PROPOSED METHOD FOR STATISTICAL ANALYSIS OF AN ON-FARM SINGLE TREATMENT TRIAL.

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Master of Science

by

Byung Jae Cho

August 2021

ABSTRACT

On-farm experimentation (OFE) has allowed farmers to improve crop and land management and increase their productivity over the years. One of the most prevalent OFE designs has been the randomized complete block design (RCBD) with full field length strips as individual plots, replicated a minimum of three-time. This design, commonly referred to as on-farm strip trials pose challenges for both farmers and scientists, due, among others, to intense planning requirements and limited statistical power. Single-strip evaluation trials are easier to implement for farmers, allowing more farmers to participate in on-farm research. Lack of replication of strips with a single-strip design has limited use of results but with growing availably of yield monitor systems that collect yield data every second that a harvester is in the field, re-evaluation of statistical approached to evaluating such designs is needed. The objective of this research, therefore, was to build a statistically appropriate framework for analyzing single-strip trials for improved OFE for both scientists and farmers.

Analyzing a single-strip treatment trial poses challenges due to spatial and temporal yield variability that could potentially interact with the treatment alternative being analyzed. To control for temporal variability, multiple years of geo-referenced yield data need to be analyzed. The first chapter of the thesis explores the three most commonly-used spatial estimation methods, namely nearest neighbor (NN), inverse distance weighting (IDW), and kriging, for converting point data gathered with yield monitors to regular, grid-based, raster maps, which are necessary for temporal analysis of yield. Seven spatial estimation methods (NN, IDW using 10, 20, 30 and all data points and kriging with exponential and Matérn covariance functions) were evaluated to determine the method that most accurately captures intra-field spatial variability of corn silage and corn grain yield in New York. Normalized root means squared error (NRMSE)

was used to evaluate the accuracy of the spatial estimation methods. Kriging with the Matérn covariance function resulted in the most accurate corn silage and grain yield raster maps at both the farm and field levels. There were statistically significant differences in NRMSE between kriging with the Matérn isotropic covariance function and all other models for both corn silage and grain, regardless of field size, year, timing of data collection, or farm that supplied the data.

While it is known that multiple years of yield maps are needed to delineate farm-specific management zones, the effect of number of years of data on the management zone delineation has not been studied. The second chapter explores this topic. Multiple years of corn silage and grain yield were analyzed to calculate average temporal yield and yield variability, as impacted by the number of years of yield data included for the computation. Results showed that data should first be examined for the presence of a yield trend when estimating average yield. Including only the most recent four to five years of data reduced the risk of misrepresenting the expected average yield when there was a yield trend. Years that had extreme weather could greatly impact average yield measurements for farms with fewer years of data. The temporal standard deviation in yield was most consistent when all years of data were included. We concluded that farms interested in developing yield stability management zones should use at least four to five years of yield data and continue to add new years of data for improved delineation of zones over time.

The third chapter explores statistical frameworks for estimating the effect of the treatment from a single-strip OFE, using georeferenced yield monitor data and a historic yield record of the farm. This study analyzed data from single strip-treatment trials on six site-years in 2018 and 2019 for two farms located in Central New York. We examined two approaches, namely Least Squares and Generalized Least Squares with spatial covariance, for estimating the effect of the

treatment, and two approaches, with the estimation assuming independence and spatial covariance, for estimating the standard errors. Results suggested Least Squares approach should be used for treatment effect estimation, while spatial covariance should be assumed when estimating the standard errors for single strip treatment trials with high resolution spatial yield data, historic yield monitor data, and yield stability-based management zones. This single-strip spatial evaluation approach allows for more field trial data to be added over time for a better understanding of drivers for outcomes such as yield and the need for site-specific management in unstable yielding zones.

BIOGRAPHICAL SKETCH

Byung Jae (Jason) Cho is currently in his 2nd year of study in pursuing his Master of Science degree in the Department of Animal Science at Cornell University and plans to graduate in August 2021. Jason attained his undergraduate degree in Statistics and a minor degree in Computer Science at Cornell University. After graduating with the Bachelor's degree, he worked as a data analyst intern at Cornell Statistical Consulting Unit during which time he helped researchers around Cornell with performing statistical analysis. Jason became interested in the application of statistics in agronomy after briefly working with Professor Quirine Ketterings who leads the Nutrient Management Spear Program in the Animal Science department at Cornell University. Jason's research interest is in building statistical tools that would allow farmers to make more informed decisions in managing their fields. After finishing his Master's degree, Jason will be starting his Ph.D. program in Statistics at Cornell University in the fall of 2021.

Dedicated to Hansol, who has brought me joy, giggles, and warm encouragement during harsh and cold winters in Ithaca.

ACKNOWLEDGMENTS

My research throughout my Master's degree program was funded with grants from USDA-NIFA, the Northern New York Agricultural Development Program (NNYADP), New York State Corn Growers Association (NYSCGA), and federal formula funds. We thank all the farmers for sharing yield data; Benjamin Lehman, Dilip Kharel, and Sheryl Swink for yield data processing; Celia Walden for helping me with data analysis; Tulsi Kharel, Sunoj Shajahan and Manuel Marcaida for reviewing data summaries and drafts; Quirine M. Ketterings, Joe Guinness, Jan van Aardt for their continuous support of my research; and finally, everyone at Nutrient Management Spear Program for constructive feedback and support.

TABLE OF CONTENTS

LIST OF FIGURES

Figure 2.5: Zone maps for a field in Grain A (referred to as Field A) as of 2019 using the most recent 3 to 12 years (Figure 2.5(a) – 2.5(j)) and a field in Silage A (referred to as Field B) using the most recent 3 to 11 years (Figure 2.5(a) – 2.5(i)). Zone 1 is colored green, zone 2 blue, zone 3 yellow, and zone 4 red. Zone 1 represents areas that are high yielding (above the farm average) with low year-to-year variability in yield (stable); zone 2 is high yielding as well but variable over years (unstable); zone 3 has low yields and high variability; and zone 4 has low yield and low variability. Difference in zone maps between two successive years are showed in Figure 2.5(k) to 2.5(s) and figure 2.5(j) – 2.5(k). Area colored in red is the part where the zone delineations differ in two successive years (e.g., L 4 vs L 3 represent difference in zone delineations using the latest four vs three years of data).

Figure 3.1: An example of a randomized compete block design (RCBD) to evaluate if a treatment impacted outcomes, such as yield. In this example there is one control and one treatment strip per block (replication) and the trial is replicated three times. When farm equipment permits, strips serve as experimental units (EUs), which typically cover the length of a field (Kyveryga et al., 2015).

Figure 3.2: Placements of N treatment strips for six fields (black) surrounded on both sides by control strips (grey) for six fields on a cash grain operation (Fields A, B, C) and a dairy farm (Fields D, E, F).

Figure 3.3: Density plot representing yield distributions of a N treatment strip (higher N) and the neighboring control strips. The dotted black and the solid grey vertical lines represent the average yield for the N treatment strip and the control strips, respectively.

Figure 3.4: Density plots representing corn yield distribution of N treatment and control strips in individual fields with up to four yield stability zones, as per Kharel et al. (2019). The dotted black and the solid grey vertical line represent the average yield for the treatment and control strip, respectively.

Figure 3.5: Spatial distribution of yield (Figure 3.5a), historic average yield (Figure 3.5b), temporal standard deviation (Figure 3.5c) and management zones (Figure 3.5d), as estimated based on previous years of yield produced for six fields (A through F). Yield stability management zones were delineated as described in Kharel et al. (2019). Zone 1 represents high-yielding-stable, Zone 2 represents high-yielding-unstable, Zone 3 represents low-yielding-unstable, and Zone 4 represents low-yielding-stable areas of the field.

Figure 3.6: Spatial distribution of coefficients for estimating treatment effect of N treatment per yield-stability based management zone (Zone 1, 2, 3, and 4) via Least Squares method (LS) and Generalized Least Squares (GLS) method for Field A. Coefficients for LS were calculated by solving $(X^T X)^{-1} X^T$ and for GLS by solving $(X^T \Sigma^{-1} X)^{-1} X^T \Sigma^{-1}$, where X represents the design matrix and Σ the spatial covariance.

Figure 3.7: The effect of N treatment and standard errors of Zones 1(d), Zone 2(b), Zone 3(a), and Zone 4(c) for Field E was estimated using 60%, 70%, 80%, and 90% of randomly sampled data, 100 times each, using the Generalized Least Squares approach and the Least Squares approach, assuming spatial covariance versus independence.

LIST OF TABLES

Table 3.1: Crop type, year, field size, average yield production, information about the nitrogen strips, soil types, and distribution of management zones, delineated as suggested by Kharel et al. (2019a), were summarized.

Table 3.2: A regression model with the yield as the response variable, temporal average yield, normalized latitude, longitude, linear interaction between latitude and longitude, four corn management zones delineated as suggested by Kharel et al. (2019) and their interaction with N treatment as the explanatory variables were summarized for six fields (Field A, B, C, D, E, and F). For LS 1, the Least Squares was used for treatment effects and independence was assumed for standard errors. For LS 2, the Least Squares was used for treatment effects and spatial covariance was used for standard errors. For GLS, the Generalized Least Squares was used for treatment effects and spatial covariance was used for standard errors.

LIST OF ABBREVIATIONS

CV, Coefficient of Variation; EC, Electrical Conductivity; GP, Gaussian Process; GNSS, global navigation satellite systems; IDW, inverse distance weighting; NDVI, Normalized Difference Vegetation Index; NN, nearest neighbor; NRMSE, normalized root mean squared error; OFE, On-farm experiment; RCBD, Randomized Complete Blocks Design; US, United States; APC, Absolute percent change.

CHAPTER 1:

SPATIAL ESTIMATION METHODS FOR MAPPING CORN SILAGE AND GRAIN YIELD MONITOR DATA

ABSTRACT

Harvester-mounted yield monitor systems are increasingly used to document corn (Zea mays L.) yield. The three most commonly used spatial estimation methods to convert point data gathered with yield monitors to regular, grid-based, raster maps include nearest neighbor (NN), inverse distance weighting (IDW) and kriging. Seven spatial estimation methods (NN, IDW using 10, 20, 30 and all data points and kriging with exponential and Matérn covariance functions) were evaluated to determine the method that most accurately captures intra-field spatial variability of corn silage and corn grain yield in New York. Yield monitor data from two dairy farms and two grain operations were cleaned using Yield Editor prior to spatial analyses. The dataset included 7-10 years of data per farm for a combined 7,484 ha (245 fields) of silage and 6,971 ha (253 fields) of grain. Data were split into training (80%) and cross-validation datasets (remaining 20% of the data). Normalized root mean squared error (NRMSE) was used to evaluate the accuracy of the spatial estimation methods. Kriging with the Matérn covariance function resulted in the most accurate corn silage and grain yield raster maps both at the farm and field level. There were statistically significant differences in NRMSE between kriging with the Matérn isotropic covariance function and all other models for both corn silage and grain, regardless of field size, year when data were obtained or farm that supplied the data. These results are beneficial to ensure accurate and precise spatial mapping of yield products toward optimized corn growth management.

INTRODUCTION

Corn is a major crop in New York, with more than 400,000 ha planted annually. In 2019, 220,000 ha (55%) were combined for grain, while 180,000 ha (45%) were harvested for corn silage (USDA 2019a). Corn silage, typically grown in rotation with hay, is especially important to the dairy industry in New York; the state is ranked third in dairy production in the United States, following California and Wisconsin, and fourth in corn silage production, following Wisconsin, California and Minnesota (USDA 2019b).

Being able to measure corn silage and grain yield at the field and within-field levels is important, as understanding yield and variability in yield over time allows for better inventory management, production optimization and improved allocation of limited resources, such as seed and fertilizer (Long et al. 2016). With greater accessibility and affordability of yield monitor systems, more corn producers are now gathering spatially explicit yield monitor data with flow and moisture sensors that record readings every second as the harvester travels through a field. The availability of spatial data over multiple years allows for analyses of both spatial and temporal variability of yield (Kharel et al. 2019a). Such knowledge can help build actionable insights to better manage nutrients and increase yield (Maestrini & Basso 2018a).

Raw yield monitor data cannot be used right away, however, because the data not only reflect systematic yield variation within a field, but also measurement errors associated with yield-mapping (Dobermann & Ping 2004; Vega et al. 2019). Kharel et al. (2019b) suggested that three main factors cause systematic measurement errors even when proper calibration procedures are implemented: (1) sensor delays, (2) velocity calibration and (3) human errors. Delays exist because the main sensors in yield monitor systems (flow rate sensors, moisture sensors and global navigation satellite systems [GNSS] units) are embedded at different locations on harvest

equipment, which causes flow and moisture values to be out of sync with corresponding GNSS readings. Velocity calibrations also heavily affect the data, as harvest equipment is calibrated for a certain velocity range (Arslan & Colvin, 2002). Theoretically, measurement errors from inadequate velocity calibrations can be reduced by driving the equipment with constant travel speed, but due to irregular field shapes and variation in elevation of many fields in New York, such practice is highly impractical. Human errors occur, among others, when the operator does not raise the harvester head after completion of a pass, in which case the pass number will not be updated in the dataset. This can cause overlapping passes and hence artificially low yield around field edges. Thus, post-harvest yield data correction and cleaning algorithms need to be applied to reduce measurement errors (Arslan & Colvin 2002; Blackmore 1999; Kharel et al. 2018; 2019b).

Yield monitor datasets consist of irregularly placed point estimates of grain flow, moisture and yield estimates; such irregularities are caused by differences in field shape, size and harvest patterns within a field. Researchers often use a rasterized yield map based on yield monitor data as a base layer in delineating zones for better field and resource management (Basso et al. 2007; Blackmore 2000; Brock et al. 2005; Buttafuoco et al. 2017; Diker et al. 2004; Hornung et al. 2006; Kharel et al. 2019a; Khosla et al. 2008) or as a means to understand variability in yield with regards to soil type, elevation and other topographical information (Anderson-Cook et al. 2002; Cox & Gerrard 2007; Kitchen et al. 1999; Maestrini & Basso 2018a 2018b; Yang et al. 2001) (Table 1.1). Yield data are typically not collected at the same GNSS locations each year, but once point data are translated into raster maps using regular grid cells, temporal yield variation can be analyzed with multiple years of data (Kharel et al. 2019a).

Independent of use, point data need to be translated into regular grids (raster maps) to generate yield maps for farmers, especially where point maps are irregular and gaps in yield data exist.

Table 1.1. Studies that used rasterized yield maps, based on yield monitor data, for various row crops toward analysis of spatial variability of crop yield.

Citation	Methodologies	Usage	Other data layers	Data
Basso et al. 2007	Kriging with exponential isotropic co-variance function	Delineating management zones	None	4 site-years of corn grain, soybean and wheat (8 ha) from Italy
Blackmore, 2000	Simple averaging	Delineating management zones	None	6 site-years of wheat (6.7 ha) from the United Kingdom
Brock et al. 2005	Inverse distance weighting	Delineating management zones	None	24 site-years of corn and soybean from (45 ha) from the United States (US)
Buttafuoco et al. 2017	Ordinary kriging	Delineating management zones	Soil characteristics	3 site-years of durum wheat (12 ha) from Italy
Diker et al. 2004	Inverse distance weighting (12 nearest points)	Delineating management zones	None	6 site-years of commercial corn grain (123.4 ha) from US
Kharel et al. 2019a	Inverse distance weighting	Delineating management zones	None	847 site-years of corn grain and silage (9084 ha) from US
Hornung et al. 2006	Median polish kriging	Delineating management zones	Soil aerial imagery and field topology	3 site-years of corn (183 ha) from US
Khosla et al. 2008	Ordinary kriging	Delineating management zones	Soil topology	15 site-years of corn grain from US
Cox and Gerrard, 2007	Nearest neighbor	Understanding interaction between yield and soil	None	12 site-years of soybean (39.4 ha) from US
Anderson-Cook et al. 2002	Nearest neighbor	Understanding interaction between yield and soil	Electromagnetic conductivity (EC) maps	2 site-year of corn grain, barley, wheat and soybean (24 ha) from US

Citation	Methodologies	Usage	Other data layers	Data
Kitchen et al. 1999	Ordinary kriging	Understanding interactions between yield, soil and landscape	Electromagnetic conductivity (EC) maps	5 site-years of corn grain, 7 site-years of soybean and 1 site-year of grain sorghum (90 ha) from US
Maestrini and Basso, 2018a	Kriging with spherical isotropic co-variance function	Understanding yield variation	Red band spectral reflectance, NDVI and surface temperature	1625 site-years of corn grain, wheat, soybean and cotton
Yang et al. 2001	Inverse distance weighting	Understanding yield and plant growth variation	Airborne digital imagery	1 site-year of sorghum (17 ha) from US
Maestrini and Basso, 2018b	Not mentioned	Understanding interactions between yield and climate, soil, topography and management	Publicly available data on topography, rain and soil information	1625 site-years of corn grain, soybean, wheat and cotton from US

The three most common approaches to developing raster maps from point data include nearest neighbor (NN), inverse distance weighting (IDW) with varying number of nearest points and kriging (Table 1.1; Ross et al. 2008). As the name suggests, NN uses the yield value of the nearest observation to estimate yield while IDW uses a weighted average of nearest neighbors, with weights proportional to the inverse distance. Assuming that there are n set of co-ordinates, $z_1, z_2, \dots, z_n$ and their yield values, denoted as $Y(z_i)$ for $i \in \{1, \dots, n\}$, at those co-ordinates, to estimate yield at co-ordinates x where the yield value is not known, the estimated yield value at location x, denoted as $\bar{Y}(x)$ can be calculated as follows:

$$\bar{Y}(x) = \frac{\sum_{i=1}^{n} \frac{Y(z_i)}{d(z_i, x)^p}}{\sum_{i=1}^{n} \frac{1}{d(z_i, x)^p}} \tag{1.1}$$

where $d(z_i, x)$ represents distance between co-ordinates z_i and x and p is some natural number. In this case, p was set to 1. By weighting sample observations by the inverse of distance,

observations that are closer to the estimated location will have higher weights than the observations that are further away. Kriging, also known as Gaussian Process (GP) regression, models spatial correlation between sample points. Spatial correlation can be modeled using various covariance functions. The Matérn and exponential functions, two commonly used covariance functions in spatial analysis, were used. Isotropy, uniformity of variance in all directions, was also assumed. The exponential covariance function is parameterized as:

$$M\left(z_i, z_j\right) = \sigma^2 \exp\left(-\frac{\left\|z_i - z_j\right\|}{\alpha}\right) \tag{1.2}$$

The Matérn covariance function is parameterized as:

$$M\left(z_i, z_j\right) = \frac{\sigma^2 2^{1-v}}{\Gamma(v)}\left(\frac{\left\|z_i - z_j\right\|}{\alpha}\right)^v K_v\left(\frac{\left\|z_i - z_j\right\|}{\alpha}\right) \tag{1.3}$$

Covariance parameters are variance, σ^2, range, α, smoothness, v and nugget, τ^2, for two GNSS co-ordinates z_i and z_j. The nugget value $\sigma^2\tau^2$ is added to the diagonal of the covariance matrix. Γ is a gamma function and K_v is the modified Bessel function of the second kind. Unlike kriging, which incurs expensive computation time, NN and IDW require computation of distances between sample points only, resulting in reduced computational complexity compared to kriging. However, both NN and IDW fail to account for complex spatial correlation within a field. Nearest neighbor becomes especially inadequate when there is high noise in the data (Wettschereck 1994). While kriging is able to account for a complex correlation structure in the data, it incurs expensive computation time and is therefore less effective when only a weak spatial dependence is present in the data.

