected to a CR1000 datalogger that records readings every hour. Total precipitation as liquid-equivalent depth is reported for each day. Maximum air temperature is the highest of all hourly maxima. Minimum air temperature is the lowest of all hourly minima. Average daily solar radiation is computed including night hours.

Figure 5.2. Distribution of soil texture in the first soil depth (0-20 cm) across all experimental plots. Texture triangle is resented by topographical position: Summit, Slope, and Depression.

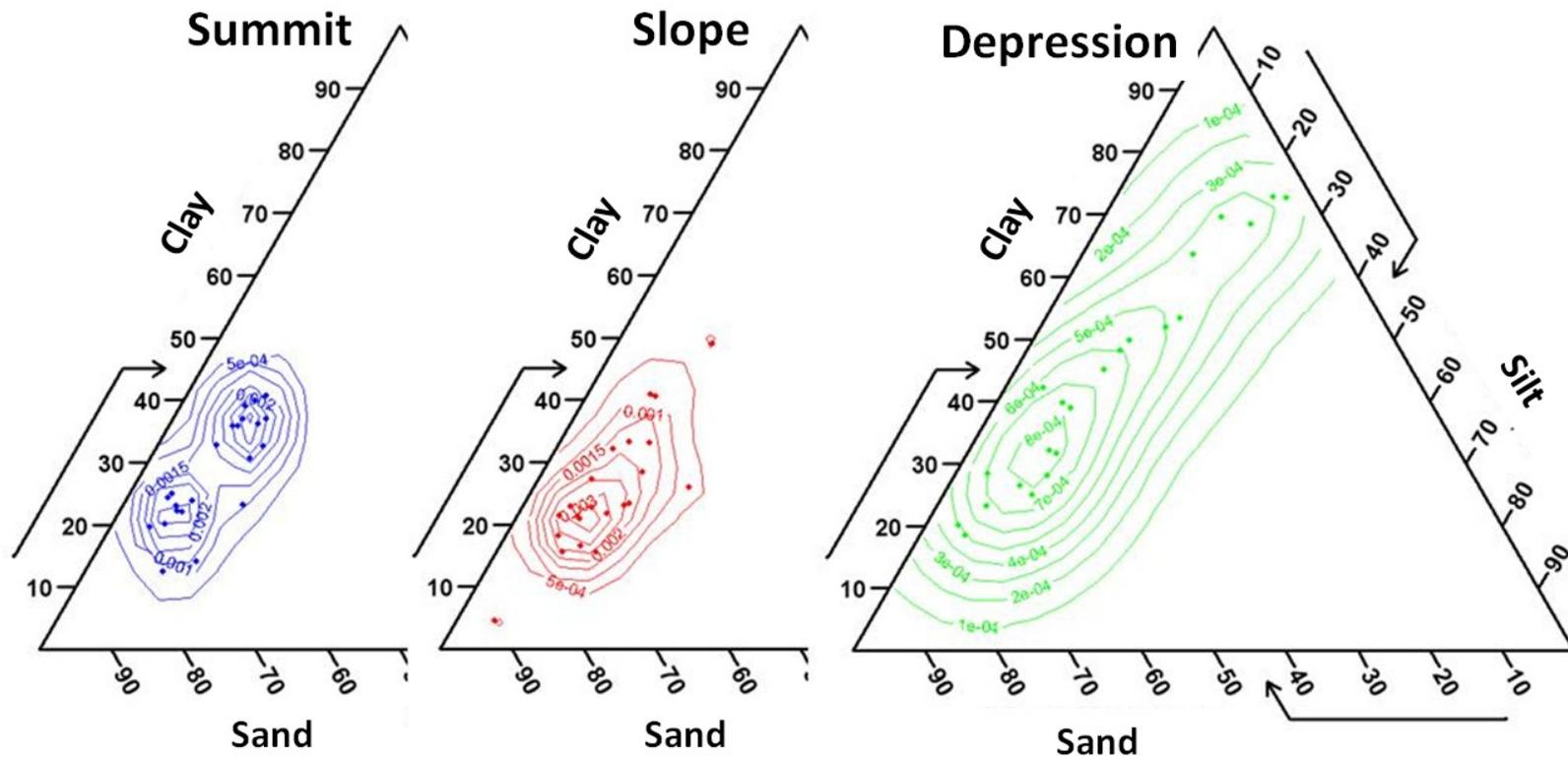


Figure 5.3. Soil properties used as input in the crop simulation model presented by soil depth: a) sand, silt, clay content, and organic carbon, b) bulk density, and water limits. Median values from all experimental plots are presented by topographical position. Medial value (line) is bounded by the 25th and 75th percentiles (shaded area). DUL is drained upper limit.

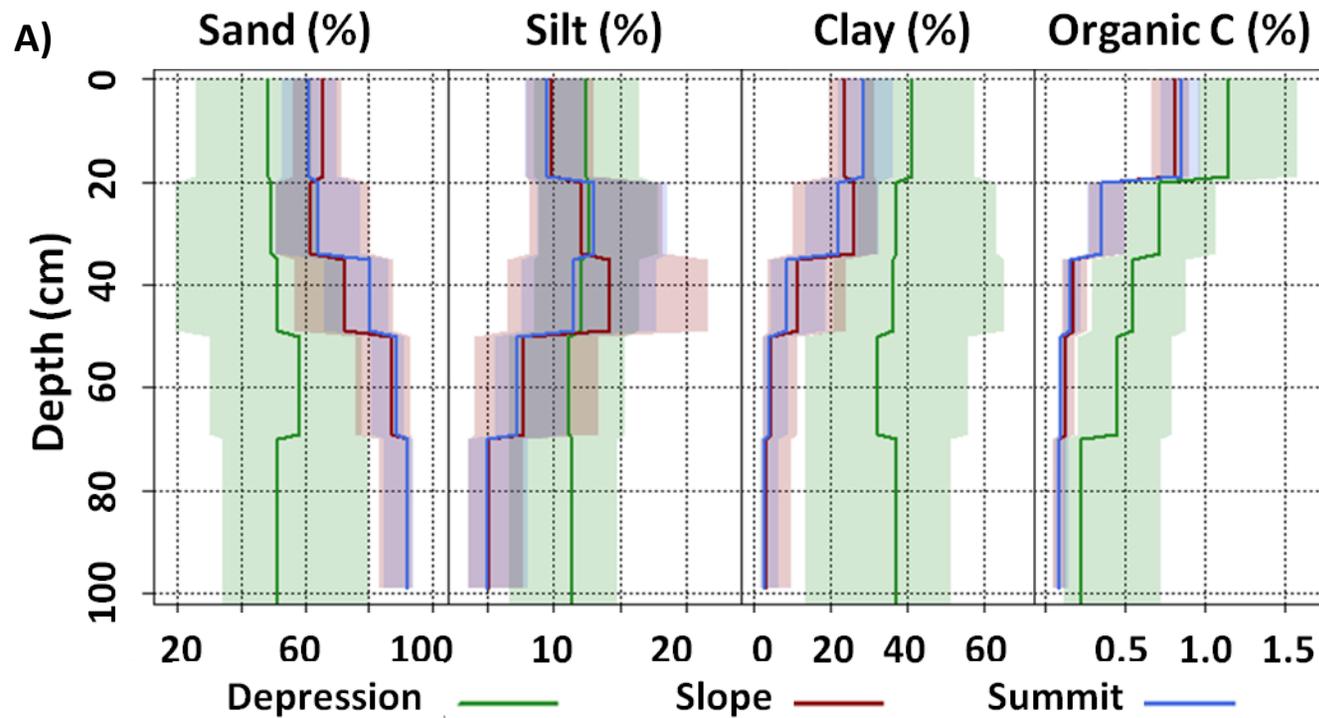
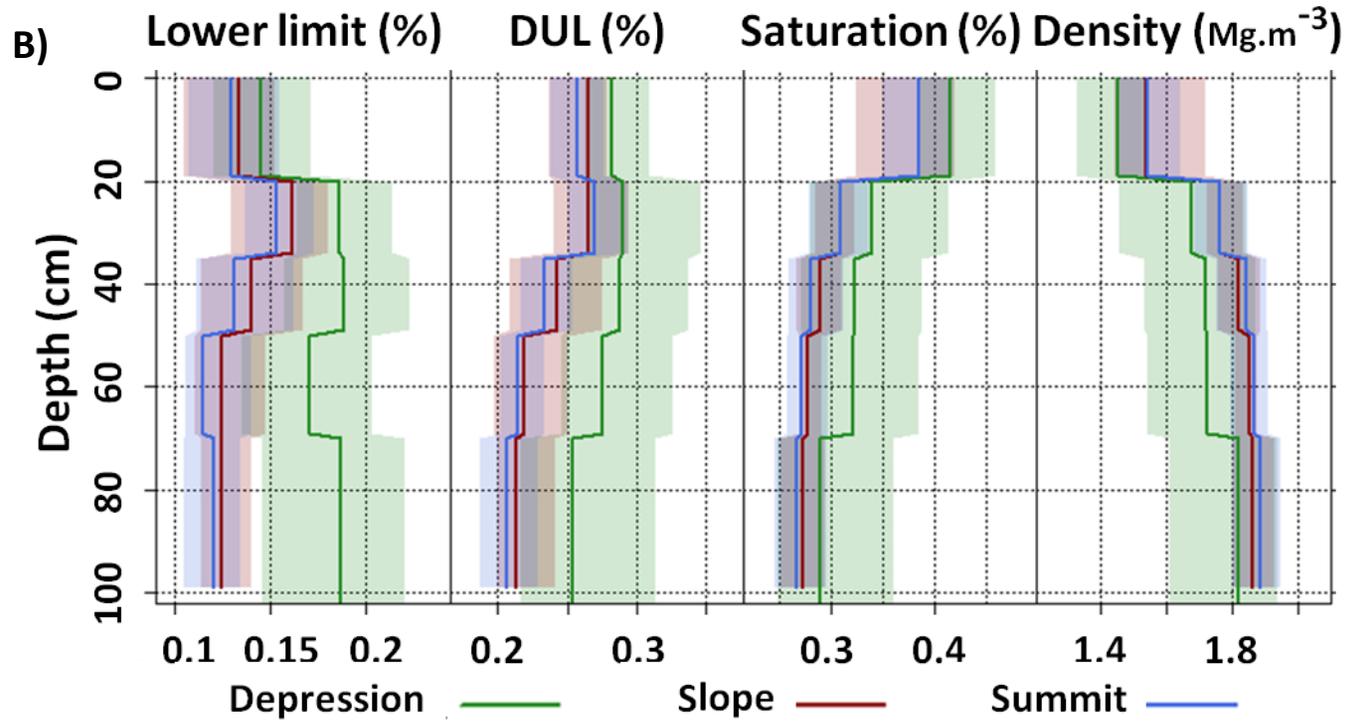