Numerous articles have been published comparing the performance of NN, IDW and kriging on a variety of data types. For example, Philips et al. (1997) and Grim & Lynch (1991) both used atmospheric data to quantify ozone exposure on forests and estimate wet deposition in

the atmosphere, respectively. Berman et al. (2015) also evaluated the performances of kriging and inverse distance weighting on interpolating ozone concentrations. Various spatial interpolation methods also were compared using soil information data, such as clay content, soil organic carbon or pH (Bhunia et al. 2016; Bregt 1992; Brus et al. 1996; Declercq 1996; Gallichand & Marcotte 1993; Laslett et al. 1987; Laslett & McBratney 1990; Van Meirvenne et al. 1994). Studies by Heine (1986), Laslett (1994), Rouhani (1986) and Weber & Englund (1994) used (water) elevation data, while Kitanidis & Shen (1996) used chemical data such as, trichloroethylene concentration, to extrapolate spatially limited contaminant concentration information at a hazardous waste site into maps. Out of aforementioned sixteen studies, nine studies (Berman et al. 2015; Bhunia et al. 2016; Grim & Lynch 1991; Heine 1986; Laslett 1994; Laslett & McBratney 1990; Laslett et al. 1987; Philips et al. 1997; Rouhani 1986) compared the performance of kriging and IDW and concluded that kriging is the better methodology. Declercq (1996) showed IDW to be superior to kriging and five studies (Bregt 1992, Brus et al. 1996; Gallichand & Marcotte 1993; Weber & Englund 1994; Van Meirvenne et al. 1994) showed little difference in performance between kriging and IDW.

While there are numerous studies on comparison of spatial estimation methods in other research disciplines, there have been only a few studies comparing different spatial estimation methods for creating a regularized crop yield map based on yield monitor data. No method is uniformly superior on all data types and it therefore is important to systematically compare methods on grain and silage data. Dobermann & Ping (2004) used corn grain and soybean (*Glycine max*. (L.) Merr.) yield data, along with vegetation indices, to analyze the effectiveness of various kriging methods. Evaluated methods included ordinary kriging, co-kriging and kriging with external drift. The study concluded that ordinary kriging led to the lowest error (Dobermann

& Ping 2004). Bazzi et al. (2015) derived profit maps from yield and economic data, such as sales price and production cost, for corn grain and soybean. Their analysis suggested that the impact of spatial estimation method (kriging versus IDW and IDW squared) on profit maps was less than US$30 ha^{-1}, considered insignificant in their study (Bazzi et al. 2015). Souza et al. (2016) concluded that corn grain and soybean yield data lacked spatial structure and, hence, kriging did not outperform IDW. It is important to note, however, that all three studies had limited datasets. All three studies focused on corn grain and/or soybean data. Dobermann & Ping (2004) used data from just two site-years. Bazzi et al. (2015) and Souza et al. (2016) used data from four site-years. None of the studies used kriging with advanced covariance functions, such as exponential and Matérn covariance functions, which is expected to produce an improved raster map. In addition, studies on forages such as corn silage are lacking.

The objective was to evaluate seven widely used spatial estimation methods in creating a rasterized corn silage or grain yield map to determine the most accurate spatial estimation method that captures intra-field spatial variability of yield for both corn silage and corn grain in the state of New York. Evaluation was done using corn silage yield monitor data from 7,484 ha (245 fields) and corn grain yield data from 6,971 ha (253 fields). The seven methods are: NN, IDW using 10, 20, 30 and all data points (IDW 10, 20, 30 and All, respectively) and kriging with exponential (Exponential) and Matérn covariance function (Matérn). The hypothesis is that of the seven methods evaluated, kriging with Matérn covariance function results in the smallest percent error regardless of field size, year when data are obtained or source of the data (farm) for both silage and grain data.

MATERIALS AND METHODS

Yield Monitor Datasets

Yield monitor data were collected from 1,318 site-years from four farms, two of which were dairy farms (hereafter referred to as Silage A and B) and two were cash grain operations (hereafter referred to as Grain A and B). Silage A and B supplied ten and nine years of data, respectively, for a total area of 7,484 ha. Grain A and B supplied seven and eight years of data, respectively, for a total of 6,971 ha. Data reflected the large variability in both yield and field size within farms in New York (Table 1.2; Figure 1.1).

Table 1.2: Summary of farm information illustrating years of record, number of fields, yield statistics, field size statistics, equipment information, location and soil type.

	Unit	Dairy farm A (Silage A)	Dairy farm B (Silage B)	Grain operation A (Grain A)	Grain operation B (Grain B)
Years of record	Years	10	9	7	8
Number of fields		155	90	163	90
Field	ha				
Average field size		13.6	9.9	10.3	9.2
Smallest field		1.5	0.9	0.3	1.1
Largest field		109.5	60.7	53.5	30.7
Total area analyzed		5192.2	2291.7	3565.0	3406.2
Yield	Mg ha^{-1}				
Average yield†		49.5	46.4	10.3	11.7
Lowest yielding field		26.1	4.7	2.4	6.3
Highest yielding field		81.9	72.0	15.3	15.4
Average spatial stdev		8.4	8.0	2.6	2.2
Equipment					
Yield monitor		John Deere Greenstar 3	John Deere Greenstar 3	John Deere Greenstar 3	John Deere Greenstar 3
Recording interval	second	1	1	1	1
Harvester width‡	rows	10, 12	10	8, 12	8, 12
Location		Central New York	Western New York	Central New York	Central New York

	Unit	Dairy farm A (Silage A)	Dairy farm B (Silage B)	Grain operation A (Grain A)	Grain operation B (Grain B)
Soil type					
Most common		Honeoye (Fine-loamy, mixed, semiactive, mesic Glossic Hapludalfs)	Erie (Fine-loamy, mixed, active, mesic Aeric Fragiaquepts)	Schoharie (Fine, illitic, mesic Oxyaquic Hapludalfs)	Ontario (Fine-loamy, mixed, active, mesic Glossic Hapludalfs)
Second most common		Lima (Fine-loamy, mixed, semiactive, mesic Oxyaquic Hapludalfs)	Langford (Fine-loamy, mixed, active, mesic Typic Fragiudepts)	Dunkirk (Fine-silty, mixed, active, mesic Glossic Hapludalfs)	Hilton (Fine-loamy, mixed, active, mesic Oxyaquic Hapludalfs)

† Area weighted average yield
‡ Rows were 0.76 m apart

Postharvest Yield Data Cleaning

Yield monitor data need to be cleaned before analysis because of the presence of systematic and random errors in the data. (Dobermann & Ping 2004; Vega et al. 2019). The raw yield monitor data were read in SMS Advanced software (Ag Leader Technology, Ames, IA, USA), exported in AgLeader format, and imported into and cleaned with Yield Editor (Sudduth et al. 2012; Sudduth & Drummond 2007) using a standardized post-harvest data cleaning protocol (Kharel et al. 2018). This data cleaning protocol addresses issues related to pass overlap (driving over areas already harvested) and yield extremes and applies sensor delays (flow delay and moisture delay) to match the position of sensors with the harvester location based on the flow or moisture pattern within the field and start- and end pass delays to eliminate inaccurate readings when the harvester is speeding up or slowing down. With the data cleaning protocol, 19, 24, 21 and 21% of data were removed for Silage A, B, Grain A and B, respectively. These values are consistent with Blackmore (1999) who removed 32%, Vega et al. (2019) who removed 30% and Thylén et al. (2000) who removed 10-50% of the erroneous yield monitor data.

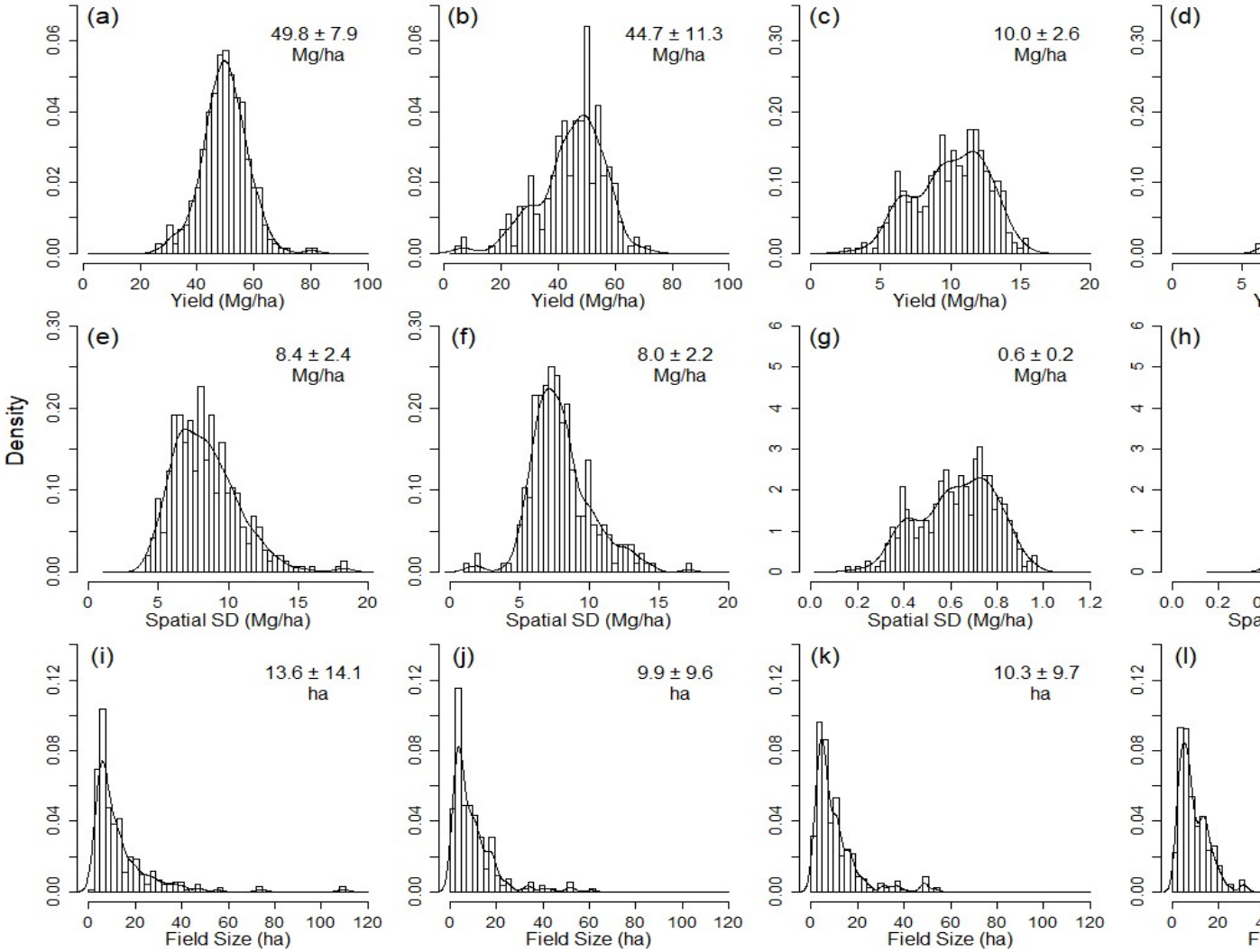

Figure 1.1: Average yield per field (a, b, c, d), spatial standard deviation of yield (e, f, g, h) and field size (i, j, k, distributions for corn silage (a, b, e, f, i, j) and corn grain (c, d, g, h, k, l) post data cleaning protocol. Results are dairy farms with silage data (Silage A and B) and two cash grain operations with corn grain data (Grain A and B the left to right

Implementation of Spatial Estimation Methods

The seven spatial estimation methods explored in this paper include NN, IDW with 10 (IDW 10), 20 (IDW 20), 30 (IDW 30) and all data points (IDW All), kriging with an exponential isotropic covariance function (Exponential) and kriging with the Matérn isotropic covariance function (Matérn), reflecting common methods used in other studies. The data were split into a training (80% of the data) and cross-validation datasets (remaining 20% of the data). Data analyses were performed with R (R Core Team 2019). Gstat package (Pebesma 2004) was used to implement NN and IDW. The GpGp package (Guinness & Katzfuss 2019) was used to implement kriging in order to reduce processing time, given the large number of data points. One of the difficulties in implementing kriging is the computation time. As Katzfuss & Guinness (2019) suggest, kriging becomes infeasible as the size of the dataset becomes larger. This is because kriging requires computation of multivariate normal distributions which incurs quadratic memory and cubic time complexity in the number of observations. In the dataset, fields averaged 6,358 data points, with a maximum of 59,971 data points per field. Implementation of kriging through the "gstat" package, therefore, was not feasible. Parameters for kriging were estimated through maximum likelihood estimation. Unlike the gstat package, the GpGp package uses a generalization of the Vecchia (1988) approach as a framework for Gaussian Process (GP) approximation, which enables fast evaluation of likelihood function resulting in shorter overall computation time.

Spatial Estimation Methods Evaluation

Cross validation was performed to evaluate the performance of each spatial estimation method. For each field, 80% of the data were randomly selected for training. The training dataset

was used to generate rasterized yield maps at 2 × 2 m spatial resolution using the various spatial estimation methods. Predictions were then compared against the validation data. Two evaluation schemes were explored: point-based and area-based. In the point-based approach, the actual yield value from the validation set was compared to the yield from the predicted rasterized yield map at the given GNSS co-ordinate. While point-based evaluation is a natural approach for validating point estimates, the approach fails to acknowledge that yield monitor data represent an average yield density over an area, instead of a point estimate at a given co-ordinate. Though the yield monitor system provides a yield estimate at a certain GNSS co-ordinate, the estimate does not represent a yield value at that specific location, but rather represent the average yield density over the distance traveled from the previous GNSS co-ordinate times the width of the harvest equipment. To correctly ascribe a yield estimate to an area, polygons were generated based on the equipment width, (swath) and distance traveled, as provided by the yield monitor. In the area-based evaluation, the actual yield value from the validation set was then compared against the average yield estimates of all 2 × 2 m pixels inside a given polygon. By accounting for the fact that a point estimate from the yield monitor represents an average yield density over a certain area, the goal was to represent yield monitor data more accurately. However, this approach was computationally expensive and more time consuming than the point-based approach. Both approaches were evaluated to determine if point-based evaluation is an appropriate approximation of area-based evaluation. Normalized root mean squared error (NRMSE) was used to evaluate the performance of each model per field. Assuming that there are m set of co-ordinates, denoted as $x_1, x_2, \dots, x_m$, in the validation dataset of a particular field:

$$\text{NRMSE} = \frac{\sqrt{\frac{1}{m} \sum_{j=1}^{m} \left(Y(x_j) - \bar{Y}(x_j)\right)^2}}{\frac{1}{m} \sum_{j=1}^{m} Y(x_j)} \times 100 \tag{1.4}$$

where $Y(x_j)$ represents the actual yield level at co-ordinate x_j and $\bar{Y}(x_j)$ represents the predicted yield level at co-ordinate x_j based on one of the methods. Because residuals, $Y(x_j) - \bar{Y}(x_j)$ for $i \in \{1,2,\dots,m\}$, are usually proportional to the yield level of that field, RMSE from a high yielding field will generally be larger than that of a lower yielding field, thus putting more weight on errors from high yielding fields. By normalizing RMSE with the average yield of the field, NRMSE provides a dimensionless measurement of error per field.

Empirical Average Analysis

Area-based evaluation resulted, on average, in a higher NRMSE than point-based evaluation. While the assumption that each observation of yield monitor data represents yield over an area rather than at a specific location is a sound data assumption, the seven spatial estimation methods explored in this paper all assume the data to be point estimates, rather than estimates over an area. This discrepancy contributed to over-estimation of error by the area-based evaluation. Thus, the empirical average analysis was performed based on point-based evaluation results only. The empirical averages of NRMSE of each spatial methods by data type (Silage A, Silage B, Grain A and Grain B), field size (up to 140 ha) and year (2009-2018) were analyzed and compared.

The analysis suggested that the average NRMSE of Grain A data was much larger than that of other three farms. Average coefficient of variation (CV) was calculated for each farm to compare variation of yield level per farm. Suppose that there are n set of co-ordinates, $z_1, z_2, \dots, z_n$ and their yield values in a single field. Coefficient of variation (CV) is defined as:

$$\mathrm{CV} = \frac{\sqrt{\frac{1}{n}\ \sum_{i=1}^{n}\left(Y(z_i) - \hat{Y}\right)^2}}{\hat{Y}} \tag{1.5}$$

where $\hat{Y}$ is the average yield level of the field. Average CV was derived by taking the arithmetic mean of CV of all fields within the farm. The relative performance of spatial methods were mostly consistent across data type, field size and year in which data were collected, except for the nearest neighbor method. The nearest neighbor method was further analyzed to account for such volatility in its performance. Coefficient of variation of yield, log of field size and year were analyzed to test the relationship between yield variability within a field and NRMSE using linear regression.

Mixed Model Analysis

For reasons explained above, the mixed model analysis was based on point-based evaluation results only. After analyzing behaviors of empirical averages of NRMSE, a mixed model was fitted using the "lme4" package in R (Bates et al. 2015) to compare differences in NRMSE for each spatial estimation methods and to test if they are statistically significant. The following R command was used to fit the linear mixed model:

$$\mathrm{lmer}\big(\mathrm{NRMSE} \sim \mathrm{Method} * \log(\mathrm{Area}) + \mathrm{Farm} + \mathrm{Year} + (1|\mathrm{Field})\big) \qquad (1.6)$$

where lmer refers to a R command to fit a linear mixed effect model in R; Method refers to the seven spatial estimation methods (Fixed effect); Area is the size of the field (Fixed effect); Farm refers to Grain A, Grain B, Silage A, Silage B (Fixed effect); Year reflects the year in which the harvest was done (Fixed effect); and Field reflects the unique combination of farm, fieldname and year of harvest (Random effect). In this model, "Area" was log transformed to normalize the data, as they were distinctly right skewed (as evident in Figure 1.1). In addition to additive effects from "spatial estimation methods", "farm", "year" and "log(area)", multiplicative effects between "spatial estimation methods" and "log(area)" were introduced, because the effect of

"log(area)" on NRMSE varied significantly depending on "spatial estimation methods". Marginal means were estimated for each spatial estimation method using the "lsmeans" package in R (Lenth 2016). Marginal means were estimated by adding average fixed effects over 10 years, 4 farms and average field size of 11 ha to the intercept for each spatial estimation method. Tukey comparisons were then performed between spatial estimation methods to elucidate the statistical difference in model performance.

RESULTS AND DISCUSSION

Area- versus Point-based Evaluation

Across IDW- and kriging-based spatial estimation methods and all fields and farms, area-based evaluation consistently led to a slightly higher average NRMSE, averaging 8.6 across these methods versus 7.9 for the point-based evaluation (Table 1.3).

Table 1.3: Comparison of average Normalized Root Mean Squared Errors (NRMSE) by farm (Silage A, Silage B, Grain A and Grain B) and evaluation methods (area-based and point-based) for seven spatial estimation methods. Estimated marginal means were generated (with point-based evaluation only) for comparisons between spatial estimation methods.

Spatial estimation methods	Silage A		Silage B		Grain A		Grain B	
	Area	Point	Area	Point	Area	Point	Area	Point
Nearest neighbor (NN)	6.94	6.89	7.55	7.76	9.73	10.74	5.84	6.72
Inverse distance weighting (IDW)								
10 nearest points (IDW 10)	7.34	6.37	7.59	6.97	10.40	9.94	6.65	6.41
20 nearest points (IDW 20)	7.94	6.78	8.19	7.22	11.18	10.46	7.25	6.86
30 nearest points (IDW 30)	8.26	7.02	8.52	7.39	11.61	10.78	7.57	7.12
All data points (IDW All)	10.60	9.13	11.20	9.26	15.20	14.05	10.06	9.47
Kriging								
Exponential isotropic (Exponential)	6.49	5.63	7.22	6.59	9.70	9.29	5.82	5.49
Matérn isotropic (Matérn)	6.10	5.42	7.14	6.54	9.53	9.19	5.31	5.14

The most likely reason for larger NRMSE in area-based evaluation in these estimation models is that none of these spatial estimation methods account for the fact that each observation

from a yield monitor system is an estimate for a certain area (product of harvester width and distance traveled per second) and not an actual point estimate with specific GNSS units.

Under both evaluation methods, Matérn consistently showed the lowest average NRMSE among all seven spatial methods, with 7.1 error under area-based evaluation and 6.6 error under point-based evaluation. Exponential resulted in the second lowest average NRMSE, followed by IDW 10, IDW 20, IDW 30 and IDW All. The estimation accuracy of IDW deteriorated with an increasing number of data points, which is plausibly due to increased smoothing as data points farther from the estimation location are captured. Performance of NN, on the other hand, varied by evaluation method. Under area-based evaluation, NN was the third best performing method behind Matérn and Exponential. Under point-based evaluation, however, NN was the fifth best method for Silage A and Grain A data, fourth best for Grain B data and sixth best on Silage B data. Given that both area- and point-based evaluations yielded the same results, with a slightly lower NRMSE for point-based evaluation for most comparisons, additional analyses on the impact of field size, year and farm specificity on model performance were performed using the point-based evaluation method for both corn silage and corn grain data.

Performance of Spatial Estimation Methods by Farm, Field, Year and Field Size

The empirical average NRMSE per farm ranged from 5-9 for Silage A, B and Grain B, and from 9-14 for Grain A (Table 1.3). This could be explained by the higher variation of yield on average for Grain A data. Coefficient of variation of yield, a measure of variation in yield per field, was calculated for each farm. Grain A had higher average CV of 27% whereas Silage A, Silage B and Grain B had 17, 19 and 20% respectively, suggesting higher level of yield variation for Grain A data, which could result in higher NRMSE across all seven methods. However,

despite an overall higher NRMSE for Grain A, the evaluation of the seven spatial estimation methods on this farm still resulted in the same ranking of methods: Matérn was the best performing method with the lowest average NRMSE, followed by Exponential, IDW 10, IDW 20, IDW 30 and IDW All. The NN results were inconsistent; it was the 2nd lowest method behind IDW All for Grain B, the 3rd lowest method behind IDW All and IDW 30 for Silage A and Grain A, and 4th lowest behind IDW All, IDW 30 and IDW 20 for Silage B.

At the individual field level, the NRMSE from Matérn was also consistently lower than that of other spatial estimation methods (Figure 1.2) as most observations, which represent NRMSE from Matérn on the y-axis and NRMSE from other models on the x-axis, on the plot are on the right hand side of the one-to-one line. Thus, not only across farms but also across individual fields, Matérn was the best performing method.

The average NRMSE ranged between 6.1 - 8.9, year-to-year. Despite the difference in NRMSE year-to-year, the relative performance of the spatial estimation methods, except NN, were consistent across years; Matérn always resulted in the lowest NRMSE, followed by Exponential, IDW 10, IDW 20, IDW 30 and IDW All (Figure 1.3). The results of the NN method showed inconsistency in ranking from year-to-year; while in most years (2011, 2014, 2016, 2017 and 2018), the NN method was one of the lowest performing methods, next to IDW All. In 2009, it was the third best method behind Matérn and Exponential. However, for every year of data and for both crop types, Matérn outperformed all other methods.

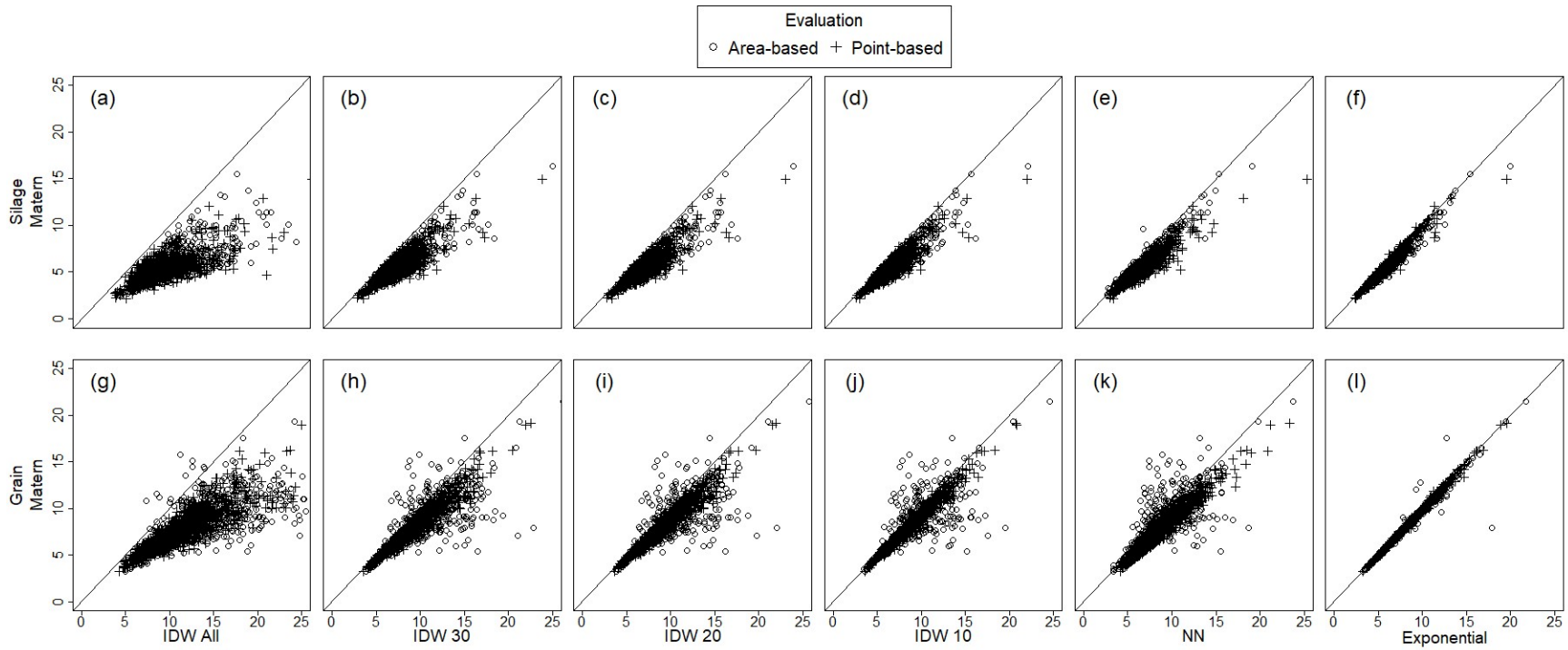

Figure 1.2: Comparison of normalized root mean squared error (NRMSE) per field. Each point on a plot represents NRMSE of a field for corn silage (a, b, c, d, e, f) and corn grain (g, h, i, j, k, l). Each dot represents NRMSE from area-based evaluation and a cross represents NRMSE from point-based evaluation. Spatial estimation methods included nearest neighbor (NN), inverse distance weighting (IDW) with varying number of nearest points (10, 20, 30, all) and kriging with exponential (Exponential) or Matérn (Matérn) covariance functions

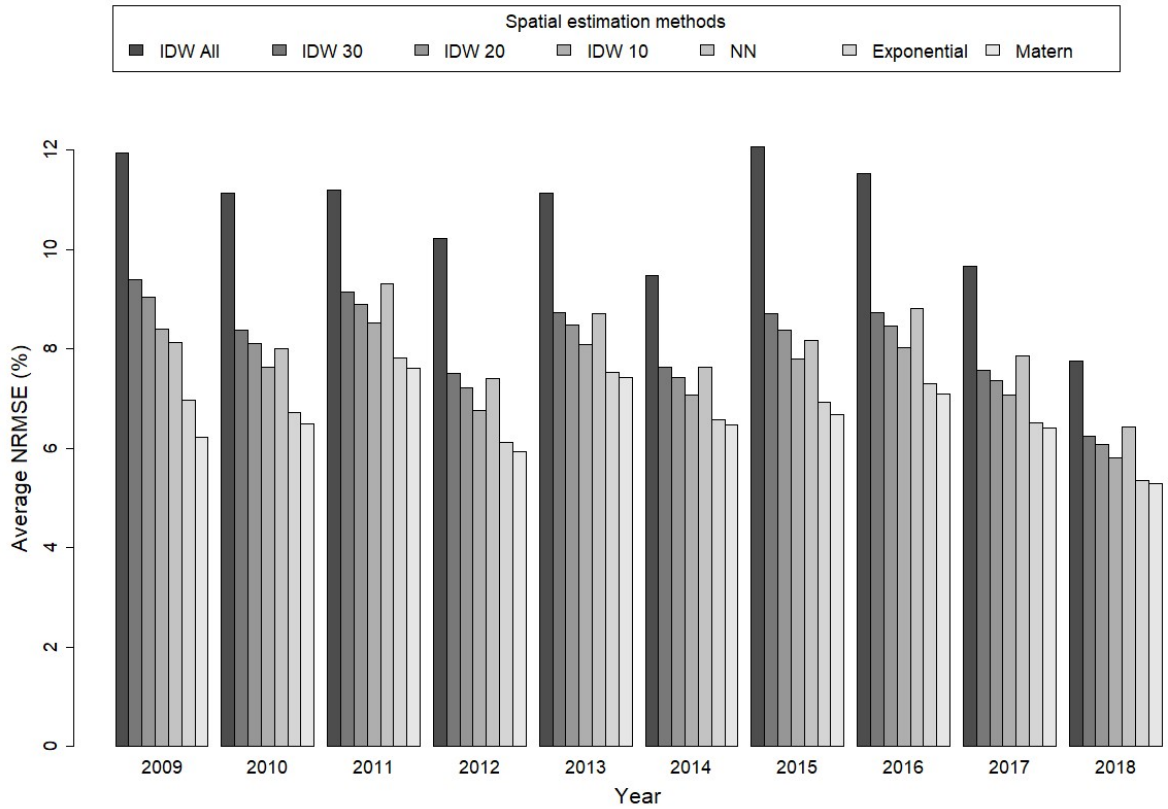

Figure 1.3: Comparison of average normalized root mean squared error from year 2009-2018 for each spatial estimation method, including nearest neighbor (NN), inverse distance weighting (IDW) with varying number of nearest points and kriging with exponential (Exponential) or Matérn (Matérn) covariance functions.

In general, all seven spatial estimation methods performed better as the size of the field increased (Figure 1.4). The degree to which NRMSE decreases as the size of the field increases differed among methods; IDW had the least steep slope of -0.64, while NN had the steepest slope of -0.99. Despite this difference in slope among spatial estimation methods, Matérn resulted in the lowest average NRMSE across fields, followed by Exponential, IDW 10, IDW 20, IDW 30 and IDW All.

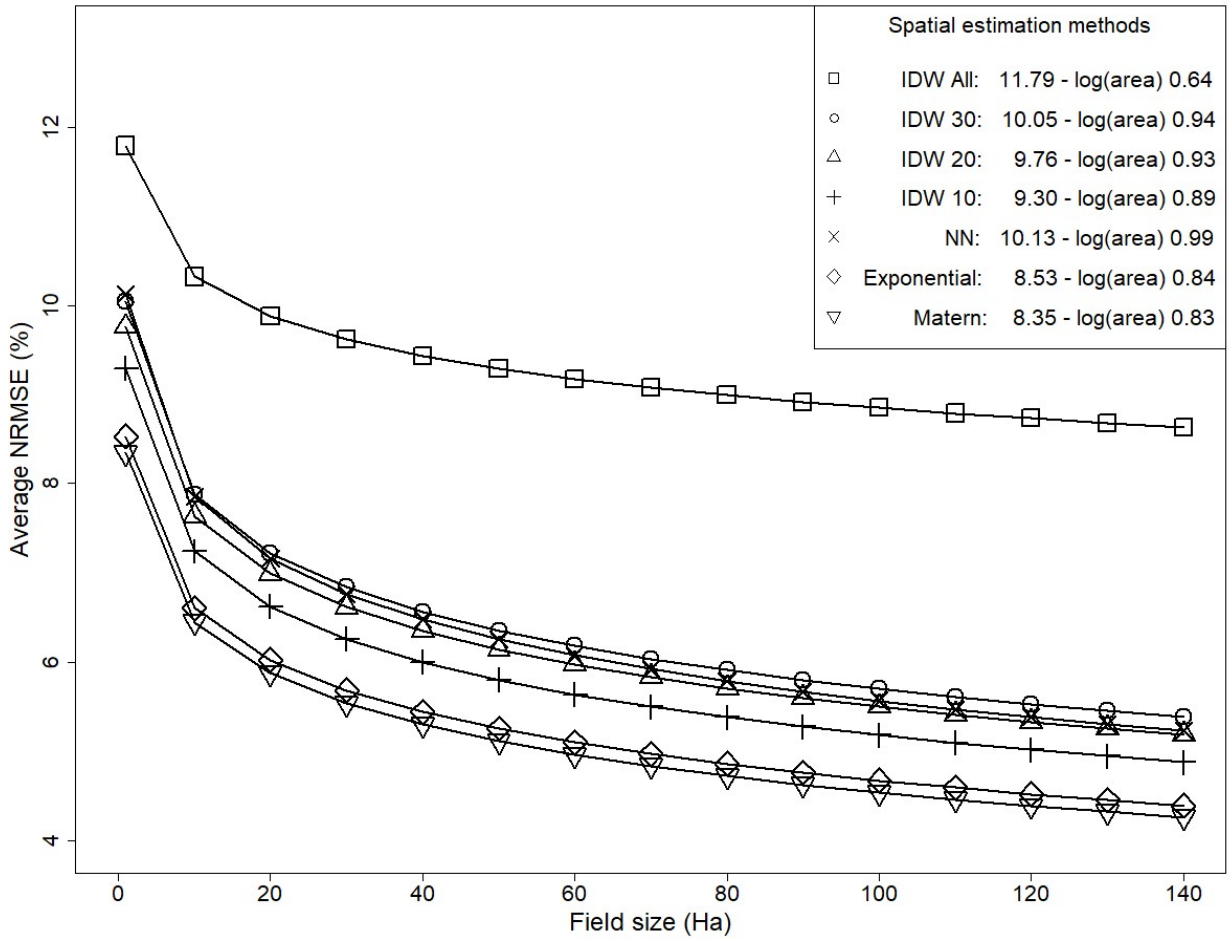

Figure 1.4. Comparison of average normalized root mean squared error (NRMSE) by corn field size, ranging up to 140 ha for each spatial estimation method, including nearest neighbor (NN), inverse distance weighting (IDW) with varying number of nearest points and kriging with exponential (Exponential) or Matérn (Matérn) covariance functions. A regular linear regression model with NRMSE as the dependent and log of area in hectare as the independent variable was fitted for each spatial estimation method to analyze the performances conditioned on field sizes.