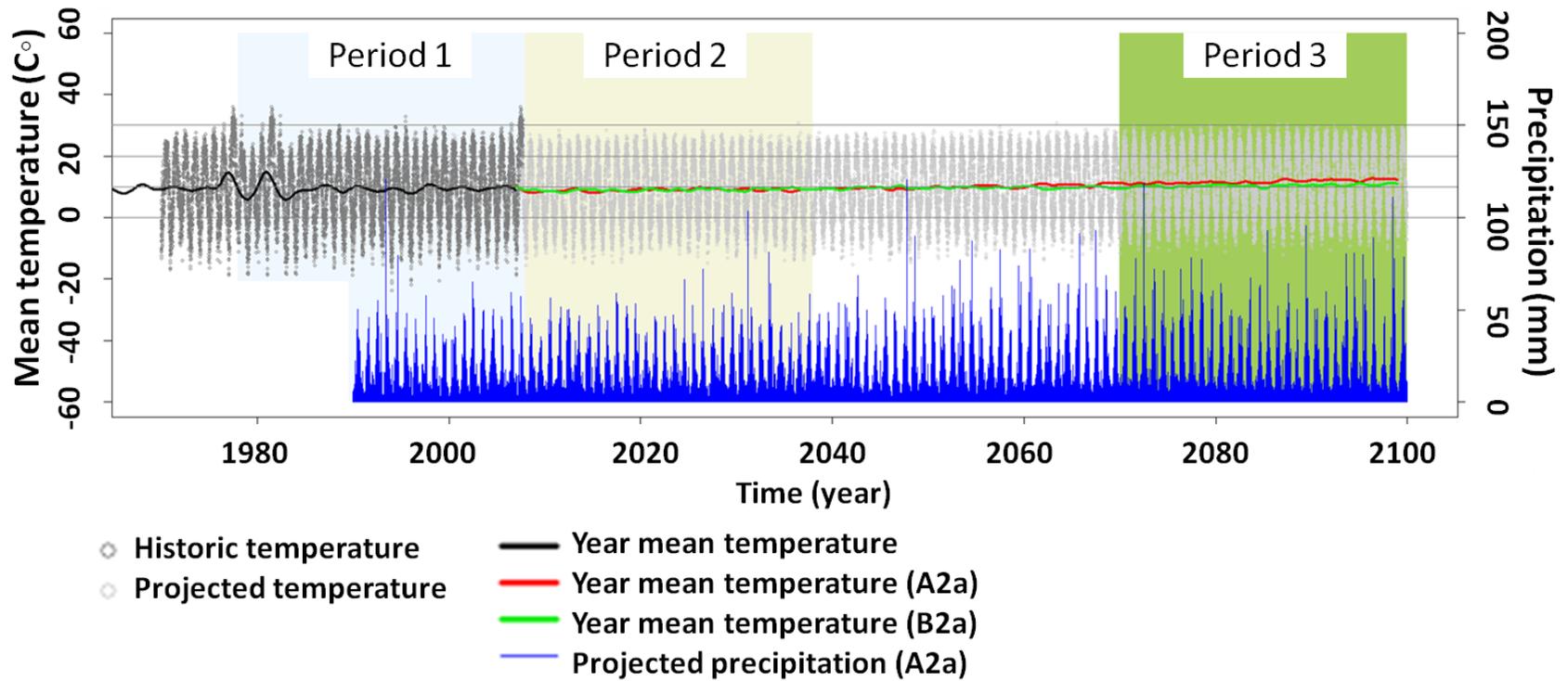
Figure 5.3. (cont'd)



5.3.4. Future climate scenarios

Future climate scenarios developed for a weather station located in Greenville, Michigan ($43^{\circ} 10' \text{ N}$, $85^{\circ} 14' \text{ W}$) were used as input in the crop simulation model to predict future crop performance. Each scenario is composed of daily values of precipitation, solar radiation, minimum, and maximum temperature during the years 2000 to 2100. Since I want to capture the major sources of uncertainty associated with future data generation I include an ensemble of climate change scenarios developed from various projections of future green house gas emissions. The ensemble of scenarios was developed using simulations from four different GCMs (CGCM2, HadCM3, ECHAM4, and CSM1.2) and two emission scenarios (A2 and B2) (Winkler et al., 2011). The two emission scenarios used here represent two different ‘storylines’: A1 represent the strong economic values and increasing globalization, B2 represent the strong environmental values and increasing regionalization. Figure 5.4 shows the projection of year mean temperature and precipitation events for A2 and B2 future climate scenarios. The ensemble of climate scenarios used in this study was developed by the Pileus project (Zavalloni et al., 2005; see www.pileus.msu.edu/climate/). Detailed descriptions of the methods used in developing climate scenarios are available in Palutikof et al. (1997) and Winkler et al. (2011).

Figure 5.4. Projection of year mean temperature and precipitation in the future climate scenarios A2 and B2 for the weather station in Greenville, Michigan.



5.3.5. Crop model and simulation experiments

SALUS software was used to simulate crop performance under current and future climate scenarios. Parameter-input files organized in four main components (soil, management, weather, and crop) contain information to initialize the model (e.g., physical soil properties) and supply driving variables (i.e., time series of weather). Table 5.2 shows the most relevant input parameters used during the period of model validation (years 2007 to 2012). A complete list with all the variables required by SALUS is presented by Ritchie et al. (1989), and in the SALUS web site (Salus software, 2013). Measured soil, weather and management parameters (except population density) were used as input in the initial simulation model (e.g. ‘blind’ simulation). I used crop parameters, and population density to ‘tune-up’ the initial simulation and therefore increase accuracy of prediction. I modified crop parameters (Table 5.2) to mimic plant development in the organic-certified crops (e.g. crops used in the treatment T4). In addition, the average value of population density was used to recreate crop populations at each topographical position. An independent simulation model was executed for each experimental plot based on its own soil and field parameters. Simulated yield values were extracted from each model to be compared with corresponding values of measured yields (see section 5.3.6).

The simulation model that produced the best prediction in the validation period was used to simulate crop performance under the future climate scenarios (years 2007 to 2099). Since the rotation is completed in 3 years, I ran three independent simulations each time starting with a different crop (corn, soybean, wheat) to avoid any effect of year by crop. The dates for field operations of planting, tillage, fertilization, and harvesting in the future were automatically selected by SALUS based on a set of decision rules. Planting dates were selected based on soil temperature and water conditions. Tillage and fertilization operations were initialized on specific

days-after-planting to match with phenological stages. Harvesting dates were selected based on crop maturity. SALUS environmental module was modified to introduce increments in CO₂ and therefore mimic the concentration of atmospheric CO₂ with the emission scenarios. Crop yield, thermal time, drought stress factor, and nitrogen stress factor were exported from each simulation.

5.3.6. Statistical analysis

The best simulation model for the validation period (2007-2012) was selected based on the differences between simulated yields from each experimental plot and the corresponding of measured yields. Five measures of error were used: the average sum of square error (*ASSE*) and the average absolute error (*AAE*) are measures of model error (off-set); the average error bias (*AEB*) indicates the direction of model error (bias); the average absolute error standardized (*AAES*) standardized the scale of error for each crop, therefore it can be used to compare model accuracy between crops; and the correlation coefficient (*r*) is a metric of association.

$$ASSE = \frac{1}{n} \sum_{i=1}^n (mY_i - sY_i)^2$$

$$AAE = \frac{1}{n} \sum_{i=1}^n |mY_i - sY_i|$$

$$AEB = \frac{1}{n} \sum_{i=1}^n (mY_i - sY_i)$$

$$AAES = \frac{1}{n} \sum_{i=1}^n |SmY_i - SsY_i|, \quad S.Y_i = \frac{(.Y_i - .\hat{Y})}{\sigma_{.Y}}$$

$$r = \frac{1}{n-1} \sum_{i=1}^n \left(\frac{mY_i - m\hat{Y}}{\sigma_{mY}} \right) \left(\frac{sY_i - s\hat{Y}}{\sigma_{sY}} \right)$$

where mY_i and sY_i are the measured and simulated yields respectively. Yield values were standardized within crop to bring each crop to the same scale of variation. Therefore $S.Y_i$ is a standardized value of yield, SmY_i and SsY_i are the standardized values of measured and simulated yields respectively. $m\hat{Y}$ and $s\hat{Y}$ are the measured and simulated mean yield respectively; and σ is the standard deviation.

Analysis of variance (ANOVA) was used to test significance in the effect of experimental factors. A linear mixed model was fitted independently for each crop, where agricultural management treatment and topographical position were fixed factors with three levels each one. Field, year, and toposequence were included as random factors. The linear mixed model was fitted independently for measured yield and simulated yield data as response variables. Mean comparisons were reported at $\alpha=0.05$.

ANOVA was also fitted on the simulated yields obtained with the future climate scenarios. To analyze the effect of climate I divided years of simulation into three climate periods: period 1 (2007-2037), period 2 (2038-2069), and period 3 (2070-2099). Climate period, topographical position, and management treatment were included as fixed factors in the linear

mixed model. A local polynomial regression was fitted to the predicted values of yield for each management treatment and topographical position (Cleveland et al., 1992). The function was fitted using 'loess' function in R (R Development Core Team, 2009). To explain the changes in simulated future yields for each crop were compared them with corresponding sources of plant stress. Therefore, ANOVA was performed on average stress values (nitrogen, drought, heat) where climate period, topographical position, and management treatments were fixed factors.

Table 5.2. Relevant model parameters used as input in SALUS during model validation. Units are presented in parenthesis

Model component	Parameter	Model component	Parameter
Soil (at each soil depth)	Bulk density (Mg m^{-3})	Management	Planting date (doy [#])
	Clay content (%)		Population density (plant m^{-2})
	Silt content (%)		Row distance (cm)
	Drained upper limit ($\text{m}^3 \text{ m}^{-3}$)		Fertilization date (doy)
	Lower limit ($\text{m}^3 \text{ m}^{-3}$)		Amount of fertilizer (kg ha^{-1})
	Organic carbon (%)		Tillage date (doy)
	Total nitrogen (%)		Tillage depth (cm)
	Sat. hydraulic conductivity (cm h^{-1})		Harvest date (doy)
Weather*	Precipitation (mm)	Crop	Development time (DD/leaf equivalent)
	Maximum temp (Celsius)		Potential kernel number (kernels ear^{-1})
	Minimum temp (Celsius)		Kernel efficiency ($\text{kernels ear}^{-1} \text{ MJ}^{-1}$)
	Solar radiation ($\text{MJ m}^{-2} \text{ d}^{-1}$)		Radiation use efficiency (g MJ^{-1})

* CO2 values in ppm were also modified in the simulations with future climate.