The analysis suggests that the performance of the NN varies by field. Wettschereck (1994) analyzed behavior of the k-nearest neighbor algorithm (1, 2… k) on various data containing noisy instances and discovered that the performance of the NN algorithm depended on number, noisiness and sparseness of the instances. He showed that NN performed poorly especially on larger dataset (>100 data points), while performing better on smaller (<100 data points) or sparsely distributed dataset. This is in line with the observation for the NN algorithm in the analysis. Sparseness, number and noisiness of instances varied greatly by field, causing the

performance of the NN method to also vary. While the performance of NN varied by field, for point-based evaluation, Matérn resulted in the lowest NRMSE for all but one of the 1,318 fields (Figure 1.2).

The strong effects of field size and year on the NRMSE were attributed to the yield variation within a field. The regression of CV in yield of each field and log of field size suggested that as the field size increased the CV decreased (Table 1.4). The regression of CV of yield and year indicated that the data from 2018 had the lowest variation in yield (Table 1.4). The correlation between CV and NRMSE showed a linear relationship between NRMSE and CV across fields (Figure 1.5), supporting the hypothesis that variation in yield affects the accuracy of all methods.

Table 1.4: Summary of linear model based on two independent variables. Year implies the year when the data were collected and log(field size) implies a log-transformed field size. The dependent variable was the coefficient of variation of yield on each field (CV), which was calculated by averaging the standard deviation of yield by the average yield of the field. The summary output shows beta estimates, standard errors, T-statistics and P-Values.

Term	Estimate	Standard Error	T-statistics	P-Value
(Intercept)	0.261	0.011	22.969	<0.001
log(field size)	-0.023	0.003	-8.298	<0.001
Year2009	0	-	-	-
Year2010	-0.039	0.011	-3.482	0.001
Year2011	0.028	0.011	2.442	0.015
Year2012	-0.001	0.012	-0.116	0.908
Year2013	0.010	0.012	0.808	0.419
Year2014	0.027	0.012	2.300	0.022
Year2015	-0.017	0.012	-1.509	0.132
Year2016	-0.006	0.032	-0.204	0.838
Year2017	0.016	0.012	1.264	0.207
Year2018	-0.056	0.012	-4.609	<0.001

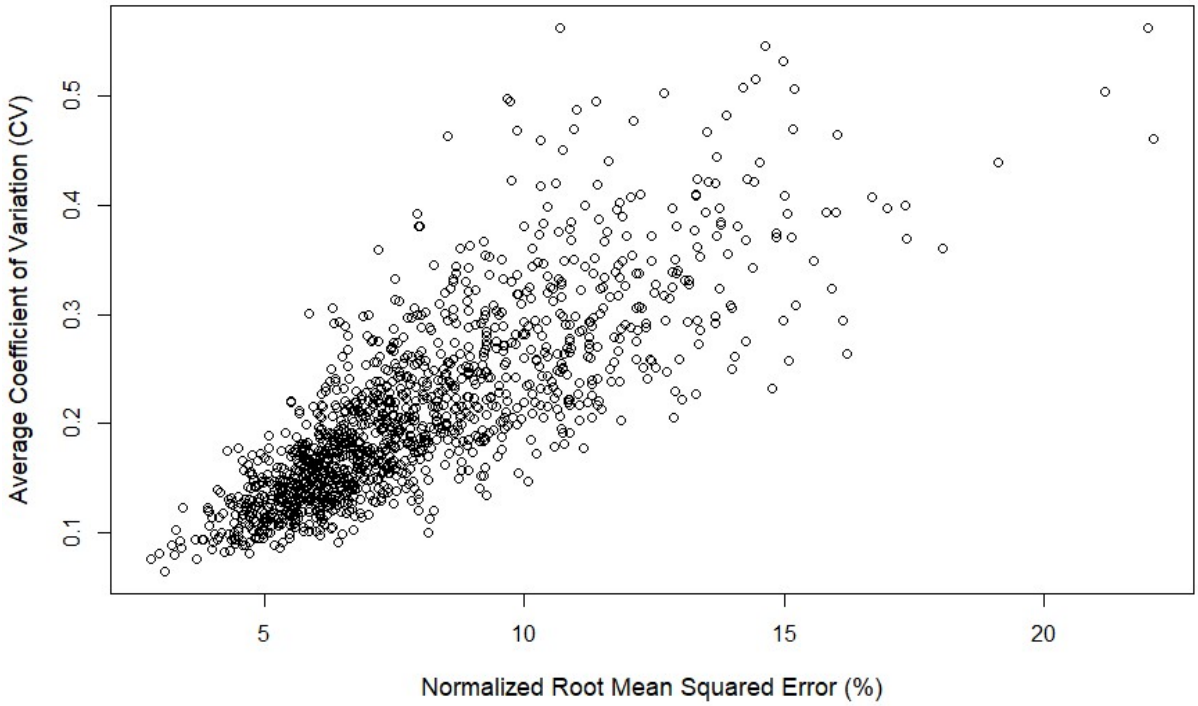

Figure 1.5: Comparison of normalized root mean squared error (NRMSE) across seven spatial estimation methods and average coefficient of variation in yield for 1,318 corn fields.

Mixed Model Results

Consistent with the observations of empirical averages, T-statistics and p-values from the mixed model indicated statistical significance of all the beta estimates in the model (Table 1.5). Kriging with the Matérn isotropic covariance function (Matérn) resulted in the lowest estimated marginal means of 6.6 error, followed by Exponential with 6.7 error, IDW 10 with 7.3 error, IDW 20 with 7.7 error, NN with 7.9 error, IDW 30 with 8.0 error and IDW All with 10.4 error (Table 1.6). In all pairwise comparisons between Matérn and six other spatial methods, the difference in NRMSE was statistically significant ($p < 0.0001$) (Table 1.7).

Table 1.5: Linear mixed effect model summary for 7 – 10 years of data per farm (7,484 ha [245 fields] of silage and 6,971 ha [253 fields] of grain) from four farms (Grain A and B, Silage A and B), where Method refers to seven spatial estimation methods (Fixed effect); log(Area) implies a log-transformed field size (Fixed effect); Farm refers to Grain A, B, Silage A, B (Fixed effect); Year reflects year of harvest (Fixed effect); and Field reflects the unique combination of farm, fieldname and harvest year (Random effect). The summary output shows beta coefficients, standard errors, degrees of freedom, t and p values for the fixed effects, as well as the residual and group variances of the random effects.

	Estimate	Standard Error	Degrees of Freedom	T value	P value
[Intercept]	13.05	0.79	1333	16.55	<0.001
Models [IDW All]	0.00	-	-	-	-
Models [IDW 30]	-1.46	0.13	7896	-11.69	<0.001
Models [IDW 20]	-1.77	0.13	7896	-14.09	<0.001
Models [IDW 10]	-2.26	0.13	7896	-18.01	<0.001
Models [NN]	-1.34	0.13	7896	-10.69	<0.001
Models [Exponential]	-3.07	0.13	7896	-24.53	<0.001
Models [Matérn]	-3.26	0.13	7896	-26.05	<0.001
log(Area)	-0.57	0.08	1702	-7.55	<0.001
Farm [Silage A]	0.00	-	-	-	-
Farm [Silage B]	-0.35	0.17	1304	-2.05	0.040
Farm [Grain A]	3.63	0.16	1304	22.46	<0.001
Farm [Grain B]	0.41	0.15	1304	2.75	0.006
Year [2009]	0.00	-	-	-	-
Year [2010]	-1.94	0.80	1304	-2.43	0.015
Year [2011]	-0.86	0.78	1304	-1.11	0.266
Year [2012]	-1.67	0.78	1304	-2.14	0.032
Year [2013]	-1.68	0.78	1304	-2.16	0.031
Year [2014]	-2.86	0.77	1304	-3.70	<0.001
Year [2015]	-1.42	0.77	1304	-1.84	0.066
Year [2016]	-1.66	0.77	1304	-2.16	0.031
Year [2017]	-1.81	0.77	1304	-2.35	0.019
Year [2018]	-2.69	0.77	1304	-3.48	0.001
Models [IDW All] : log(Area)	0.00	-	-	-	-
Models [IDW 30] : log(Area)	-0.31	0.04	7896	-7.55	<0.001
Models [IDW 20] : log(Area)	-0.29	0.04	7896	-7.09	<0.001
Models [IDW 10] : log(Area)	-0.26	0.04	7896	-6.33	<0.001
Models [NN] : log(Area)	-0.35	0.04	7896	-8.70	<0.001
Models [Exponential] : log(Area)	-0.20	0.04	7896	-4.94	<0.001
Models [Matérn] : log(Area)	-0.19	0.04	7896	-4.69	<0.001
Random Effects					
Residual (σ^2)	0.69				
Intercept [Field] (τ_{Field})	3.84				
Number of observations for Field	1318				
Total number of observations	9226				

Table 1.6: The least square mean estimates, their respective standard errors and 95% confidence intervals of normalized root mean squared error (NRMSE) for seven spatial estimation methods. The response for the linear mixed model was NRMSE; four farms (Grain A, B, Silage A, B), the seven spatial estimation methods, harvest year (2009 ~ 2018) and the logged transformed size of the field in hectare were treated as additive fixed effects.

Spatial estimation methods	Estimates	Standard Error	95% confidence interval
Nearest neighbor (NN)	7.94	0.10	7.75 - 8.13
Inverse distance weighting (IDW)			
10 nearest points (IDW 10)	7.34	0.10	7.15 - 7.53
20 nearest points (IDW 20)	7.73	0.10	7.54 - 7.92
30 nearest points (IDW 30)	7.97	0.10	7.77 - 8.16
All data points (IDW All)	10.44	0.10	10.25 - 10.64
Kriging			
Exponential isotropic (Exponential)	6.71	0.10	6.52 - 6.90
Matérn isotropic (Matérn)	6.55	0.10	6.36 - 6.74

Table 1.7: Tukey comparison of least square estimates of normalized root mean squared error (NRMSE) between kriging with Matérn isotropic covariance function (Matérn) and six other methods, including inverse distance weighting (IDW) with varying number of points, nearest neighbor (NN) and kriging with exponential isotropic covariance function (Exponential).

Contrast	Estimate	Standard Error	Z.ratio	P-value
IDW All - Matérn	3.892	0.035	111.854	<0.0001
IDW30 - Matérn	1.415	0.035	40.679	<0.0001
IDW20 - Matérn	1.176	0.035	33.810	<0.0001
IDW10 - Matérn	0.787	0.035	22.629	<0.0001
NN - Matérn	1.387	0.035	39.860	<0.0001
Exponential - Matérn	0.156	0.035	4.497	<0.0001

On average, NRMSE from Matérn was 17% lower than NRMSEs of the other six spatial estimation methods. The difference was the largest with 37% decrease when compared against IDW All and the smallest when compared against Exponential with just 2% decrease. These results suggest kriging to be the most consistent spatial estimation method in mapping yield monitor data for corn grain and silage across a large range of field sizes and yield levels. Contrary to the finding by Souza et al. (2016) that yield monitor data lack spatial structure, the result suggests that spatial information can be used to better estimate yield at the field- and within-field levels for both corn grain and silage. The results also contradict Bazzi et al. (2015),

who stated that the spatial estimation method was of peripheral importance in generating a yield map. The contradictory results may stem from varying data cleaning protocol, such effects were not tested. The analysis shows that the difference between Matérn, the best performing method, and IDW All, the lowest performing model, averaged 46% per field, suggesting that spatial estimation method is a significant factor when generating a yield map. Both Souza et al. (2016) and Bazzi et al. (2015) included only a limited number of fields and focused on grain crops (soybean and corn). The apparent inconsistency in conclusions between this study and the work by Souza et al. (2016) and Bazzi et al. (2015) may be due to differences in location and the size and source of the data, including crop type.

A rasterized yield map based on yield monitor data often is used, along with other data, such as vegetation indices or electrical conductivity maps, to delineate management zones (Basso et al. 2007; Blackmore 2000; Brock et al. 2005; Diker et al. 2004; Hornung et al. 2006; Kharel et al. 2018; Khosla et al. 2008) or to understand the interaction between yield and other features such as soil, landscape and topography (Anderson-Cook et al. 2002; Cox and Gerrard 2007; Kitchen et al. 1999; Maestrini and Basso 2018b; Yang et al. 2001). Out of the 12 studies listed above, only five studies (Basso et al. 2007; Hornung et al. 2006; Khosla et al. 2008; Kitchen et al. 1999; Maestrini and Basso 2018b) used a form of kriging to generate a rasterized yield map. The analysis suggests that greater attention is required to yield mapping by both researchers and practitioners who aim to use yield data to develop management zones and/or prescription maps, given that the choice of estimation method affects the rasterized yield maps generated from the yield monitor data. Findings in this paper span a large number of fields, variety in field sizes and yield levels, as well as corn harvested for grain and corn grown for

silage and suggest the need for kriging with Matérn isotropic covariance function to account for the spatial structure of yield within fields.

CONCLUSIONS

Out of the seven spatial estimation methods tested, kriging with Matérn isotropic covariance function resulted in the lowest NRMSE across four farms, ten years of silage yield data, nine years of grain yield data and across a wide range of field sizes (1 - 140 ha), reflecting the diversity of fields in corn production in New York. On average, Exponential was the second-best method, followed by IDW 10, IDW 20, NN, IDW 30 and IDW All. These results support the original hypothesis. Kriging with Matérn covariance function is highly recommended to derive single year corn yield raster maps for corn grain and corn silage yield monitor data and development of multi-year yield stability maps that include not only spatial, but also temporal variation in yield.

REFERENCES

Anderson-Cook, C.M., Alley, M.M., Roygard, J.K.F., Khosla, R., Noble, R.B. & Doolittle, J.A. (2002). Differentiating soil types using electromagnetic conductivity and crop yield maps. *Soil Science Society of America Journal, 66*, 1562-1570. https://doi.org/10.2136/sssaj2002.1562.

Arslan, S., & Colvin, T. (2002). An evaluation of the response of yield monitors and combines to varying yields. *Precision Agriculture. 3,* 107-122. https://doi.org/10.1023/A:1013887801918.

Basso, B., Bertocco, M., Sartori, L., & Martin, E. C. (2007). Analyzing the effects of climate variability on spatial pattern of yield in a maize–wheat–soybean rotation. *European Journal of Agronomy, 26*(2), 82-91. https://doi.org/10.1016/j.eja.2006.08.008.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1-48. https://doi.org/10.18637/jss.v067.i01.

Bazzi, Claudio L, Souza, Eduardo G, Khosla, Raj, Uribe-Opazo, Miguel A, & Schenatto, Kelyn. (2015). Profit maps for precisión agriculture. *Ciencia e investigación agraria, 42*(3), 305-330. https://dx.doi.org/10.4067/S0718-16202015000300007.

Blackmore, S., (1999). Remedial correction of yield map data. *Precision Agriculture, 1*, 53–66. https://doi.org/10.1023/A:1009969601387.

Blackmore, S., (2000). The interpretation of trends from multiple yield maps. *Computers and Electronics in Agriculture, 26*(1), 37–51. http://dx.doi.org/10.1016/S0168-1699(99)00075-7.

Bregt, A.K. (1992). *Processing of soil survey data*. Ph.D. thesis, Department of Environmental Sciences, Agricultural University of Wageningen, The Netherlands.

Brock, A., Brouder, S.M., Blumhoff, G., & Hofmann, B.S. (2005). Defining yield-based management zones for corn-soybean rotations. *Agronomy Journal, 97*(4), 1115-1128. https://doi.org/10.2134/agronj2004.0220.

Brus, D.J., DeGruijter, J.J., Marsman, B.A., Visschers, R., Bregt, A.K., & Breeuwsma, A. (1996). The performance of spatial interpolation methods and choropleth maps to estimate properties at points: A soil survey case study. *Environmetrics* 7, 1-16. https://doi.org/10.1002/(sici)1099-095x(199601)7:1<1::aid-env157>3.0.co;2-y.

Cox, M., & Gerard, P. (2007). Soil management zone determination by yield stability analysis and classification. *Agronomy Journal*, *99*, 1357-1365. https://doi.org/10.2134/agronj2007.0041.

Declercq, F.A.N. (1996). Interpolation methods for scattered sample data: Accuracy, spatial patterns, processing time. *Cartography and Geographic Information Systems, 23*(3), 128-144. https://doi.org/10.1559/152304096782438882.

Diker, K., Heermann, D., & Brodahl, M. (2004). Frequency analysis of yield for delineating yield response zones. *Precision Agriculture, 5*. 435–444. https://doi.org/10.1007/s11119-004-5318-9.

Dobermann, A., & Ping, J. (2004). Geostatistical integration of yield monitor data and remote sensing improves yield maps. *Agronomy Journal*, *96*, 285-297. https://doi.org/10.2134/agronj2004.0285.

Gallichand, J., & Marcotte, D. (1993). Mapping clay content for subsurface drainage in the Nile Delta. *Geoderma* 58, 165-179. https://doi.org/10.1016/0016-7061(93)90040-r.

Grim, J.W., & Lynch, J.A. (1991). Statistical analysis of errors in estimating wet deposition using five surface estimation algorithms. *Atmospheric Environment. Part A, 26*(2), 317-327. https://doi.org/10.1016/0960-1686(91)90303-o.

Guinness, J and Katzfuss, M. (2019) *GpGp: Fast Gaussian Process Computation Using Vecchia's Approximation.* R package version 0.2.1 https://CRAN.R-project.org/package=GpGp.

Heine, G.W. (1986). A controlled study of some two-dimensional interpolation methods. *COGS Computer Contributions, 2*(2), 60-72.

Hornung, A., Khosla, R., Reich, R., Inman, D., & Westfall, D. (2006). Comparison of site-specific management zones: Soil-color-based and yield-based. *Agronomy Journal, 98*(2), 405-417. https://doi.org/10.2134/agronj2005.0240.

Katzfuss, M., & Guinness, J. (2019). A general framework for Vecchia Approximations of Gaussian processes. Retrieved June 29, 2020, from https://arxiv.org/abs/1708.06302.

Kharel, T., Swink, S.N., Youngerman, C., Maresma, A., Czymmek, K.J., Ketterings, Q.M., Kyveryga, P., Lory, J., Musket, T.A., & Hubbard, V. (2018). *Processing/cleaning corn silage and grain yield monitor data for standardized yield maps across farms, fields, and years*. Cornell University, Nutrient Management Spear Program, Department of Animal Science, Ithaca NY. http://nmsp.cals.cornell.edu/publications/extension/ProtocolYieldMonitorDataProcessing1_15_2020.pdf.

Kharel, T.P., Marema, A., Czymmek, K.J., Oware, E.K., & Ketterings, Q.M. (2019a). Combining spatial and temporal corn silage yield variability for management zone development. *Agronomy Journal, 111*(6), 2703-2711. https://doi.org/10.2134/agronj2019.02.0079.

Kharel, T.P., Swink, S.N., Maresma, A., Youngerman, C., Kharel, D., Czymmek, K.J., & Ketterings, Q.M. (2019b). Yield monitor data cleaning is essential for accurate corn grain/silage yield determination. *Agronomy Journal, 111*(2), 509-516. https://doi.org/10.2134/agronj2018.05.0317.

Kitanidis, P.K., & Shen, K.F. (1996). Geostatistical interpolation of chemical data. *Advances in Water Resources* 19, 369-378. https://doi.org/10.1016/0309-1708(96)00016-4.

Kitchen, N.R., Sudduth, K.A. & Drummond, S.T. (1999), Soil electrical conductivity as a crop productivity measure for claypan soils. *Journal of Production Agriculture, 12*(4), 607-617. https://doi.org/10.2134/jpa1999.0607.

Khosla, R., Inman, D.J., Westfall, D.G., Reich, R.M., Frasier, M., Mzuku, M., Koch, B., & Hornung, A. (2008). A synthesis of multi-disciplinary research in precision agriculture: site-specific management zones in the semi-arid western Great Plains of the USA. *Precision Agriculture, 9*, 85-100. https://doi.org/10.1007/s11119-008-9057-1.

Laslett, G.M., McBratnety, A.B., Pahl, P.J., & Hutchinson, M.F. (1987). Comparison of several spatial prediction methods for soil pH. *Journal of Soil Science,* 38, 325-341. https://doi.org/10.1111/j.1365-2389.1987.tb02148.x.

Laslett, G.M. (1994). Kriging and splines: An empirical comparison of their predictive performance in some applications. *Journal of the American Statistical Association* 89: 391-409. https://doi.org/10.2307/2290840.

Laslett, G.M., & McBratnety, A.B. (1990). Further comparisons of spatial methods for predicting soil pH. *Soil society of America Journal* 54, 1553-1558. https://doi.org/10.2136/sssaj1990.03615995005400060007x.

Lenth, R. V. (2016). Least-Squares Means: The R Package lsmeans. *Journal of Statistical Software, 69*(1), 1-33. doi:10.18637/jss.v069.i01

Long, E., Ketterings, Q.M., Russell, D., Vermeylen, F., & DeGloria, S.D. (2016). Assessment of yield monitoring equipment for dry matter and yield of corn silage and alfalfa/grass. *Precision Agriculture*, 7, 546-563. https://doi.org/10.1007/s11119-016-9436-y.

Maestrini, B., & Basso, B. (2018a). Drivers of within-field spatial and temporal variability of crop yield across the US Midwest. *Scientific Reports, 8,* 106-112. https://doi.org/10.1038/s41598-018- 32779-3.

Maestrini, B., & Basso, B. (2018b). Predicting spatial patterns of within-field crop yield variability. *Field Crops Research, 219*, 106-112, https://doi.org/10.1016/j.fcr.2018.01.028.

Pebesma, E.J., 2004. Multivariable geostatistics in S: the gstat package. *Computers & Geosciences, 30*, 683-691. https://doi.org/10.1016/j.cageo.2004.03.012.

Philips, D.L., Lee, E.H., Herstrom, A.A., Hogsett, W.E., & Tingey, D.T. (1997). Use of auxiliary data for spatial interpolation of ozone exposure in southeastern forests. *Environmetrics, 8*, 43-61. https://doi.org/10.1002/(sici)1099-095x(199701)8:1<43::aid-env237>3.0.co;2-g.

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Ross, K.,Morris, D., & Johannsen, C.J. (2008). A review of intra-field yield estimation from yield monitor data. *Applied Engineering in Agriculture. 24.* 309-317. https://doi.org/10.13031/2013.24496.

Rouhani, S. (1986). Comparative study of ground-water mapping techniques. *Ground Water 24*(2), 207-216. https://doi.org/10.1111/j.1745-6584.1986.tb00996.x.

E. G. Souza, C. L. Bazzi, R. Khosla, M. A. Uribe-Opazo & Robin M. Reich (2016) Interpolation type and data computation of crop yield maps is important for precision crop production, *Journal of Plant Nutrition, 39*(4), 531-538. https://doi.org/10.1080/01904167.2015.1124893.

Sudduth, K. A., & Drummond, S. T. (2007). Yield Editor. *Agronomy Journal*, *99*(6), 1471–1482. https://doi.org/10.2134/agronj2006.0326.

Sudduth, K. A., Drummond, S. T., & Myers, D. B. (2012). Yield Editor 2.0: Software for Automated Removal of Yield Map Errors. ASABE Paper No. 121338243. St. Joseph, MI: ASABE. https://doi.org/10.13031/2013.41893.

Thylén, L., Algerbo, P.A., Giebel, A., Robert, P.C., Rust, R.H., and Larson, W.E. (2000). An expert filter removing erroneous yield data. *Proceedings of the 5th International Conference on Precision Agriculture and Other Precision Resources Management.* Robert, P.C., R.H. Rust, W.E Larson (eds.). ASA/CSSA/SSSA, Madison, Wisconsin, USA – CDROM.

USDA. (2019a). United States Department of Agriculture - New York State Agriculture Overview. Retrieved from https://www.nass.usda.gov/Quick_Stats/Ag_Overview/stateOverview.php?state=NEW%20YORK.

USDA (2019b). United States Department of Agriculture - Crop Production 2019 Summary. Retrieved from https://www.nass.usda.gov/Publications/Todays_Reports/reports/cropan20.pdf.

Van Meirvenne, M., Scheldeman K., Baert,G., & Hofman G. (1994). Quantification of soil textural fractions of Bas-Zaire using soil map polygons and/or point observations. *Geoderma, 62*, 69-82. https://doi.org/10.1016/0016-7061(94)90028-0.

Vecchia, A. (1988). Estimation and model identification for continuous spatial processes. *Journal of the Royal Statistical Society. Series B (Methodological), 50*(2), 297-312. https://doi.org/10.1111/j.2517-6161.1988.tb01729.x.

Weber, D., and E. Englund. (1994). Evaluation and comparison of spatial interpolators II. *Mathematical Geology, 26*(5), 589-603. https://doi.org/10.1007/bf02089243.

Wettschereck D. (1994). *A Study of Distance-Based Machine Learning Algorithms*. (Ph.D thesis), Oregon State University.

Yang, C., Bradford, J., Wiegand, C. (2001). Airborne multispectral imagery for mapping variable growing conditions and yields of cotton, grain sorghum, and corn. *Transactions of the ASAE, 44*(6), 1983-1994. https://doi.org/10.13031/2013.6997.

CHAPTER 2:

TEMPORAL VARIABILITY OF CORN YIELD AND ITS EFFECT ON SITE-SPECIFIC MANAGEMENT ZONE DELINEATION

ABSTRACT

Yield data from previous years can be used to delineate farm-specific management zones for a variety of crop types. Recent work in New York showed that management zones for corn (*Zea mays* L.) should consider both average temporal yield and yield variability over time. In this study, multiple years of corn silage and grain yield were analyzed to calculate average temporal yield and yield variability, as impacted by the number of years of yield data included for the computation. Yield data were supplied by three dairy farms and three cash grain operations in New York. Farm-level temporal average yield and standard deviation were determined per farm using all data (seven to twelve years depending on the farm), as well as subsets of the latest three, four, or five years of data. Results showed that data should first be examined for the presence of a yield trend when estimating average yield. Including only the most recent four to five years of data reduced the risk of misrepresenting expected average yield when there was a yield trend. Otherwise, both the temporal average and standard deviation in yield were most consistent when all years of data were included. Up to 46% difference in year-to-year zone delineation was observed when an additional year of data were introduced. We conclude that farms interested in developing yield stability zones use at least four to five years of yield data and continue to add new years of data for improved delineation of zones over time.

INTRODUCTION

Under traditional farm management, one fertilizer rate (blanket application) is applied to a field, ignoring the variability in soil, topography, or local weather conditions (Srinivasan, 2006). The site-specific zone-based management approach aims to improve upon this one-rate approach by recognizing and managing heterogeneity within a field. Under a zone-based approach, a field is divided into areas of similar characteristics and resource needs (management zones) for individualized management. Using a zone-based management approach can help reduce costs for farmers resulting from better resource allocation, as well as reduce the farm's environmental footprint, while maximizing yield (Hornung et al., 2006).

Management zones can be generated based on one or more data layers. A yield map is among the most prominent data sources used by scientists to delineate management zones (Table 2.1). Yield maps are typically generated based on yield monitor data. Some have used single or multiple year yield maps derived from yield monitor data as the only data source to generate management zones (Anastasiou et al., 2017; Basso et al., 2007; Blackmore, 2000; Brock et al., 2005; Cox & Gerald, 2007; Diker et al., 2004; Johnson et al., 2003; Kharel et al., 2019a; Maestrini & Basso, 2018). Others have used various remotely sensed data layers, such as electrical conductivity (EC), topography and aerial imagery, and survey data, such as soil characteristics and texture, as well as qualitative assessments, such as growers' past experience in addition to a yield map, to delineate management zones (Fraisse et al., 2001; Hornung et al., 2006; Khosla et al., 2008; Moshia et al., 2015).

Table 2.1: Summary of farm information illustrating years of record, number of fields, field size statistics, yield statistics, crop rotation, location, soil type, crop type, and harvested area per year.

Citation	Data Source	Site-years	Crop Type	Zone Delineation
Anastasiou et al. (2017)	Yield monitor data	2	Grape	2 zones: high and low yield
Basso et al. (2007)	Yield monitor data	5	Corn, wheat, soybean	2 zones: high and low yield
Blackmore (2000)	Yield monitor data	6	Winter wheat, oilseed rape	3 zones: higher-yielding and stable, lower-yielding and stable, and unstable
Brock et al. (2005)	Yield monitor data	24	Corn, soybean	4 ~ 6 zones per field based on c-means clustering
Cox & Gerald (2007)	Yield monitor data	12	Soybean	4 yield zones: consistent-high, consistent-average, consistent-low, inconsistent
De Lara et al. (2018)	Soil water content based on a neutron probe	1	Corn	2 ~ 4 zones
Diker et al. (2004)	Yield monitor data	6	Corn	3 zones: stable high yield, stable low yield and unstable
Fraisse et al. (2001)	Yield monitor data, soil electrical conductivity (EC), slope, compound topographic index based on catchment area, and slope gradient	10	Corn, soybean, grain sorghum	4 or 5 zones
Gili et al. (2017)	Manually harvested yield and soil characteristic including altimetry, clay, silt, soil organic matter, phosphorus, EC, and pH	1	Corn	2 ~ 4 zones
Hoffmann et al. (2017)	Soil samples, EC, and simulated yield	9	Wheat, barley	3 zones: dune, mid-slope, and swale
Hornung et al. (2006)	Yield monitor data, bare soil imagery, soil organic matter, soil texture	3	Corn	3 zones: low, medium, high based on soil characteristics and previous yield
Kharel et al. (2019a)	Yield monitor data	847	Corn grain and silage	4 zones: zone 1 (high yield, low variability), zone 2 (high yield, high variability), zone 3 (low yield, high variability), and zone 4 (low yield, low variability)
Khosla et al. (2008)	Aerial imagery, producer experience, topography, and yield monitor data	15	Corn	3 zones: low, medium, high yield

Citation	Data Source	Site-years	Crop Type	Zone Delineation
Moshia et al. (2015)	Aerial imagery, topography, producer's past experience, and yield monitor data	2	Corn	3 zones: low medium and high yielding
Johnson et al. (2003)	Yield monitor data	2	Corn, winter wheat	4 zones: low, medium low, medium high, and high electrical conductivity (EC)
Leroux et al. (2018)	Aerial imagery, soil electrical conductivity (EC), topography, vine vigor, soil type, and water availability	2	Banana, grape, wheat	5 zones: low, low-medium, medium, medium-high, and high risk based on soil EC, elevation, and normalized difference vegetation index (NDVI)
Santesteban et al. (2013)	Soil conductivity, elevation, and yield response based on NDVI	27	Grape	3 Zones: class 1 (high NDVI, low to medium EC), class 2 (medium NDVI and EC), class 3 (low NDVI, high EC).
Maestrini and Basso (2018)	Yield monitor data	1625	Corn, soybean, wheat, cotton	3 zones: low and stable, high and stable, unstable yield
Melo Damian et al. (2016)	Wheat grain yield components: spikes per square meters, grains per spike, grain weight, soil and chemical properties	1	Wheat	3 zones: low intra-zonal, intermediate intra-zonal, high inter-zonal variation (based on grain yield, mean grain weight, pH, spikes per square meter, C)
Monzon et al. (2018)	Soil depth, frost risk and water table data	14	Soybean	4 zones: zone 1 (shallow soils, low frost risk, deep water table); zone 2 (intermediate soil depth, low frost risk, deep water table); zone 3 (deep soils, low frost risk, deep water table); and zone 4 (deep soils, high frost risk, water table < 3 m from surface)

When delineating management zones based on past yield data, it is important to account for both spatial and temporal yield variability (Kharel et al., 2019a). Spatial yield variability represents within-field variability in yield caused by a non-uniform distribution of soil

properties, soil moisture, pest pressure, rooting depth, and other factors (Sawyer, 1994). In addition to spatial yield variability, yield can vary greatly across years due to a variety of factors including weather, management, and topography. Year-to-year variation in weather can be substantial and can interact with both biotic and abiotic factors in the field (Andresen et al., 2001; Kravchenko et al., 2005; Mallory & Porter, 2007). Field management can also contribute to this temporal variability in yield. For example, Smith et al. (2007) compared four management systems: conventional tillage, no-till, low-input, and organic management, and followed a three-year corn, soybean [*Glycine max* (L.) Merr.], and winter wheat (*Triticum aestivum* L.) rotation over a 12-year period. The results showed that the temporal yield variability was 40% lower under low-input management and 33% lower in the no-till management compared to the conventional system. Kim et al. (2020) performed spatio-temporal analysis of corn grain yield data from six fields over a six-year period through empirical orthogonal functions. They concluded that topography had a large impact on temporal yield variability, explaining between 52 and 56% of the variability of yield over multiple years.

Some studies showed that temporal yield variability can be greater than spatial variability in yield. For example, Porter et al. (1998) analyzed ten years of corn grain and soybean yield data from three farms in the northern Corn Belt and found the temporal yield variability to be three times greater than the spatial yield variability. Eghball & Varvel (1997) also showed temporal yield variability to be more prominent than spatial variability in yield based on analyses of 20 years of corn grain, soybean, and sorghum [*Sorghum bicolor* (L.) Moench] yield data from Nebraska. The study concluded that site-specific management can be extremely difficult in the presence of high temporal yield variability. In New York, Kharel et al. (2019a) showed greater temporal variability than spatial variability in corn silage yield, based on

78 fields with three years of yield data. The same study furthermore showed that spatial and temporal variability in corn yield were not well-correlated.