DOY=day of the year

5.4. Results

5.4.1. Model validation-measures of error

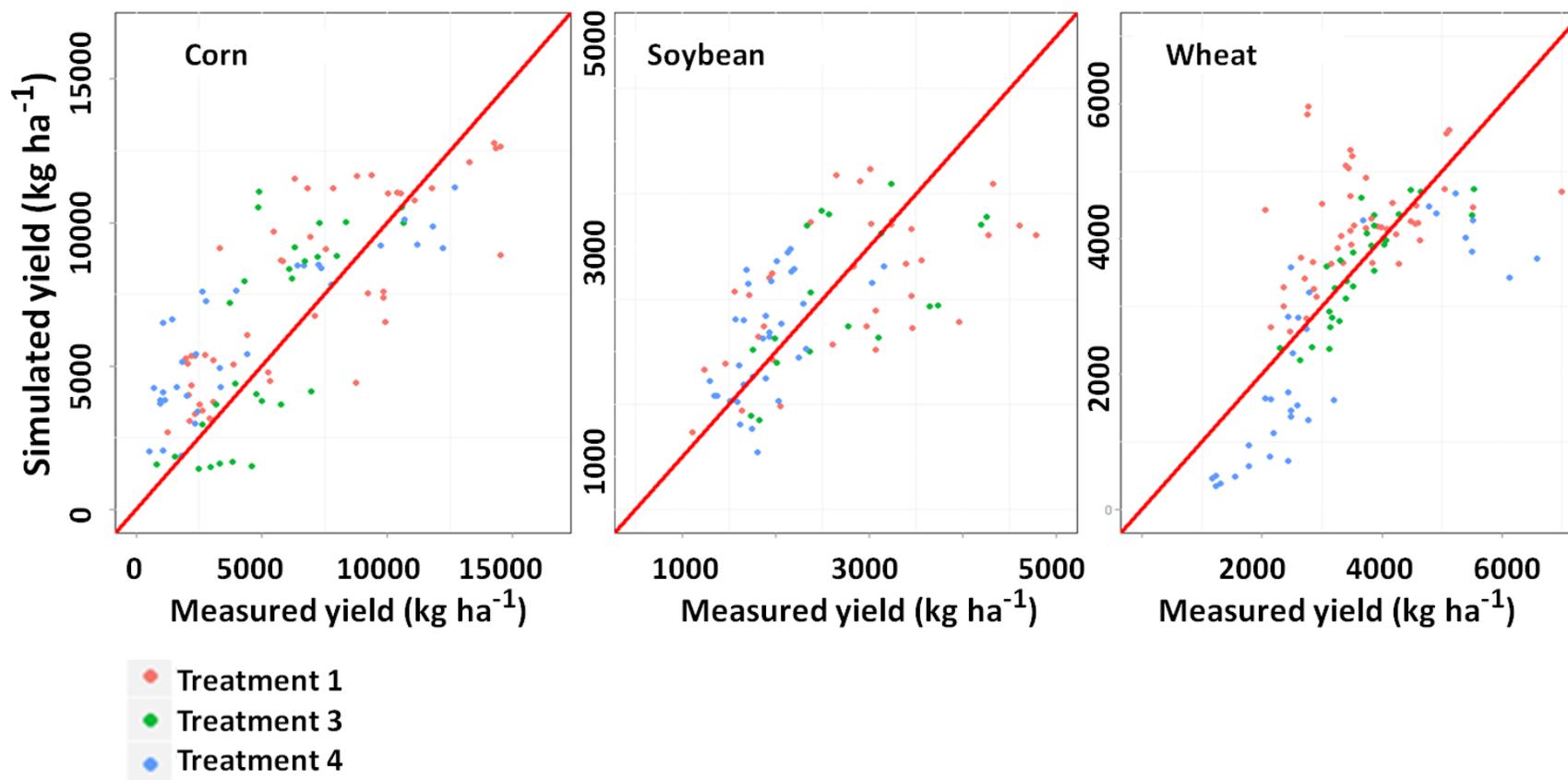
Crop performance of corn, soybean, and wheat was closely predicted by SALUS as indicated by the measures of model error (Table 5.3). The initial (i.e. 'blind') simulation, which use measured soil parameters, historical weather, and crop management data was enough to produce acceptable model predictions of crop yields. The correlation coefficients between measured and simulated yields range from 0.42 to 0.55 (Table 5.3) indicating a moderate association. The absolute error bias (AEB) showed that SALUS tend to overestimate corn yields ($+2.4 \text{ t ha}^{-1}$) and slightly underestimate soybean (-0.11 t ha^{-1}) and wheat yields (-0.15 t ha^{-1}). Model prediction considerably improved once the crop parameters and population density were modified to represent actual conditions of crop variety and crop density at each topographic position. Measures of error in this final validation improved for all crops. The correlation coefficients between measured and simulated yields increase considerably, now showing stronger associations that range from 0.62 to 0.79 (Table 5.3). Figure 5.5 shows the linear association between measured and simulated yields from the final validation model. The AEB showed a reduced bias in corn ($+1.0 \text{ t ha}^{-1}$), soybean (-0.03 t ha^{-1}), and wheat (-0.03 t ha^{-1}) compared to the initial validation. Similarly, the absolute error (AAE) in the final validation showed a reduced error in prediction of approx. 2.09, 0.57, and 0.75 t ha^{-1} for corn, soybean, and wheat respectively. The absolute error standardized (AAES) which allows comparisons of model error between crops indicate that the prediction in corn (0.52) and wheat (0.59) was more accurate than the prediction in soybean (0.73). These results suggest that the final validated

model can be used to produce point estimates of crop yield for a particular condition of topography and management that are comparable to measured yields.

Table 5.3. Measures of model error for the initial and final simulations presented by crop.

Simulation	Crop	ASSE	AEB	AAE	AAES	r
Initial validation	Corn	22.3 x10 ⁶	2401.4	3875.9	0.74	0.51
	Soy	7.2 x10 ⁵	-109.9	699.6	0.89	0.42
	Wheat	15.2 x10 ⁵	-148.4	939.0	0.67	0.55
Final validation	Corn	6.4 x10 ⁶	1036.6	2086.8	0.52	0.79
	Soy	4.8 x10 ⁵	-34.8	570.4	0.73	0.62
	Wheat	10.1 x10 ⁵	-27.5	754.0	0.59	0.68

Figure 5.5. Measured yield vs. simulated yield from the final validated model for: Corn, Soybean, and Wheat. The 1:1 line is presented in red. Dot color represents the agricultural management treatment: T1 (Red), T3 (Green), and T4 (Blue).



5.4.2. Model validation-ANOVA

Analysis of variance on the simulated yields indicates that the validated model is capable to reproduce similar effects to the ones observed in the measured yields. Table 5.4 presents the summary of the ANOVA table (only F-value and probability are shown) from the linear mixed model fitted on simulated and measured yields of corn, soybean, and wheat. Reported F-values and probabilities represent the test of hypothesis for the two-way interaction and the main effects of management treatment and topographical position. Hypothesis testing in simulated yields match closely the significant effects found in the hypothesis testing in measured yields. This is an indication that the validated model not only is capable to reproduce similar point estimates of yield, but also is capable to reproduce the variability observed within a particular management treatments and topographical position. A significant two-way interaction between position and treatment was found in the ANOVA tables of corn and wheat measured yields. However, only the main effect of position was significant in the ANOVA for soybean measured yields. Using the simulated yield values I was able to detect the same significant terms: the interaction between treatment and position in simulated corn and wheat; and the effect of position in simulated soybean. Moreover, comparison of simulated yields in topographical positions and management treatments produced a similar interpretation of significant differences as the comparisons of measured yields. In other words, I was able to reach similar conclusions regarding yield differences with either measured yields or simulated yields. Figure 5.6 shows the simulated and measured least squares means by treatment and position. Mean separation in simulated yields match closely the mean separation in measured yields for all crops, topographical positions, and most of the management treatment. For instance, the significant lower corn yields in slope positions were clearly detected in the simulate data. Significant differences between management

treatments in soybean and wheat yields were also detected in the simulated data. However, with the simulated data I was not able to detect the difference in corn yields between summit-T1 and summit-T4, and the difference in soybean yields between depression-T3 and summit-T3. The overall result from this mean separation is a good indication that I can now use the validated model to predict future crop yields and potentially use these predicted yields to analyze differences between management treatments and topographical positions under future climates.

Table 5.4. ANOVA F-value and associated probability (Pr>F) for the statistical term in the linear mixed model. F-values are presented for the models with measured yield and simulated yields respectively.

Effect	Measured		Simulated	
	F-value	Pr > F	F-value	Pr > F
Corn				
Treatment	9.91	0.113	1.27	0.423
Position	76.75	<.0001	30.17	<.0001
Position*Treatment	3.09	0.021	3.69	0.018
Soybean				
Treatment	0.78	0.564	7.94	0.162
Position	7.01	0.028	29.65	<.0001
Position*Treatment	0.18	0.942	1.29	0.323
Wheat				
Treatment	0.35	0.740	19.36	0.047
Position	9.63	0.003	92.97	<.0001
Position*Treatment	3.62	0.032	5.18	0.001

Figure 5.6. Simulated yield and measured yield by topographic position and management treatment for: a) corn, b) soybean, and c) wheat. Mean separation based on Fisher's LSD ($\alpha=0.05$). Upper-case letters compare management treatments within a particular topographic position. Lower-case letter compare topographic positions within a particular management treatment.