Many have attempted to account for yield variability over time when delineating management zones. Blackmore (2000), Diker et al. (2004) and Maestrini & Basso (2018) used a three-zone approach: higher-yielding and stable; lower-yielding and stable; and unstable. Blackmore (2000) based its conclusion on six site-years of wheat (*Triticum spp* L.) and canola (*Brassica napus* L.) in England. Diker et al. (2004) used six site-years of corn in Colorado, USA, and Maestrini & Basso (2018) used 1,625 site-years of corn, soybean, wheat, and cotton (*Gossypium spp* L.) from Arkansas, Kansas, Colorado, Michigan, Illinois, Iowa, and Indiana, USA. Others, such as Cox & Gerald (2007) and Kharel et al. (2019a) proposed a four-zone approach: stable high, unstable high, stable low, and unstable low yielding. Cox & Gerald (2007) used twelve site-years of soybean in Mississippi, while Kharel et al. (2019a) used 847 site-years of corn silage in New York, USA for zone delineation. These studies reflect that temporal yield variability is increasingly taken into consideration in developing site-specific management zones.

A few studies have stated the importance of having a sufficient number of years of data when analyzing temporal yield variability (Bakhsh et al., 2000; Kitchen et al., 2005; Maestrini & Basso, 2018). Bakhsh et al. (2000) analyzed temporal yield variability by studying two years of corn grain and one year of soybean from a field located in Iowa, USA. The study noted the lack of temporal stability in yield in the analyzed data and suggested more years of data are needed. Kitchen et al. (2005) analyzed two fields with ten site-years of corn and seven site-years of soybean, both located in Missouri, USA. This study also questioned the number of years of data required to appropriately represent long-term climate trends and their effect on yield. Similarly, Maestrini & Basso (2018) stressed the importance of having a sufficient number of years of data

when delineating management zones in their study of 1,625 site-years of corn grain, soybean, wheat and cotton across the Midwest. They stated that using fewer years of data reduced the accuracy when estimating both the average yield and standard deviation of yield.

Some have attempted to propose a range of number of years necessary to appropriately account for temporal yield variability. Leroux et al. (2018) analyzed temporal variability of wheat and canola based on six and eight years of data from two fields located in England and France. Temporal standard deviation for wheat was around 1.43 Mg ha^{-1} while it was around 1.71 Mg ha^{-1} for Canola. The study concluded that at least four years of yield data were needed to delineate reliable management zones. A five-year study from 1991 to 1995 by Lamb et al. (1997) with corn grain in Minnesota suggested that at least five years of data were required to appropriately account for variation in yield over time. The average grain yield over the five-year period from 1991 to 1995 were 4.52 Mg ha^{-1}, 2.76 Mg ha^{-1}, 2.82 Mg ha^{-1}, 3.20 Mg ha^{-1}, and 3.58 Mg ha^{-1} with the standard deviation of around 0.72 Mg ha^{-1} It is important to note that the fields studied by Lamb et al. (1997) were managed as uniformly as possible in terms of water, pesticide, and fertilizer application. Ping & Dobermann (2005) analyzed six years of yield monitor data of corn grain from an irrigated field located in Nebraska, USA. The average yields for corn ranged between 11.5 Mg ha^{-1} to 13.5 Mg ha^{-1} with the temporal standard deviation of around 0.71 Mg ha^{-1}. The study suggested using five to ten years of yield monitor data to properly account for year-to-year yield variability, especially for fields with frequent management changes and/or wide climatic variation. Corn silage and grain studies are needed in the Northeast region of the USA because temporal yield variability is expected to be large due to larger variability in weather, landscape, soils, and management practices, as well as the lack of irrigation systems on most dairy and cash grain farms in the Northeast (USDA, 2020).

The objective of this study was to analyze the impact of number of years of data on estimating the farm-level temporal average yield and its variability for both corn grain and corn silage grown in farms in New York State. We first assessed the impact of using three-, four-, and five-year moving averages, and all data (extended window) on the stability of temporal average yield and standard deviation. We hypothesized that both the farm-level temporal average yield and farm-level temporal standard deviation would become more consistent over time as we increase the number of years of data included in the assessment. The impact of excluding older years of data on the estimation of temporal average yield and standard deviation was also analyzed to determine the impact of a trend on average yield and yield variability. Our hypothesis was that using all available data would under- or over-estimate the farm-level temporal average yield, if a yield trend was present. The farm-level temporal standard deviation was expected to become more consistent as more years of data are included. Lastly, the number of years of data needed to delineate management zones that were consistent over time, while also accurately reflecting yield trends, was determined for each farm based on the analyses above.

MATERIALS AND METHODS

Yield Monitor Datasets

Yield monitor data were obtained from six farms: three dairy farms and three cash grain operations all located in New York. Dairy farms provided silage yield monitor data (hereafter referred to as Silage A, B, and C), and cash grain operations provided grain yield monitor data (hereafter referred to as Grain A, B, and C). Each of the farms had at least seven years of yield records, with actual number of years ranging from seven to twelve (Table 2.2). The number of fields, area harvested, soil types, practiced crop rotations, and geographical locations within New

York were vastly different among farms (Table 2.2). The area harvested by each farm also varied yearly. Missing data are present in Silage B in 2014 due to equipment breakdown at harvest.

Data Processing

Before analysis, yield monitor data need to be cleaned due to the presence of systematic and random errors (Dobermann & Ping, 2004; Vega et al., 2019; Kharel et al., 2019b). To read and export the yield monitor data in AgLeader format, SMS Advanced software (Ag Leader Technology, Ames, IA, USA) was used. A standardized post-harvest data cleaning protocol, as described by Kharel et al. (2018) and based on Kharel et al. (2019b), was performed using Yield Editor (Sudduth et al., 2012; Sudduth & Drummond, 2007). This data cleaning protocol aims to remedy issues such as pass overlap (driving over areas already harvested), yield estimate extremes due to equipment slowing down or speeding up, and inconsistencies between sensor delays (flow and moisture delays) and the harvester location. Cleaned yield monitor data were then used to create rasterized yield maps. Yield monitor datasets consist of irregularly placed point estimates of yield that need to be interpolated into a rasterized yield map to compare multiple years of yield data. Cho et al. (2021) explored widely used methods in creating rasterized yield maps for corn grain and silage and concluded that kriging with the Matérn covariance function resulted in the most accurate yield maps. This approach was used to generate yield maps with four square meter pixels (2 x 2 m) from yield monitor data collected from the six farms. All analyses were performed with R (R Core Team, 2019). The GpGp package (Guinness & Katzfuss, 2019) was used to create rasterized yield maps. The data structure supported by "raster" package in R (Hijmans, 2019) was used to represent yield maps.

Table 2.2: Summary of farm information illustrating years of record, number of fields, yield statistics, field size statistics, location, soil type, and the area of crop harvested yearly.

	Unit	Silage A	Silage B	Silage C	Grain A	Grain B	Grain C
Years of record	Years	11	9	10	12	7	8
Number of fields		238	117	132	273	51	222
Average field size	Ha	10.0	9.2	7.4	9.6	19.9	4.8
Smallest field	Ha	0.9	0.6	0.4	0.3	2.8	0.3
Largest field	Ha	93.1	56.9	44.0	52.2	72.9	28.4
Total area analyzed	Ha	4,698	2,633	1,975	5,355	2,416	2,097
Area-weighted average yield	Mg ha^{-1}	50.3	42.0	47.2	10.9	11.6	11.2
Lowest yielding field	Mg ha^{-1}	27.3	21.0	16.1	2.8	8.3	2.8
Highest yielding field	Mg ha^{-1}	81.7	59.8	72.1	16.5	14.0	14.7
Average spatial standard dev.	Mg ha^{-1}	9.1	9.4	8.1	2.3	1.7	2.0
Crop rotation		3-4 yrs corn; 4-5 yrs hay; single yr corn	2-3 yrs corn; 3-5 yrs hay; single yr corn	Corn/soybean rotation; 3-4 yrs corn, hay	Corn/soybean rotation; 3-4 yrs corn, hay	2~3 yrs of corn; 2-3 yrs soybean	Corn/soybean rotation
Location in New York		Central	Northern	Western	Central	Central	Eastern
Most common soil type		Honeoye (fine-loamy, mixed, semi-active, mesic Glossic Hapludalfs)	Hogansburg (coarse-loamy, mixed, semi-active, frigid Aquic Eutrudepts)	Conesus (fine-loamy, mixed, active, mesic Glossaquic Hapludalfs)	Ontario (fine-loamy, mixed, active, mesic Glossic Hapludalfs)	Honeoye (fine-loamy, mixed, semi-active, mesic Glossic Hapludalfs)	Lima (fine-loamy, mixed, semi-active, mesic Oxyaquic Hapludalfs)
Second most common soil type		Lima (fine-loamy, mixed, semi-active, mesic Oxyaquic Hapludalfs)	Swanton (coarse-loamy over clayey, mixed over illitic, super-active, nonacid, frigid Aeric Epiaquepts)	Lansing (fine-loamy, mixed, active, mesic Glossic Hapludalfs)	Edwards (marly, euic, mesic Limnic Haplosaprists)	Kendaia (fine-loamy, mixed, semi-active, nonacid, mesic Aeric Endoaquepts)	Kendaia (fine-loamy, mixed, semi-active, nonacid, mesic Aeric Endoaquepts)
Average area harvested 2008	Ha	-	-	-	130	-	-

	Unit	Silage A	Silage B	Silage C	Grain A	Grain B	Grain C
2009	Ha	186	-	-	164	-	-
2010	Ha	253	-	171	354	-	-
2011	Ha	381	336	161	322	-	-
2012	Ha	323	305	171	451	-	-
2013	Ha	108	348	170	494	363	277
2014	Ha	222	-	248	430	339	203
2015	Ha	396	307	204	644	322	301
2016	Ha	542	544	388	471	304	239
2017	Ha	785	170	190	628	428	254
2018	Ha	754	366	73	761	325	251
2019	Ha	748	258	199	505	335	299

Measuring Farm-Level Temporal Average Yield and its Variation

Our project aims to provide key insight with regards to temporal yield and its variation to allow for better farm-level management. In the analyses, x_{ijk} represents the k-th raster pixel in year j on field i. Subscript i was used to index fields within a farm ($i = 1, \ldots, N$), N is the number of fields for a given farm, j is used to reference the year in which the data were collected ($j = 1, \ldots, T_i$), T_i is the number of years of data for the given field within the farm, k is used to reference a location (pixel) within a field ($k = 1, \ldots, M_i$), and M_i is the number of raster pixels for the field. Thus, $Y(x_{ijk})$ represents the yield estimate at location k, for field i, collected in year j.

Temporal average yield of a raster pixel, $\hat{Y}(x_{ik})$ was calculated to measure the expected yield at a given location for a field at any given year (1). Temporal variance of a raster pixel, $\mathrm{var}(Y(x_{ik}))$, was measured to estimate the variability of yield over time at a given location within a field (2).

$$\hat{Y}(x_{ik}) = \frac{1}{T_i}\sum_{j=1}^{T_i} Y(x_{ijk}) \tag{2.1}$$

$$\mathrm{var}(Y(x_{ik})) = \frac{1}{T_i}\sum_{j=1}^{T_i}\left(Y(x_{ijk}) - \hat{Y}(x_{ik})\right)^2 \tag{2.2}$$

These two measures of temporal average yield and temporal variance at the pixel level can be averaged over all fields harvested by the farm and farm-level estimates can be measured using the equations (3) and (4).

$$\text{Farm-level temporal average yield} = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{M_i}\sum_{k=1}^{M_i}\hat{Y}(x_{ik}) \tag{2.3}$$

$$\text{Farm-level temporal standard deviation} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\frac{1}{M_i}\sum_{k=1}^{M_i}\mathrm{var}(Y(x_{ik}))} \tag{2.4}$$

Notice that both measures were weighted by the number of pixels in a field to account for the wide variety in field sizes within a farm (Table 2.1). These whole farm average yields and temporal standard deviations were calculated at the farm-level, instead of at the field-level, as farms typically manage just a handful of different rates for resources like manure, fertilizer, and seed.

Year-to-Year Yield Comparisons

Year-to-year variation in yield can affect the stability of both farm-level temporal average yield and standard deviation. Yearly area-weighted average yield was calculated for each farm to estimate whole farm yield per year. For a given farm and year, *j*, farm-level area-weighted average yield was calculated by determining the average yield of all fields harvested that year (field-size weighted average) as shown in equation (5).

$$\text{Yearly area-weighted average yield} = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{M_i}\sum_{k=1}^{M_i}Y\left(x_{ijk}\right) \tag{2.5}$$

It is important to distinguish farm-level area weighted average yield and farm-level temporal average yield. The former represents overall production of a farm at a given year, while the latter attempts to represent overall production of a farm over multiple years, thus at any given time frame. The farm-level area-weighted average yield was calculated to compare the level of production among three dairy farms (Silage A, B, and C) and three cash grain operations (Grain A, B, and C). The area-weighted farm-level average yields were compared against the New York state-wide average yield for corn silage and corn grain production as well (USDA, 2020).

Forward analysis: Impact of number of years of yield data on estimating farm-level temporal average yield and standard deviation

Temporal average yield and standard deviations change every year because the farm updates its data yearly. Depending on how many years of data are used to estimate the farm-level temporal average yield and standard deviation, the degree to which the incoming data affect those measures will vary greatly. Both (2.1) and (2.2) are inversely proportional to T_i, the number of years of data, suggesting that the longer the number of years of data, smaller the impact of one individual year's data on the overall value. The impact of the number of years used on the stability of the farm-level temporal average yield and temporal standard deviation was analyzed (hereafter referred to as forward analysis). While it is important to make sure that the measures accurately reflect a farm's yield trend and its variability, measures that fluctuate greatly by year can hinder management decisions. In Forward Analysis, the temporal average yield and temporal standard deviation were calculated per farm using three, four, and five years, as well as all available data (extended window) throughout each farm's data, as shown in Figure 2.1a, b, c, and d. The analysis aims to compare stability of temporal average yield and standard deviation impacted by the number of years of data included. The term "*stability*" will be used to describe the amount of change in each measure year-to-year. A measure is deemed stable if the amount of change in temporal average yield or standard deviation is small. Stability was quantified through the absolute percent change (APC), which denotes percent change in both measures, temporal average yield and standard deviation, in two successive years and allows unit-less comparison of stability among farms, as shown in equation (5):

$$\text{Absolute Percent Change (APC)} = \left|\frac{\text{Measure at Year (N + 1)}}{\text{Measure at Year N}} - 1\right| \times 100 \qquad (2.6)$$

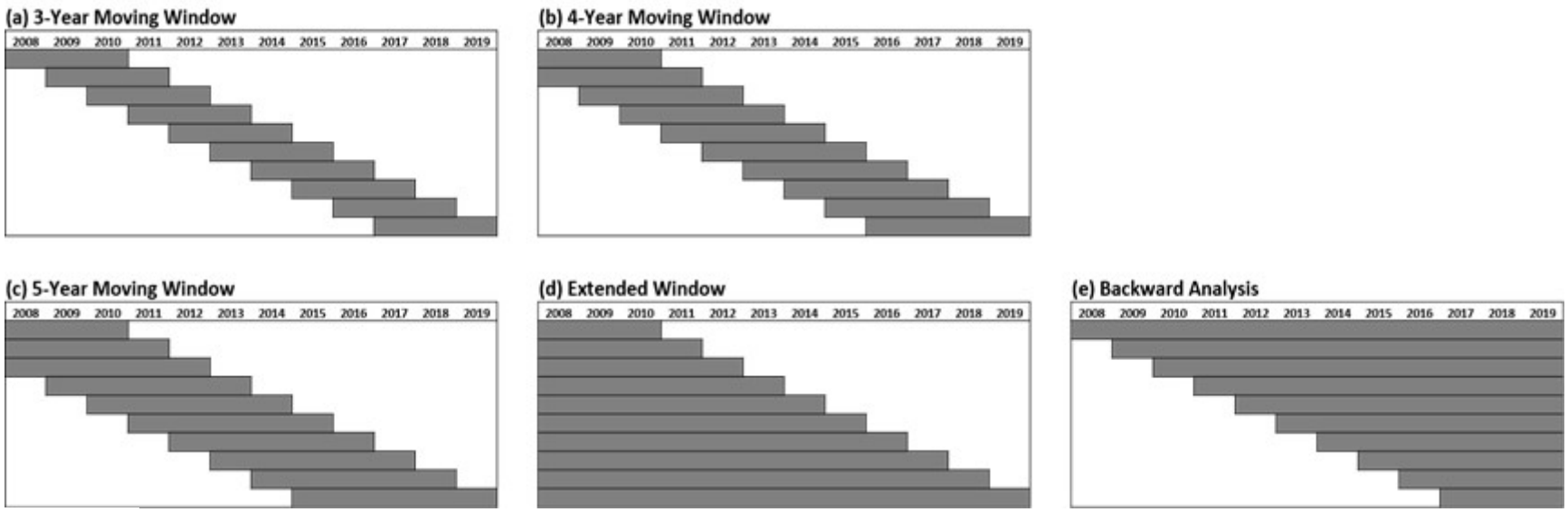

Figure 2.1: Visualization of forward analysis using three, four, five years moving window (a, b, and c, respectively), extended window (d), and backward analysis (e). Farm-level temporal average yield and farm-level temporal standard deviations are calculated using the data collected in highlighted regions.

Backward analysis: Impact of using older data on estimating temporal average yield and standard deviation

Past studies (Lamb et al., 1997; Leroux et al., 2018; Ping & Dobermann, 2005) all showed wide year-to-year variations in yield. Moreover, yield variation in New York is expected to be even greater due to variability in weather, landscape, soils, and management practices (Kharel et al., 2019a). Considering the yield variation and trend over the past decade (USDA, 2020), inclusion of the older data on temporal average yield and standard deviation was evaluated, as including data from many years ago could misrepresent the farm's more current overall yield and its variation. For the analysis, temporal average yield and temporal standard deviation were calculated per farm as of 2019, using the latest three years through all years of available data. This analysis, referred to as Backward Analysis (Figure 2.1e), was designed to quantify the impact of including older years of data on temporal average yield and standard deviation and to decide how many previous years of data should be used to appropriately reflect the farm's year-to-year production and its variability.