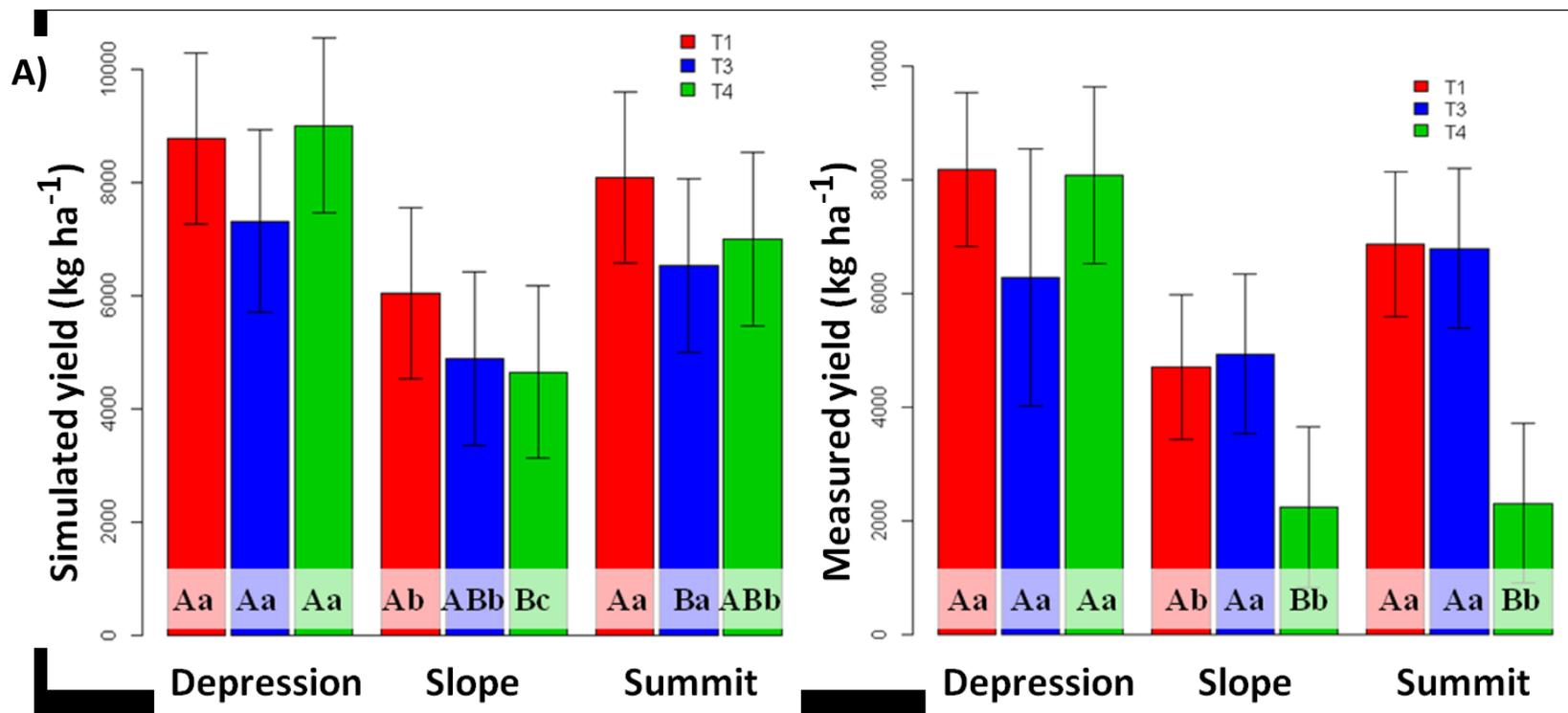


Figure 5.6. (cont'd)

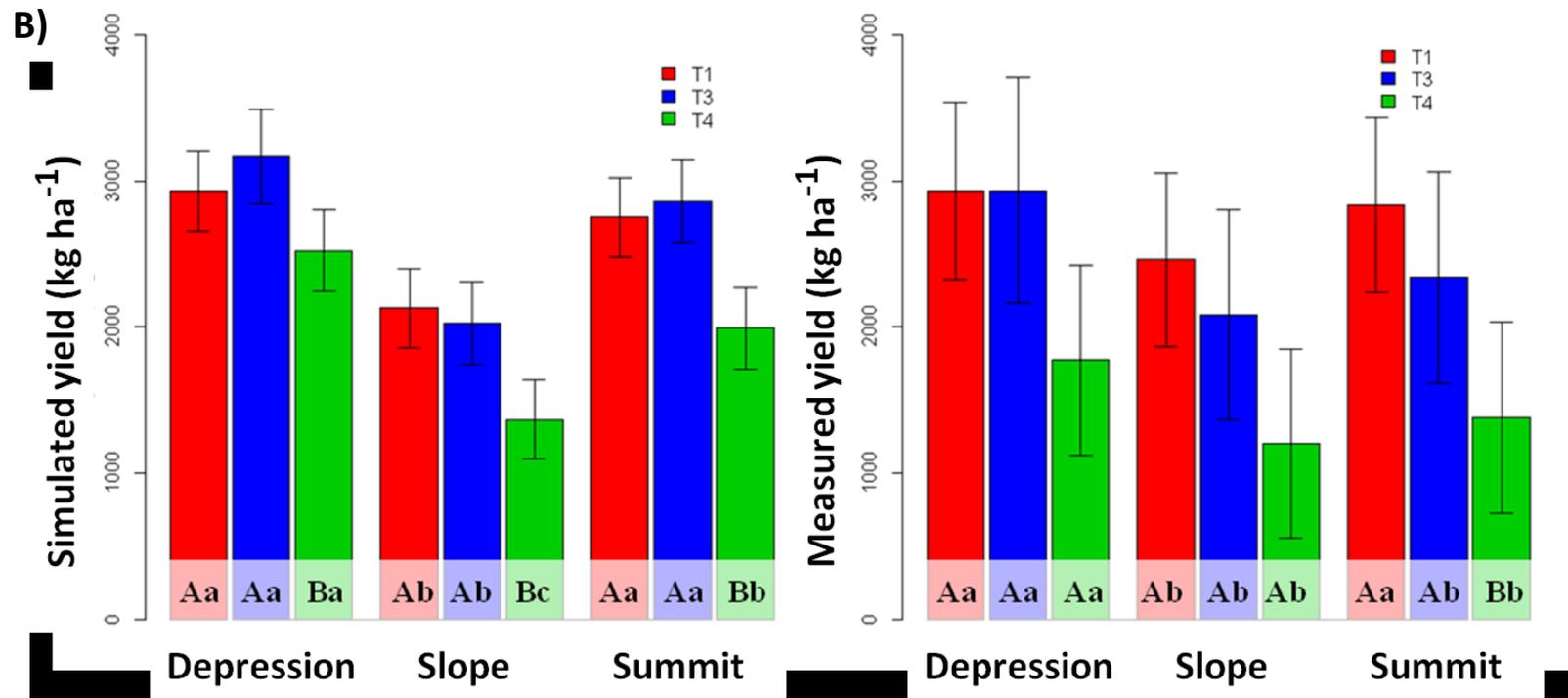
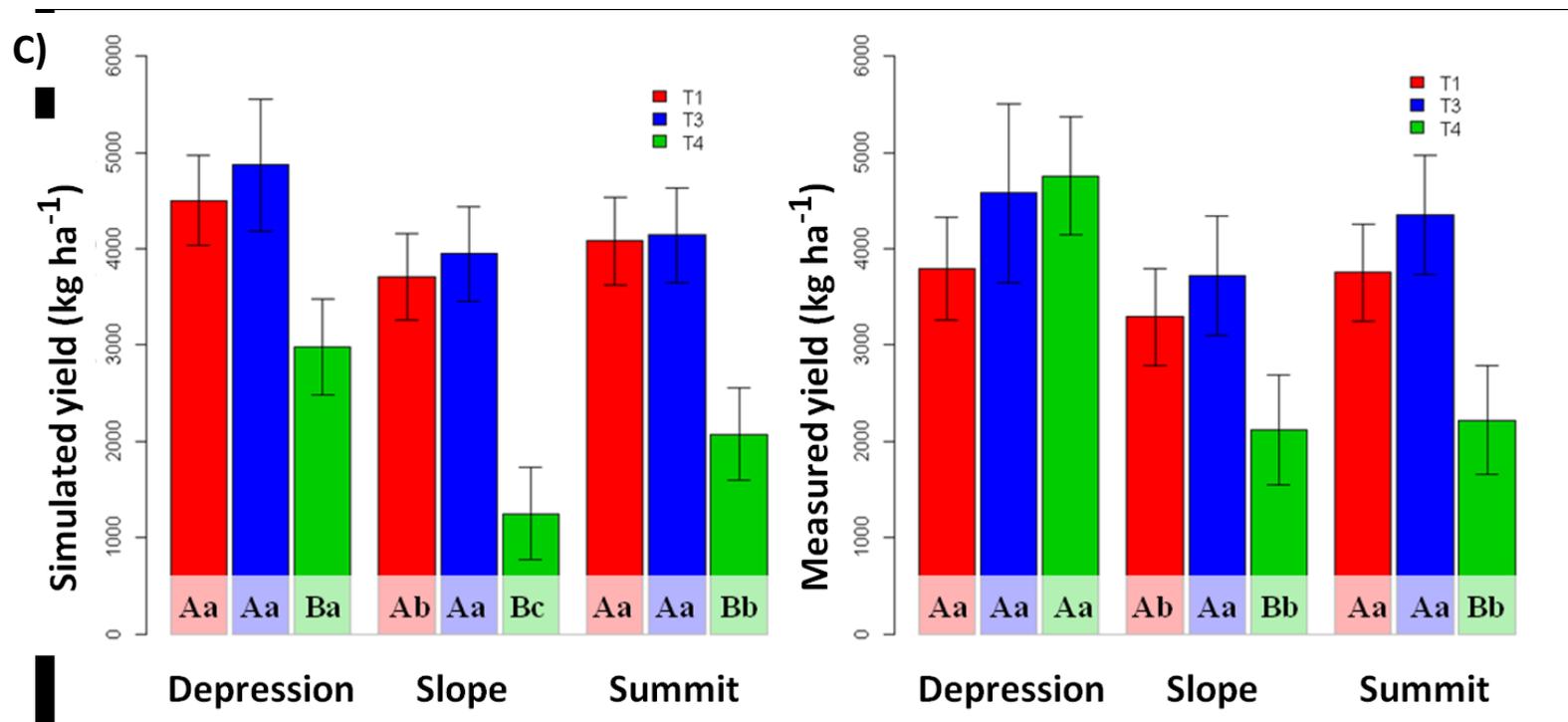


Figure 5.6. (cont'd)



5.4.3. Simulation with future climate scenarios

5.4.3.1. Projections of crop yield

Projections of crop performance under 100 years of future climate scenarios showed a significant decline in corn yields, in particular for the organic-based treatment. On the other hand, soybean and wheat yields showed a slight increment (Fig. 5.7). The analysis of variance showed a significant change of simulated yields across the climate periods for all three crops (Table 5.5). The significant two-way interaction between management treatment and climate period indicates that the change in crops yields observed across time depends on the particular management treatment. In addition, change in soybean and wheat yields across time also depend on the particular topographical position (Table 5.5). Crop yield comparisons between climate periods broadly indicate a significant decline in corn yields for treatment T4, and a significant increase in T1 and T3 for soybean and wheat (Fig. 5.7). In particular I noticed that for corn yields, the treatment T4 (certified-organic) significantly decline after climate period 1 in slope and summit position; and decline after climate period 2 in depression position. The total decline in corn yields in the climate period 3 of about 1.7, 2.3, and 2.1 t ha⁻¹ respects the climate period 1, represent a loss of 35, 55, and 53% in depression, summit, and slope respectively. For soybean, depression and summit position in treatments T1 and T3 increase yields after climate period 1. The total increment of soybean yields in by the climatic period 3 range between 230 to 248 kg ha⁻¹, which represent a gain from 5.9 to 6.5% respect the climate period 1. For wheat, the treatment T1 slightly increase yields after climate period 1 in summit and slope, whereas treatment T3 increase yields in all topographical positions. Wheat yields also decline in the treatment T4. Wheat yields increases in treatment T3 represent a gain from 28 to 39%, whereas yield declines in treatment T4 represents a loss between 14 to 24%.

Yield increase observed in soybean and wheat can be explained by the increased atmospheric CO₂ using in the simulations. Soybean and wheat are C₃ crops that respond positively to CO₂ fertilization. I used a future climate scenario that have an increment of about 3.5 ppm of CO₂ every year, therefore increasing the concentration of CO₂ to 700 ppm by the year 2099. Parry et al., (2004) reported about 15% gain in wheat yield when increasing CO₂ from 350 to 750. They also report about 25% gain in soybean yields with the same increment in CO₂. Similar effect of CO₂ has been reported by Long et al., (2006), in soybean and wheat; and by Hatfield et al., (2011) in several grain cereals. Therefore, C₃ crops can still increase photosynthetic activity and grain yields in increased CO₂ if this offset is not lost due to any stress. However, I noticed lower yields in soybean and wheat in some particular management treatments and positions that can be explained by a higher stress factor. For example, lower soybean yields in slope positions across all treatments are explained by the higher drought stress factor (Fig. 5.8). Projected climate scenarios for the Great Lakes Region have been shown a slightly increase in total precipitation by the year 2070; however the frequency of droughts is expected to increase (IPCC, 2007); and rain events are expected to be more uneven distributed since the number of wet days per year will decline in 13 to 60 days (depend on the climate scenario) respect to what is observed today (Winkler, 2006). Therefore, the higher capacity of water storage in depression and summit position reduce considerably the stress by drought. Similarly, lower wheat yields in the treatment T4 can be explained by the higher nitrogen stress experienced in these treatments (Fig. 5.8). Management strategies to improve soybean yields are

related to water management in particular for the slope positions. In wheat we must procure an improved management of nitrogen in particular for the organic-based system.

In the case of corn I noticed a general decline in yields, more evident in the treatment T4. Maize is a C₄ crop in which the respond to CO₂ fertilization is considerably small (Long et al. 2006), therefore the photosynthetic activity in corn does not increased substantially with increased CO₂ and any gain can easily be overcome by losses caused by stresses. Parry et al., (2004) showed that increasing CO₂ from 350 to 750 produce a gain of only 7% in corn yields. On the other hand, increased temperatures could reduce the life cycle and the duration of the reproductive phase which cause a reduction in grain yield. Hatfield et al., (2011) reported a list of studies that demonstrated the negative effect of increased temperature on corn yields. For example, they reported that an increase of 2 °C in the temperature decreased from 5 to 8% corn yields in the Central Corn Belt region. The decline in corn yields are also explained by the higher nitrogen stress particularly observed in the treatment T4, and summit and slope positions of treatment T3 (Fig. 5.8). Mitigation strategies for corn will require improvements in several management practices. For example, corn yield decline is associated to nitrogen stress but also to increasing temperatures that reduce the life cycle in hybrids of slow development (i.e. certified-organic crops). Therefore, developing new materials according to the new short growing season and improving management of nitrogen could help overcome the projected decline.

Figure 5.7. SALUS projection of crop yields for a) corn, b) soybean, and c) wheat under future climate scenarios. Topographic position is color-coded: depression (green), summit (blue), and slope (red). Projection of mean yield in management treatments are plotted with different line type: continuous (T1), dash (T3), and dotted (T4).

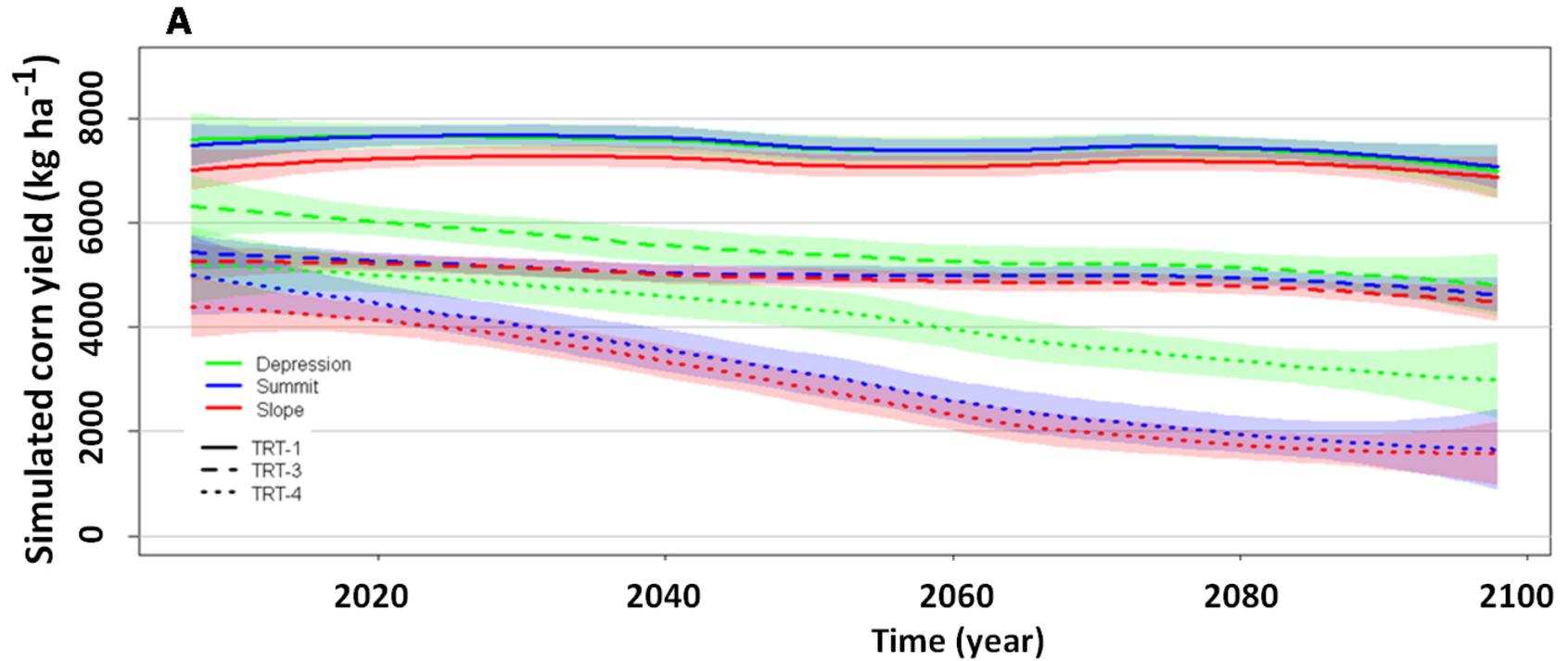


Figure 5.7. (cont'd)

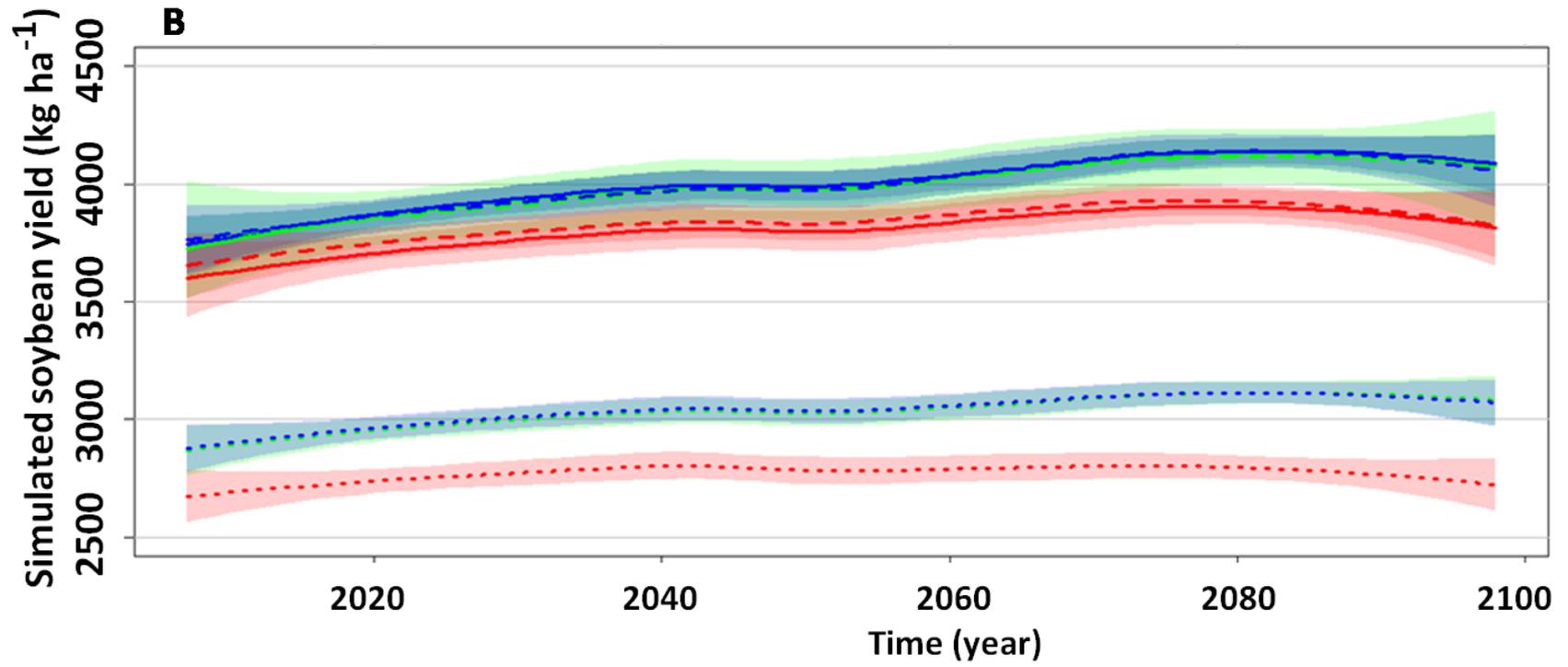


Figure 5.7. (cont'd)