Years of data and delineation of yield stability-based management zones

Multi-year rasterized yield maps based on yield monitor data can be used as a base layer in delineating management zones. Kharel et al. (2019a) proposed a yield stability-based management zone, in which a field can be broken up into four different zones at the farm level. This zone delineation method uses the farm-level temporal average yield and farm-level temporal standard deviation as cut-offs to classify parts of a field into four zones. Zone 1 represents high-yielding stable field sections, zone 2 represents high-yielding unstable field sections, zone 3 represents low-yielding unstable field sections, and zone 4 represents low-

yielding stable parts of a field. This particular zone delineation method assumed that yield and yield variability differences among zones reflect the ability of a crop to yield and buffer weather extremes and suggests that zones may be managed differently for the greatest return on investment (Kharel et al., 2019a). Yield-stability based management zones described in Kharel et al. (2019a) were used to generate zone maps using a range in numbers of years (three, four, and five most recent years, and all the available data) for two fields, one from Grain A and another from Silage A. These farms and fields had the largest number of years of data. Changes in management zones were quantified by calculating the total area and percent change in area classified into different yield stability zones when more years of data (harvest seasons) were included.

RESULTS AND DISCUSSION

Year-to-Year Yield Comparisons

Area-weighted average yields among the three grain farms ranged from 10.9 Mg ha^{-1} for Grain A to 11.6 Mg ha^{-1} to Grain B. It should be noted that Grain A had the longest yield record (twelve years) versus seven and eight years for Grain B and C, respectively. For the three dairies, yields for Silage B averaged 42 Mg ha^{-1}, versus 47.2 Mg $ha^{-1}$for Silage C and 50.3 Mg ha^{-1} for Silage A (Table 2.2). The variability among farms was not unexpected, given the diversity of soil types on the farms, inconsistent yearly weather patterns in different locations, differences in total land-base in corn among the farms, and varying number of years of data, in addition to rotation and management differences among farms.

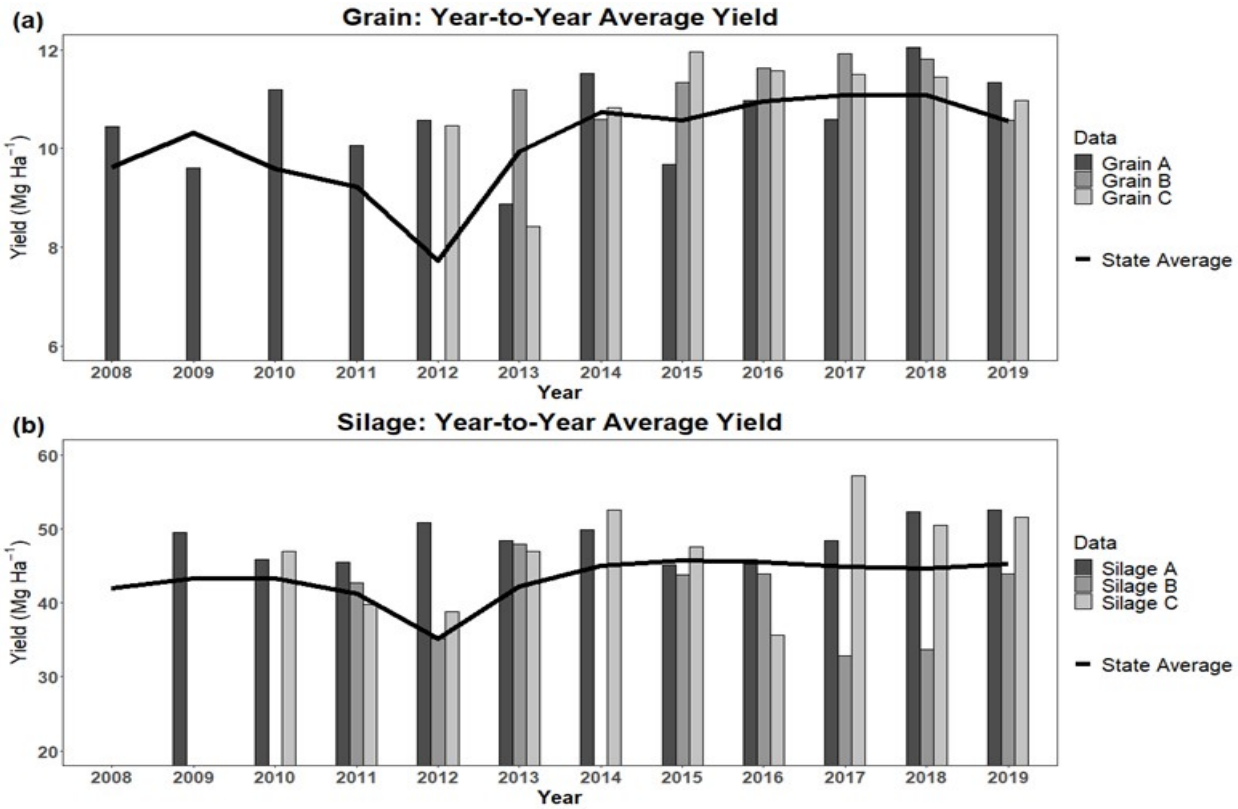

Figure 2.2: Year-to-year yield trend of six farms over the past eleven years for grain data (a), and ten years for silage data (b), compared to the New York state-wide average yield (USDA, 2020).

While both state-wide average yield and the area-weighted average yield of the six farms varied yearly (Figure 2.2), there are a few noticeable differences. The year 2012 was classified as a drought year, which is reflected in low state-wide corn grain and silage yields (Figure 2.2). However, yields recorded for the farms in this study were less impacted by the drought than was shown for statewide averages. This may reflect an improved biological buffer capacity, described as a soil's and plant's ability to adjust to changes in weather (Meisinger et al., 2008; Long & Ketterings, 2016). This may especially be the case for corn silage fields that are more likely to have a manure history, which can contribute to greater buffer capacity (Long and Ketterings, 2016). State-wide weather and yield data are not necessarily consistent with data trends from individual farms, as shown by yield data from Grain A and C.

The temporal yield variation differed by farm. For Grain A, B, and C, the temporal standard deviation was 2.51 Mg Ha^{-1}, 1.45 Mg Ha^{-1}, and 2.30 Mg Ha^{-1} and for Silage A, B, and C, it was 10.04 Mg Ha^{-1}, 10.71 Mg Ha^{-1}, and 12.43 Mg Ha^{-1} respectively. Temporal yield variations of fields located in New York were larger than those of fields located in Minnesota or Nebraska. Lamb et al. (1997) and Ping & Dobermann (2005) analyzed year-to-year yield variation of corn grain based on fields in Nebraska and Minnesota respectively and showed that the year-to-year variation to be around 0.71 Mg Ha^{-1}, which is noticeably smaller than the amount of temporal variation exhibited by the three cash grain operations. High year-to-year yield variation is expected given large variability in weather, landscape, soils, and management practices in New York (Kharel et al., 2019a). Due to such large temporal yield variation, we may also expect large variation in yield stability-based management zone delineation, as it is impacted by the farm-level temporal average yield and standard deviation.

Assessment of Forward Analysis

The degree to which both farm-level temporal average yield and standard deviation are affected by new incoming data varied by farm and by years (Figure 2.3). However, on average, using all the available data resulted in most consistent estimates for both farm-level temporal average yield and standard deviation. Across fields, the yearly change in farm-level temporal average yield for grain was 26, 16, and 8% for three-, four-, and five-year moving average, respectively, versus 5% when all years of data (extended window) were used. For silage, extended window resulted in a 4% average yearly change in farm-level temporal average yield versus 17, 8, and 7% for three, four, and five-year moving averages, respectively. For grain, farm-level temporal standard deviation change was 3.7, 2.5, 2.3, and 1.6% for the three-, four-,

five-year moving average and extended window, respectively, versus 5.3, 3.0, 2.2, and 1.8% for three-, four-, five-year moving averages, and an extended window, respectively (Figure 2.3).

Estimating temporal standard deviation becomes a challenge when only a fixed number of previous years (three, four, or five) of data are include, since there tend to be only a few fields with more than three years of data in any given time frame due to crop rotations on both grain and dairy farms. This was notably a problem for Grain C, a farm with a corn and soybean rotation; from 2014 to 2016, estimation of farm-level temporal standard deviation using the three most recent years of data for Grain C was impossible, because there were no fields with three consecutive years of corn (Figure 2.3x). Based on forward analysis, we concluded that the extended window approach (using all years of data) results in more stable measurements of both temporal average yield and standard deviation over multiple years for both silage and grain.

These results support the assessment made in Leroux et al. (2018), namely that the analysis of historical yield data is challenging due to a wide range of growing conditions across years, and that obtaining reliable conclusions using fewer than four years of historic yield data is risky. Our data suggest that estimations of both farm-level temporal average yield and farm-level temporal standard deviation are impacted by year-to-year APC and that using all available data can lead to most stable measurements. By including more years of data, both measures became more robust to new incoming data, making them less susceptible to outlier years due to extreme weather conditions. Also, farm-level temporal standard deviations were especially unstable when only a few years of data were used, because the measure requires at least three years of data per field to estimate. As we limit the number of years used to estimate the measure, the number of fields that have more than three years of data also becomes limited.

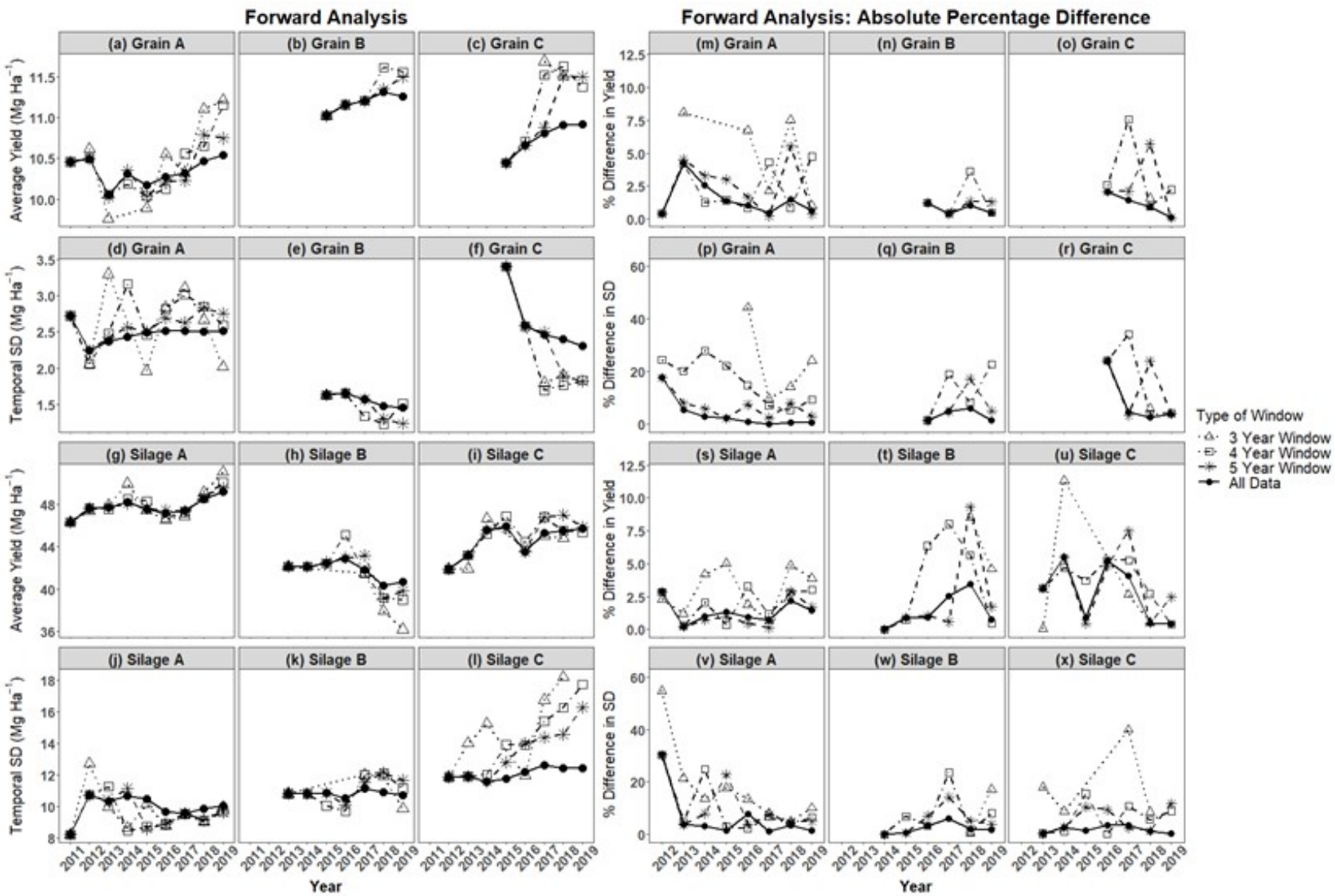

Figure 2.3: Change in farm-level temporal average yield (a, b, c, g, h, i) and temporal standard deviation (d through l) for grain (a through f) and silage (g through l). Absolute percent change (APC), calculated by taking the absolute value of the ratio between measures in two successive years minus 1, were calculated and visualized (m through z).

Assessment of Backward Analysis

For grain data, the temporal average yield across years was generally higher when only the most recent three to five years of data were included in the assessment, as opposed to using six or more years of data (Figure 2.4a). This is consistent with our observation of the area-weighted average yield of the three grain datasets (Figure 2.2a), which showed a weak but consistently positive yield trend.

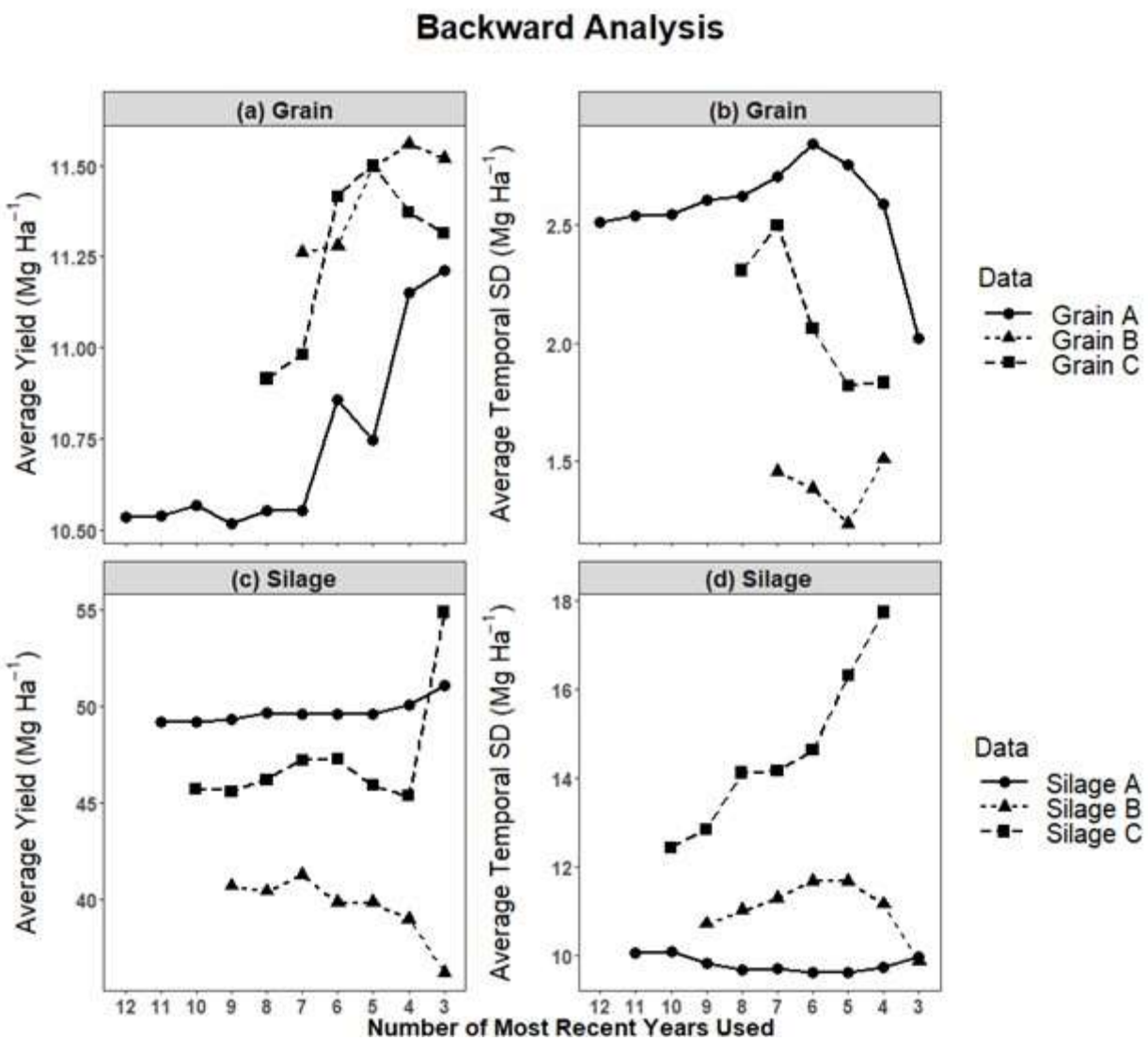

Figure 2.4: Line graphs of farm-level average yield (a and c) and farm-level temporal standard deviation (b and d) as of 2019 using the most recent three to twelve years of data for grain (a and b) and three to eleven years of data for silage (c and d).