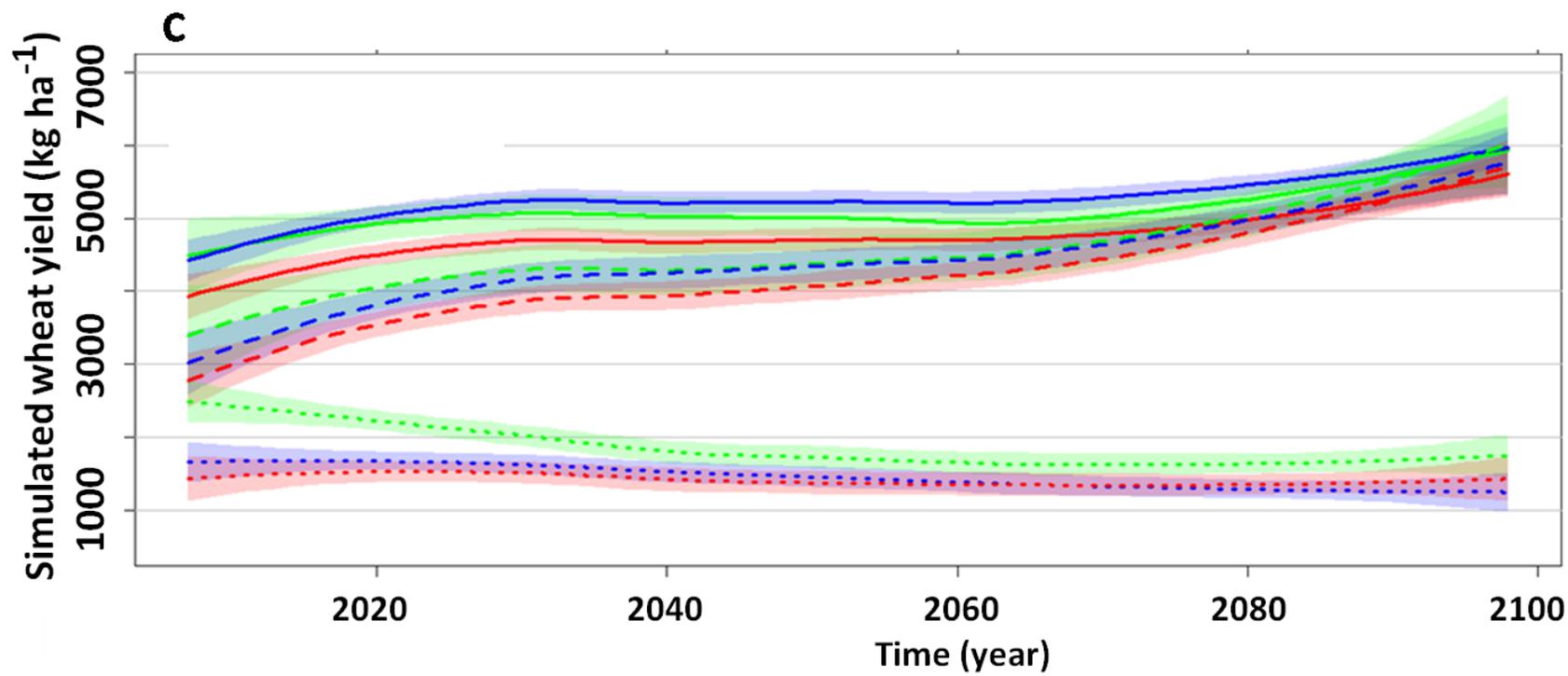


Figure 5.8. Relevant stress factors by management treatment, position and climate period in a) nitrogen stress in corn, b) drought stress in soybean, and c) nitrogen stress in wheat. Mean stress values with different letters within same treatment and position indicate differences between climate periods after Tukey's HSD ($\alpha=0.05$). Stress factors range from 0 to 1, where 0 indicates the maximum stress and 1 indicates no stress.

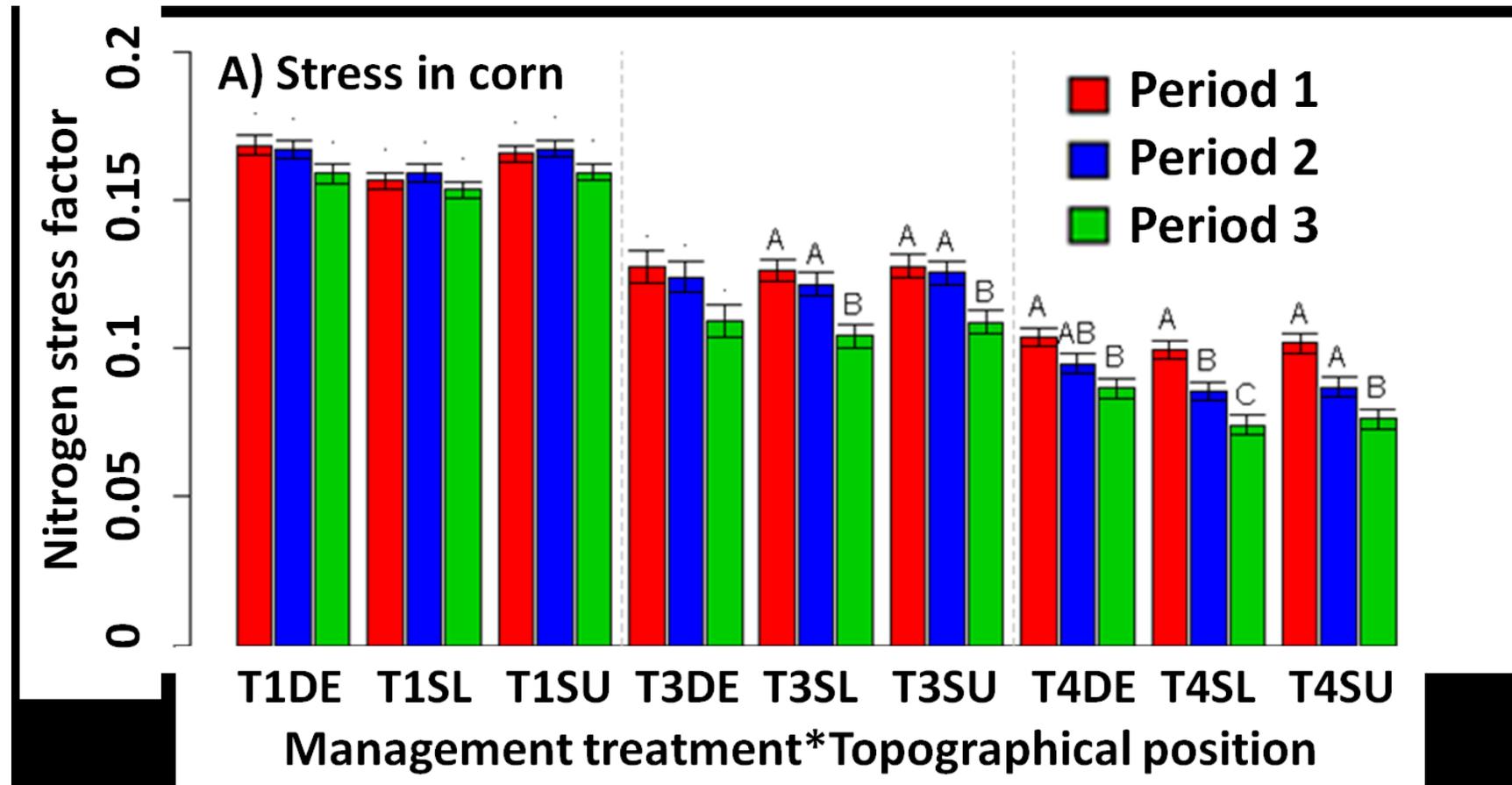


Figure 5.8. (cont'd)

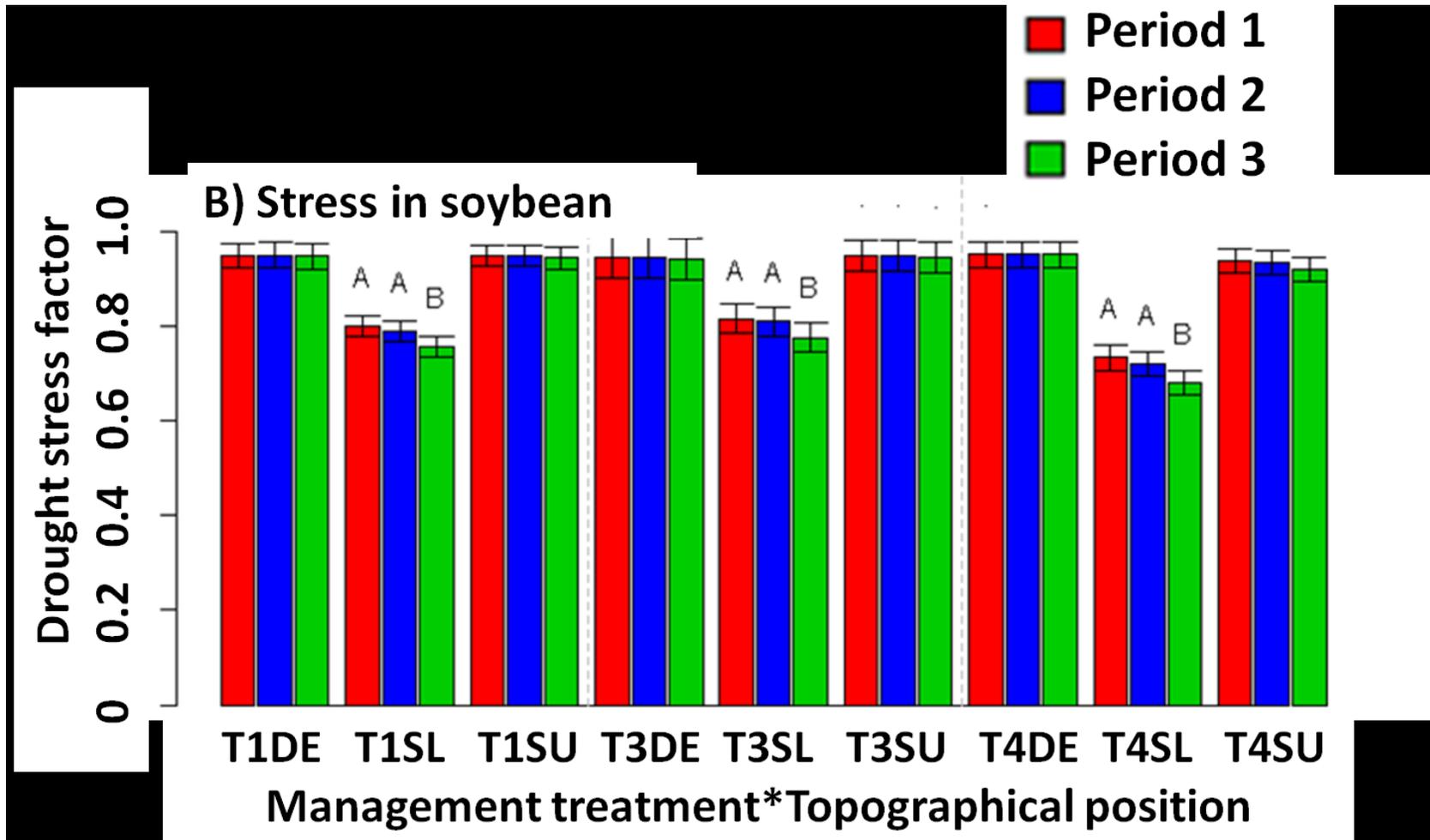


Figure 5.8. (cont'd)

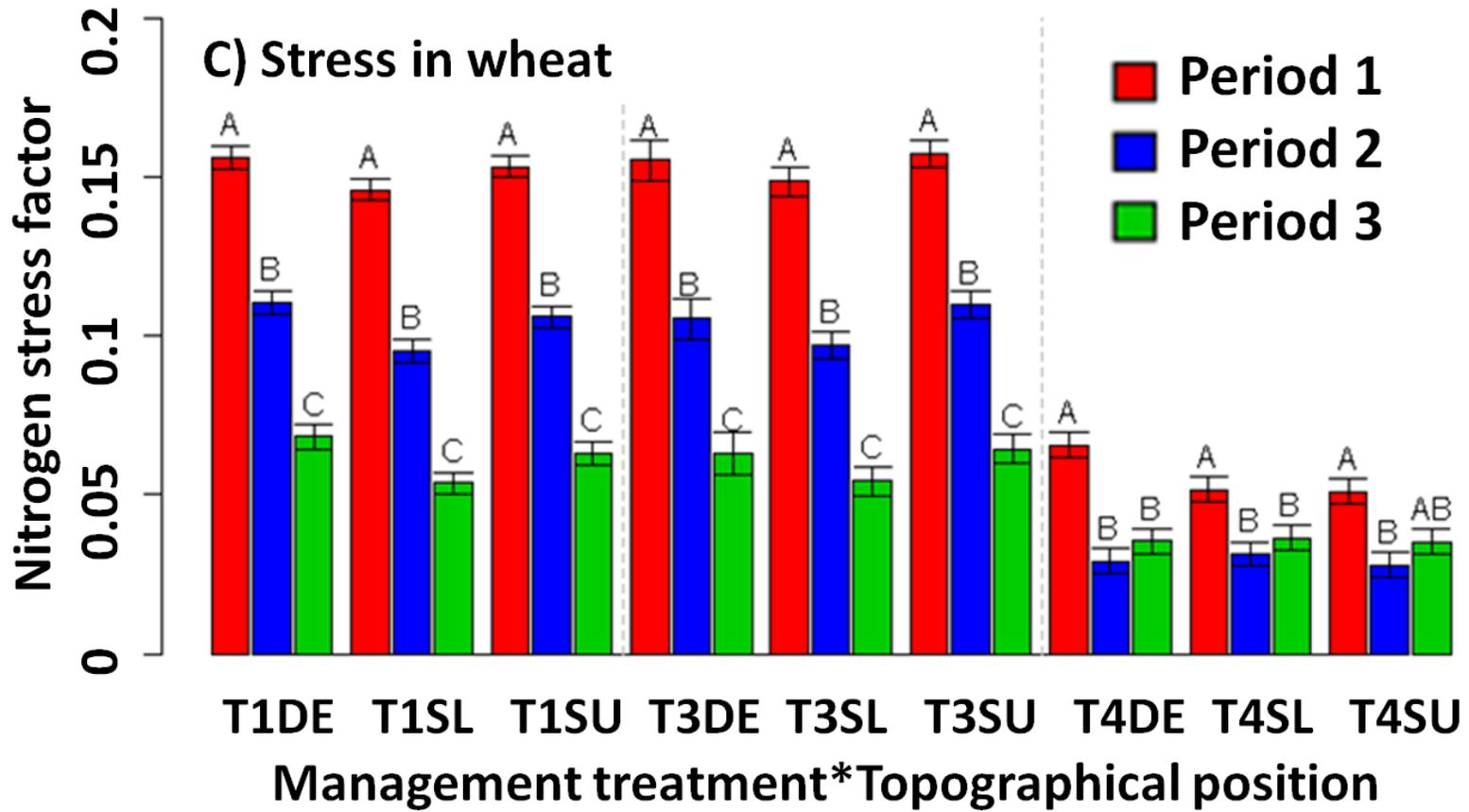


Table 5.5. ANOVA F-value and associated probability (Pr>F) for the linear mixed model on simulated future yields.

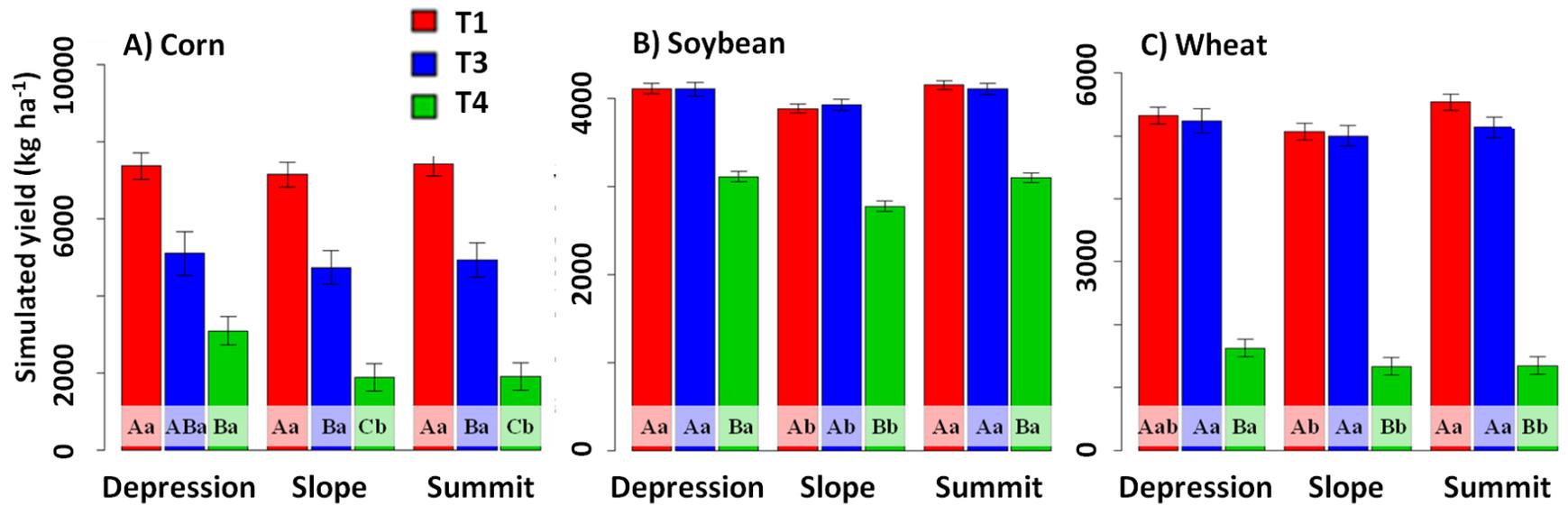
Effect	Corn		Soybean		Wheat	
	F-value	Pr > F	F-value	Pr > F	F-value	Pr > F
Treatment	67.39	<.0001	540.50	<.0001	393.78	<.0001
Position	3.60	0.07	20.39	0.00	22.50	0.00
Treatment*Position	1.15	0.39	1.01	0.45	6.13	0.01
Period	19.48	0.00	6.41	0.01	9.44	0.00
Treatment*Period	28.03	<.0001	32.91	<.0001	102.26	<.0001
Position*Period	0.13	0.97	23.50	<.0001	3.00	0.04
Treatment*Position*Period	1.73	0.09	0.28	0.97	0.28	0.97

5.4.3.2. Comparison of crop yields in the climate period 3 (2070-2099)

Figure 5.9 shows simulated crop yields for each topographical position and management treatment in the period 3 (years 2070 to 2099). Comparison of projected crop yields across management treatments showed significantly lower yields in treatment T4 compare to treatments T1 and T3 for all three crops. Surprisingly, no yield differences were found between treatments T1 and T3 in soybean and wheat, indicating that T3 is an attractive option of agricultural management to reduce the impact produced by conventional systems without losing significant amounts of cash crop. Comparison of projected yields across topographical positions is crop specific. In corn, slope and summit positions produce considerably lower yields compare to the

depression position in the treatment T4. In soybean, significantly lower yields were observed in slope position across all management treatments. In wheat, lower yields were observed in slope and summit positions compare to depression for the treatment T4; and lower yields were obtained in the slopes of treatment T1. Higher drought and nitrogen stress explain the significant losses in crop yields in summit and slope positions; therefore I consider that conservation strategies can be implemented in these areas to mitigate this negative effect. These results indicate that both management treatment and topographical position will play an important effect on future crop yields, therefore they should be considered during the definition of mitigation strategies for climate change.

Figure 5.9. Future simulated yields by topographic position and management treatment in a) corn, b) soybean, and c) wheat. Mean separation is based on Tukey's HSD ($\alpha=0.05$). Upper-case letters compare management treatments within a particular topographic position. Lower-case letter compare topographic positions within a particular management treatment.



5.5. Conclusions

The crop simulation model SALUS have been proved to be an optimal tool to reproduce point estimates of crop yields at diverse topographical conditions and agricultural management systems in Southwest, Michigan. Moreover, I demonstrated that predictions obtained from the simulations are good enough to represent the variability within agronomical treatments and topographical positions. Treatment yield differences observed in simulated experiments matched closely the differences observed in the real experiments. Therefore, I conclude that predicted yields obtained from simulation models can be used to analyze differences between management treatments and topographical positions under future climates.

Projections of crop performance under 100 years of future climate scenarios showed a significant decline in corn yields, in particular for the organic-based management; whereas soybean and wheat yields slightly increased. Increased temperatures and higher nitrogen stress explained the decline in corn yields. Higher temperatures in the late-half of the century will reduce the life cycle of corn causing reductions in grain yield. Increased nitrogen stress in corn was particularly observed in the organic-based management for corn, therefore explaining the larger decline in this treatment. Increased atmospheric CO₂ explained the yield increase observed in the C₃ crops (soybean and wheat) which respond positively to CO₂ fertilization. However, drought stress in the last third of the century negatively impact soybean yields in slope positions; whereas nitrogen stress impact negatively wheat yields after the year 2037 particularly in the organic-based treatments. Mitigation strategies for corn will require the developing of new materials matching the new short growing season and improving management of nitrogen.

REFERENCES

REFERENCES

- Ahuja, L.R., Ma, L., Howell, T.A. 2002. Agricultural system models in field research and technology transfer. CRC Press, USA. 357pp.
- Basso, B., Ritchie, J.T., Pierce, F.J., Braga, R.P., Jones, J.W. 2001. Spatial validation of crop model for precision agriculture. *Agr. Systems*. 68: 97-112
- Basso, B., J. Ritchie, P. Grace, and L. Sartori. 2006. Simulation of tillage systems impact on soil biophysical properties using SALUS. *Ital. J. Ag.* 4:677-688.
- Bergsma, 2013. KBS002: MCSE: Weather Station Sensors. KBS research protocols Available at <http://lter.kbs.msu.edu/protocols/8> (Accessed: 2013, Aug 26)
- Carter, T. R., La Rovere, E. 2001. Developing and applying scenarios. In: McCarthy, J.J., Canziani, O.F., Leary, N.A., Dokken, D.J. White, K.S. (eds) *Climate change 2001*. Cambridge University Press, pp. 145–190.
- Cleveland, W., Grosse E., Shyu, W.M. 1992. Local regression models. In: *Statistical Models in S*. eds J.M. Chambers and T.J. Hastie, Wadsworth & Brooks/Cole.
- Crum, J.R., Collins, H.P. 1995. KBS soils. Kellogg Biological Station Long-Term Ecological Research. Available at <http://lter.kbs.msu.edu/research/site-description-and-maps/soil-description> (Accessed: 2013, Aug 2)
- Gee, G.W. Bauder, J.W. 1986. Particle-size analysis. In A. Klute (ed.) *Methods of soil analysis*. Part 1. Physical and Mineralogical Methods. 2nd ed. Agron. 9. ASA/SSSA, Madison, WI.
- IPCC, 1995. *Climate change 1995*. J.T. Houghton, L.G. Meira Filho, B.A. Callander, N. Harris, A. Kattenberg, and K. Maskell, Eds., Cambridge Uni. Press, UK, 531pp.
- IPCC, 2007. *Climate change 2007*. M. Parry, O. Canziani, J. Palutikof, P. van der Linden and C. Hanson, Eds., Cambridge Uni. Press, UK, 976pp.
- Hatfield, J.L., Boote, K.J., Kimball, B.A., Ziska, L.H., Izaurralde, R.C., Ort, D., Thomson, A.M., Wolfe, D. 2011. Climate Impacts on Agriculture: Implications for Crop Production. *Agron. J.* 103(2): 351-370
- Long, S.P., Ainsworth, E.A., Leakey, A.D., Nosberger, J., Ort, D.R. 2006. Food for thought: lower-than-expected crop yield stimulation with rising CO₂ concentrations. *Science*. 312: 1918-1921.
- LTER, 2013. Baseline Soil Sampling and Analysis. GLBRC Intensive and Scale-Up Sites. Available at <http://lter.kbs.msu.edu/protocols/158> (Accessed: 2013, Aug 26)

- Munoz, J.D., Kravchenko, A. 2012. Deriving the optimal scale for relating topographical attributes and cover crop plant biomass. *Geomorphology*, 179: 197-207
- Munoz, J.D., Steibel, J.P., Snapp, S., Kravchenko, A.N. 2013. Cover crop effect on corn growth and yield in agricultural fields is influenced by topography. *Agr. Ecos. Env.* (accepted)
- Palutikof, J.P., Winkler, J.A., Goodess, C.M., Andresen, J.A. 1997: The simulation of daily time series from GCM output. Part 1: Comparison of model data with observations. *Journal of Climate*, 10, 2497-2513.
- Parry, M.L., Rosenzweig, C., Iglesias, A., Livermore, M., Fischer, G. 2004. Effects of climate change on global food production under SRES emissions and socio-economic scenarios. *Global Env. Change*, 14: 53-67
- R Development Core Team. 2009. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org>.
- Ritchie, J.T., Gerakis, A., Suleiman, S. 1999. Simple model to estimate field-measured soil water limits. *Trans. ASAE*. 42(6): 1609-1614
- Ritchie J.T., Singh U., Godwin D.C., Hunt L. 1989. A user's guide to CERES Maize-V.2.10. Muscle Shoals, AL. International Fertilizer Development Center.
- Robertson, P., Bronson, J. 2011. KBS059: LTER Scale-up experiment. LTER research protocol. Available at <http://lter.kbs.msu.edu/datatables> (Accessed: 2011, Oct 18).
- Robertson, P., Kravchenko, S., Sippel, S. 2012. KBS056: Raw Geo-referenced Agronomic Yields. LTER research protocols. Available: <http://lter.kbs.msu.edu/datatables/297> (Accessed: 2012, Nov 20).
- Rosenzweig, C, Parry, M.L. 1994. Potential impact of climate change on world food supply. *Nature*, 367: 133-138.
- SALUS software. 2013. System Approach for Land Use Sustainability. (Accessed: 2013, Aug, 25. <http://nowlin.css.msu.edu/salus/>)
- Schimel, D. 2006. Climate change and crop yields: beyond Cassandra. *Science*, 312: 1889-1890.
- Simmons, J. 2012. KBS004: Agronomic Protocol: Main Cropping System Experiment. LTER research protocols. Available at <http://lter.kbs.msu.edu/protocols/104> (Accessed: 2012, Jun 20).
- Sudduth, K., Drummond, S. 2007. Yield editor: Software for removing errors from crop yield maps. *Agron. J.* 99:1471-1482.
- Tsuji, G.Y., Hoogenboom, G., Thornton, P.K. 1998. Understanding options for agricultural production. Kluwer Academic Pub. GB. 399pp.

Venables, W., Ripley, B. 2002. *Modern Applied Statistics with S* (4th ed.). Springer, USA. 495pp.

Winkler, J. 2006. Pileus Project: Climate science for decision support. Available at <http://pileus.msu.edu/about/resources/presentations.htm> (Accessed: 2013, Aug 20).

Winkler, J., J. Palutikof, J. Andresen, C. Goodess, 1997: Simulation of Daily Temperature Time Series from GCM Output. *J. Climate*, 10, 2514–2532.

Winkler, J.A., Guentchev, G.S., Perdinan, Tan, P., Zhong, S., Liszewska, M., Abraham, Z., Niedźwiedź, T., Ustrnul, Z. 2011: Climate scenario development and applications for local/regional climate change impact assessments: An overview for the non-climate scientist. Part I: Scenario development using downscaling methods. *Geography Compass*, 5/6, 275–300.

Zavalloni, C., Andresen, J.A., Winkler, J.A., Flore, J.A., Black, R., Beedy, T.L. 2005. The Pileus Project: Climatic impacts on sour cherry production in the Great Lakes region in past and projected future time frames. *Acta Horticulture*.

Chapter 6: Conclusions

Assessing cover crop biomass across the landscape was required to study the effect of cover crops in agricultural ecosystems. Red clover biomass was derived fast and inexpensively from the Normalized Difference Vegetative Index (NDVI). I demonstrated that NDVI is a reliable source for predicting red clover biomass across the field landscape. The nonlinear relationship between NDVI and biomass was adequately represented by the Richards function. Moreover, the best model fit was found with the Richard function after accommodating the non-constant variance observed in the residuals. From a practitioners' perspective these results indicate that addressing heteroscedasticity using a functional variance can improve model fit and predictions of red clover biomass. These predictions and associated uncertainty are critical data products that will ultimately help to quantify the effects of cover crop biomass on main crop yields. In addition, I showed that nonlinear hierarchical models are an interesting alternative to model different sources of variability. Future investigations of cover crop effect in agricultural systems can take advantage of this approach to include experimental factors in their statistical models. I encourage the use of hierarchical structures whenever field variability is relevant, particularly for experiments that include fields from different environmental and soil conditions.

Since my main objective was to relate topography with red clover biomass and corn yields, I studied these relationships at different geographical scales. My results indicated that the strength of the relationship between topography and biomass varied with the scale of geographical derivation. Changes in the strength of the relationship were explained by the change of information across geographical scale. For example, curvature, flow accumulation and flow length were more affected by the scale of derivation than relative elevation, slope and the

potential solar radiation index. As a result, relative elevation, slope and potential solar radiation were significant predictors for biomass at most studied scales; while flow accumulation, curvature, and flow length were significant predictors for biomass mainly at large scales. I concluded that predictive models that use topographic attributes will be most accurate when used at the optimal scales. In my study, I showed that increasing the scale tended to increase the strength of the relationship. For example, the scales 20 and 40 m led to the regression models with better prediction performance across most of the studied fields. Therefore, I conclude that using the original DEM for analyzing the relationships between topography and cover crop biomass, may not always be appropriate, since not all possible associations with topographic attributes are depicted at the smallest scale. Thus, prior to developing predictive regression models I recommend exploring the relationships between topography and biophysical variables of interest at a range of scales. For future works, a good starting point is to evaluate the topographic “complexity” of the area since it brings indication of where filtering is most needed. Topographic “complexity” can be assessed by using the loading values of principal components, and the semivariogram parameters of the topographic and studied biophysical variables. For example, I showed that partial sill and nugget/sill ratios in terrain slope semivariograms are good indicators of the optimal derivation scale.

The previous conclusions regarding biomass prediction and identification of optimal scale for derivation were the basis for studying the effect of red clover biomass on crop yields in fields with diverse terrain. I derived red clover biomass and topographic attributes maps at the optimal scale to study the multiple relationships between topography, red clover biomass, and corn yields. I showed that topography has significant effect on red clover biomass and corn yields; in particular, attributes associated with water accumulation and water availability played

an important role in corn production under dry conditions, whereas attributes related to drainage and water re-distribution played a more important role under wet conditions. I also showed that red clover biomass had a positive effect on corn yield, but, the magnitude of that effect varied temporally and spatially. For example, the positive effect of red clover was significant only in the years with average precipitation, and the magnitude of this effect was clearly more pronounced in summit and slope than in depression positions. Since the independent effect of red clover biomass on corn performance was more pronounced on slopes and summits, I conclude that it is worth the effort to ensure a good cover crop stand on summit and slope areas, because this is where it can be most beneficial to the subsequent cash crop. Efforts to increase red clover biomass in summit and slope positions supported gains in corn yield, helping to reduce the gap in corn yields between unfavorable (summit and slopes) and favorable (depression) areas. This study also shows that using hierarchical structures to account for the variability within experimental factors, in particular for field and year effects, considerably improved the model fit and my understanding of the agricultural ecosystems. Therefore, I encourage the use of hierarchical models to include different sources of variability when analyzing the relationships between topography, red clover, and corn yield.

The newly generated knowledge about the studied agricultural ecosystems was used to simulate crop production in current and future climate scenarios. I showed that simulated values closely represent the variability observed within agronomical treatments and topographical positions. Therefore, I conclude that predicted yields obtained from simulation models can be used to analyze differences between management treatments and topographical positions under future climates. Projections of crop performance under future climate scenarios showed a general decline in corn yields; while soybean and wheat yields slightly increased. I demonstrated that

increased temperatures and higher nitrogen stress explained the decline in future corn yields. Therefore, potential strategies for corn will require the development of new genetic materials and improving management of nitrogen. On the other hand, increased atmospheric CO₂ explained the general increase in yields observed in soybean and wheat. However, drought stress in the last third of the century negatively impact soybean yields in slope positions; whereas nitrogen stress impact negatively wheat yields particularly in the organic-based treatments.

This study contributes to understand the multiple interactions between agricultural management systems and topography. I showed that the use of cover crops substantially contribute to the performance of the following cash crop. It is important to highlight the fact that agricultural systems are dynamic in space and time; therefore the effect of cover crop biomass varied spatially and temporal. Moreover, the potential impact of conservation agricultural practices (i.e. use of cover crops) may depend on the particular landscape position, weather conditions, and management system. I showed that slope and summit positions will benefit the most by the use cover crop practices. This potential effect will be even more important under the future climate. For example, crops like corn and wheat will benefit from practices that aim the maintaining or enhancing of nitrogen levels in the soils. On the other hand, crops like soybean will benefit from practices that improve water storage to mitigate water stress. The use of cover crops will certainly contribute to mitigate these sources of stress.