For silage, the yield data did not show a yield trend over time; yields differed less than 2 Mg ha^{-1} across years, regardless of how many years of data were included to derive average temporal yield, except for Silage B (Figure 2.4c). Yield for Silage B in the last three years (2016 to 2019; 50 Mg ha^{-1}) was considerably higher compared to the average production of 40 Mg ha^{-1} from 2010 to 2015 (Figure 2.2). The improved production level may be due to weather patterns in the last three years for this farm and/or a shift in management and variety selection. For Silage A and C, no yield trend was present over the span of the data provided.

The farm-level temporal standard deviations varied considerably among farms for both grain and silage (Figure 2.4b and 2.4d). For Grain A and Silage B data, standard deviations increased when three to six years of data were included in the estimation, but then converged to 5.6 Mg ha^{-1} and 10 Mg ha^{-1}, respectively, when more years of data were used. Silage A data, on the other hand, show more consistency in farm-level temporal standard deviation over time (less than 0.5 Mg ha^{-1} difference over time). The standard deviation decreased for Silage B as more years of data were included. For Grain B and C, no clear trend was detected, but for these two datasets only four and five data points were available for analysis.

The study by Maestrini & Basso (2018) stressed the importance of having a sufficient number of years of data to estimate both average yield and standard deviation for delineating management zones. Our study similarly suggested that use of all available years of yield data on a farm is beneficial to obtain stable measurements of farm-level temporal average yield and standard deviation. However, as shown in Figure 2.4a, estimation of average yield using all the available data can lead to inaccurate estimation of yield when yield trends are present on a farm.

Impact on yield stability-based management zone delineation

Changes in both the farm-level temporal average yield and standard deviation, due to inclusion of additional years of data, impacted zone delineation (Figure 2.5). The example field from Grain A showed differences in zone classification between two successive years of 43%, 46%, 19%, 12%, 4%, 9%, 4%, 1%, and 2%, respectively, for zone maps shown in Figure 2.5(k) ~ 5(s). For the example field from Silage A, the differences were 11%, 13%, <0.5%, 1%, 1%, 3%, 2%, and <0.5%, respectively, for zone maps shown in Figure 2.5(j) ~ (q). The larger changes shown for the field of Grain A reflects the yield trend of the farm. Changes in zone delineations at the field level were most drastic (Figure 2.5 (k) and (l); Figure 2.5 (j) and (k)) where farm-level temporal average yield and standard deviation were highly variable, as shown in Figure 2.4.

More than 20 years ago, Eghball &Varvel (1997) concluded that using yield maps to guide site-specific management may not be useful when temporal variability is high. Kharel et al. (2019a) showed that temporal and spatial variability are not correlated, suggesting the need for zone delineation methodologies that take average yield and variability in yield over time (yield stability) into account. Data shown in Figure 2.5 suggest that year-to-year yield variation can be large and that delineation of yield stability zones will be impacted by number of years of data included in the assessment. However, it is also clear from these data that there are parts within the field that are consistently high-yielding (zone 1) versus other parts that are consistently low-yielding (zone 4), regardless of how many years of data were used in the assessment. As crops grown in each of the management zones will likely respond differently to input and weather patterns during the growing season (Kharel et al. 2019a; Long and Ketterings, 2016), zone delineation will allow for more informed site-specific management.

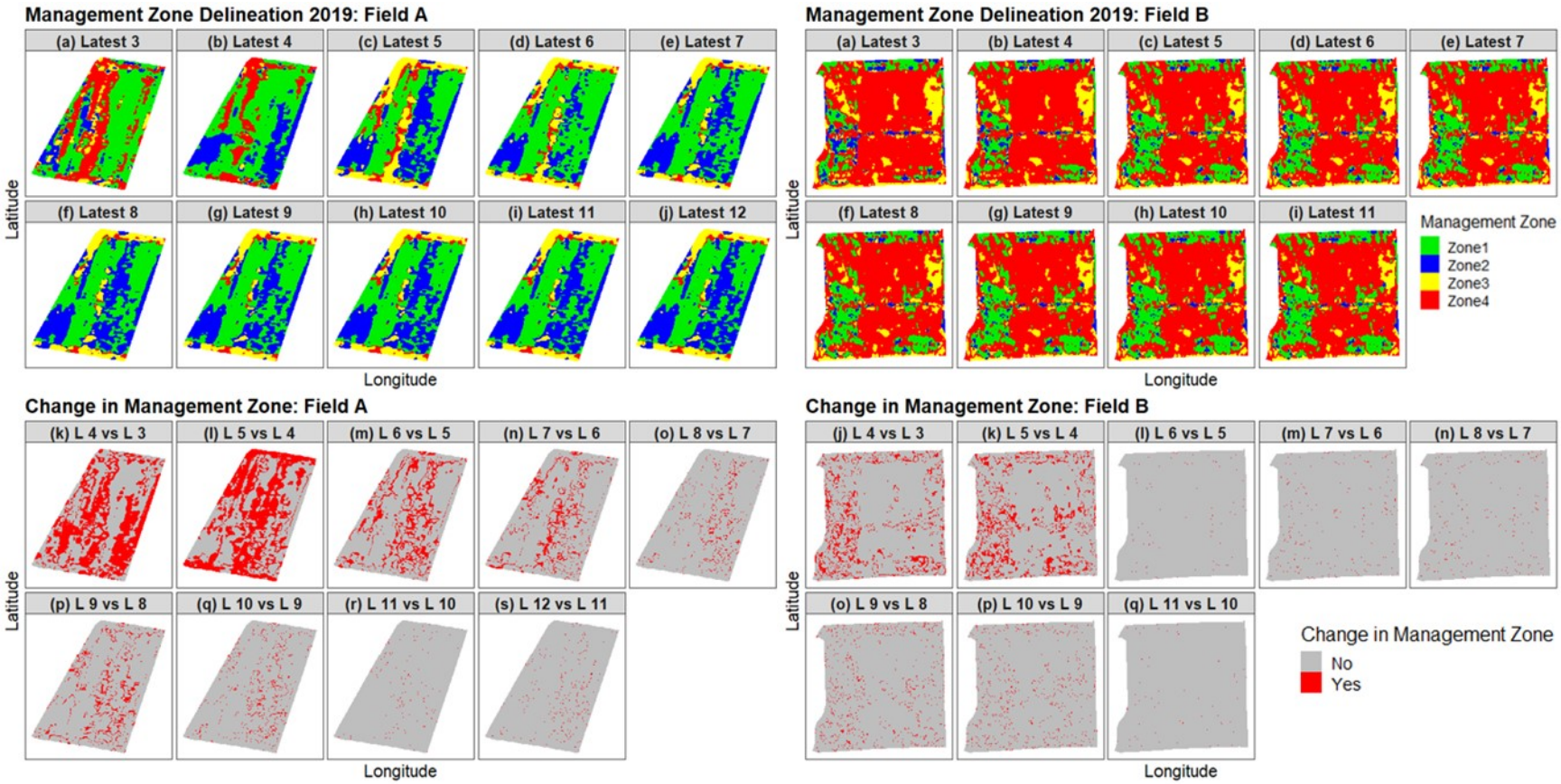

Figure 2.5: Zone maps for a field in Grain A (referred to as Field A) as of 2019 using the most recent 3 to 12 years (Figure 2.5(a) – 2.5(j)) and a field in Silage A (referred to as Field B) using the most recent 3 to 11 years (Figure 2.5(a) – 2.5(i)). Zone 1 is colored green, zone 2 blue, zone 3 yellow, and zone 4 red. Zone 1 represents areas that are high yielding (above the farm average) with low year-to-year variability in yield (stable); zone 2 is high yielding as well but variable over years (unstable); zone 3 has low yields and high variability; and zone 4 has low yield and low variability. Difference in zone maps between two successive years are showed in Figure 2.5(k) to 2.5(s) and figure 2.5(j) – 2.5(k). Area colored in red is the part where the zone delineations differ in two successive years (e.g., L 4 vs L 3 represent difference in zone delineations using the latest four vs three years of data).

Adjustments of management and resource allocation can be tested per zone once zones are delineated. In-season, weather informed adjustments are most likely to benefit yield, yield quality, and resource use efficiency in areas that are delineated as relatively unstable over years (zones 2 and 3), while for more stable yielding areas (zones 1 and 4), adjustments might be smaller. Research is ongoing in New York to establish an analytical approach that takes yield and yield stability into account when analyzing the response of a crop to specific resource allocation (on-farm trials) such as seed or fertilizer.

CONCLUSIONS

Forward analysis showed that use of all available data (≥5 years) leads to more stable measurements of temporal average yield and temporal standard deviations for both silage and grain. Backward analysis showed that use of older years of data can lead to underestimation of the overall temporal average yield when farm-level improvements in yield over time have occurred. For farms with a yield trend, the temporal standard deviation of yield was not impacted by inclusion of older years of data in the assessment. Based on these findings, we conclude that estimations of farm-level temporal average yield become more accurate when more years of data are included (≥5 years), unless a trend in yield has occurred. In the latter case, it is recommended to use the most recent four to five years of data to derive a farm-average yield. The standard deviation in yield was not impacted by yield trends over time and showed more stable values with additional years of data. While whole farm average yield and standard deviation of yield can be determined based on as few as three years of data, it is recommended to include at least four to five years of data for more accurate zone delineation.

REFERENCES

Anastasiou, E., Tsiropoulos, Z., Balafoutis, T., Fountas, S., Templalexis, C., Lentzou, D., & Xanthopoulos, G. (2017). Spatiotemporal stability of management zones in a table grapes vineyard in Greece. *Advances in Animal Biosciences, 8*(2), 510–514. https://doi.org/10.1017/S2040470017000632

Andresen, J. A., Alagarswamy, G., Rotz, C. A., Ritchie, J. T., & LeBaron, A.W. (2001). Weather impacts on maize, soybean, and alfalfa production in the Great Lakes region, 1895–1996. *Agronomy Journal, 93*(5), 1059–1070. https://doi.org/10.2134/agronj2001.9351059x

Bakhsh, A., Jaynes, D. B., Colvin, T. S., & Kanwar, R. S. (2000). Spatio-temporal analysis of yield variability for a corn-soybean field in Iowa. *Transactions of the ASAE, 43*(1), 31–38. https://doi.org/10.13031/2013.2684

Basso, B., Bertocco, M., Sartori, L., & Martin, E. C. (2007). Analyzing the effects of climate variability on spatial pattern of yield in a maize–wheat–soybean rotation. *European Journal of Agronomy, 26*(2), 82–91. https://doi.org/10.1016/j.eja.2006.08.008

Blackmore, S. (2000). The interpretation of trends from multiple yield maps. *Computers and Electronics in Agriculture, 26*(1), 37–51. http://dx.doi.org/10.1016/S0168-1699(99)00075-7

Brock, A., Brouder, S. M., Blumhoff, G., & Hofmann, B. S. (2005). Defining yield-based management zones for corn-soybean rotations. *Agronomy Journal, 97*(4), 1115–1128. https://doi.org/10.2134/agronj2004.0220

Cho, J. B., Guinness, J., Kharel, T. P., Sunoj, S., Kharel, D., Oware, E. K., Aardt, J. A. V., & Ketterings, Q. M. (2021). Spatial estimation methods for mapping corn silage and

grain yield monitor data. *Precision Agriculture*. https://doi.org/10.1007/s11119-021-09793-z

Cox, M., & Gerard, P. (2007). Soil management zone determination by yield stability analysis and classification. *Agronomy Journal*, *99*, 1357–1365. https://doi.org/10.2134/agronj2007.0041

De Lara, A., Khosla, R., & Longchamps, L. (2018). Characterizing spatial variability in soil water content for precision irrigation management. *Agronomy Journal, 8*(5), 1–8. https://doi.org/10.3390/agronomy8050059

Diker, K., Heermann, D., & Brodahl, M. (2004). Frequency analysis of yield for delineating yield response zones. *Precision Agriculture, 5*. 435–444. https://doi.org/10.1007/s11119-004-5318-9

Dobermann, A., & Ping, J. L. (2004). Geostatistical integration of yield monitor data and remote sensing improves yield maps. *Agronomy Journal*, *96*(1), 285–297. https://doi.org/10.2134/agronj2004.0285

Eghball, B., & Varvel, G. E. (1997). Fractal analysis of temporal yield variability of crop sequences: Implications for site-specific management. *Agronomy Journal, 89*(6), 851–855. https://doi.org/10.2134/agronj1997.00021962008900060001x

Fraisse, C. W., Sudduth, K. A., & Kitchen, N. R. (2001). Delineation of site-specific management zones by unsupervised classification of topographic attributes and soil electrical conductivity. *Transactions of the ASAE, 44*(1), 155–166. https://doi.org/10.13031/2013.2296

Gili, A., Álvarez, C., Bagnato, R., & Noellemeyer, E. (2017). Comparison of three methods for delineating management zones for site-specific crop management. *Computers and*

Electronics in Agriculture, *139*, 213–223. https://doi.org/10.1016/j.compag.2017.05.022

Guinness, J., & Katzfuss, M. (2019) *GpGp: Fast Gaussian process computation using Vecchia's approximation.* R package version 0.3.2. Retrieved April 21, 2021, from https://CRAN.R-project.org/package=GpGp

Hijmans, R. J. (2019). R*aster: geographic data analysis and modeling*. R package version 2.9-23. Retrieved April 21, 2021, from https://CRAN.R-project.org/package=raster

Hoffmann, M. P., Llewellyn, R. S., Davoren, C.W., & Whitbread, A.M. (2017). Assessing the potential for zone-specific management of cereals in low-rainfall south-eastern Australia: Combining on-farm results and simulation analysis. *Journal of Agronomy and Crop Science, 203*(1), 14–28. https://doi.org/10.1111/jac.12159

Hornung, A., Khosla, R., Reich, R., Inman, D., & Westfall, D. (2006). Comparison of site-specific management zones: Soil-color-based and yield-based. *Agronomy Journal, 98*(2), 405–417. https://doi.org/10.2134/agronj2005.0240

Srinivasan, A. (2006). *Handbook of Precision Agriculture: Principles and Applications*. Food Products Press.

Johnson, C. K., Mortensen, D. A., Wienhold, B. J., Shanahan, J. F., & Doran, J. W. (2003). Site-specific management zones based on soil electrical conductivity in a semiarid cropping system. *Agronomy Journal, 95*(2), 303–315. https://doi.org/10.2134/agronj2003.0303

Kharel, T. P., Marema, A., Czymmek, K. J., Oware, E. K., & Ketterings, Q. M. (2019). Combining spatial and temporal corn silage yield variability for management zone

development. *Agronomy Journal, 111*(6), 2703–2711. https://doi.org/10.2134/agronj2019.02.0079

Kharel, T. P., Swink, S. N., Maresma, A., Youngerman, C., Kharel D., Czymmek, K. J. & Ketterings Q. M. (2019b). Yield monitor data cleaning is essential for accurate corn grain/silage yield determination. *Agronomy Journal 111*(2), 509–516. doi:10.2134/agronj2018.05.0317.

Kharel, T., Swink, S. N., Youngerman, C., Maresma, A., Czymmek, K. J., Ketterings, Q. M., et al. (2018). *Processing/cleaning corn silage and grain yield monitor data for standardized yield maps across farms, fields, and years*. Cornell University, Nutrient Management Spear Program, Department of Animal Science, Ithaca, NY, USA. Retrieved April 21, 2021, from http://nmsp.cals.cornell.edu/publications/extension/ProtocolYieldMonitorDataProcessing1_15_2020.pdf

Khosla, R., Inman, D. J., Westfall, D. G., Reich, R. M., Frasier, M., Mzuku, M., Koch, B., & Hornung, A. (2008). A synthesis of multi-disciplinary research in precision agriculture: site-specific management zones in the semi-arid western Great Plains of the USA. *Precision Agriculture, 9*, 85–100. https://doi.org/10.1007/s11119-008-9057-1

Kim, S., Daughtry, C., Russ, A., Pedrera-Parrilla, A., & Pachepsky, Y. (2020). Analysis of Spatiotemporal Variability of Corn Yields Using Empirical Orthogonal Functions. *Water, 12*(12), 3339. https://doi.org/10.3390/w12123339

Kitchen, N. R., Sudduth, K. A., Myers, D. B., Drummond, S. T., & Hong, S. Y. (2005). Delineating productivity zones on claypan soil fields using apparent soil electrical

conductivity. *Computers and Electronics in Agriculture, 46*(1-3), 285–308. https://doi.org/10.1016/j.compag.2004.11.012

Kravchenko, A. N., Robertson, G.P., Thelen, K. D., & Harwood, R. R. (2005). Management, topographical, and weather effects on spatial variability of crop grain yields. *Agronomy Journal, 97*(2), 514–523. https://doi.org/10.2134/agronj2005.0514

Lamb, J. A., Dowdy, R. H., Anderson, J. L., Rehm, & G. W. (1997). Spatial and temporal stability of corn grain yields. *Journal of Production Agriculture*, *(10)*3, 410–414. https://doi.org/10.2134/jpa1997.0410

Leroux, C., Jones, H., Taylor, J., Clenet, A., & Tisseyre, B. (2018). A zone-based approach for processing and interpreting variability in multi-temporal yield data sets. *Computers and Electronics in Agriculture, 148*, 299–308. https://doi.org/10.1016/j.compag.2018.03.029

Long, E., & Ketterings Q. M. (2016). Factors of yield resilience under changing weather evidenced by a 14-years record of corn-hay yield in a 1000-cow dairy farm. *Agronomy for Sustainable Development*. DOI:10.1007/s13593-016-0349-y.

Maestrini, B., & Basso, B. (2018). Drivers of within-field spatial and temporal variability of crop yield across the US Midwest. *Scientific Reports, 8,* 106–112. https://doi.org/10.1038/s41598-018-32779-3

Mallory, E. B., & Porter, G. A. (2007). Potato yield stability under contrasting soil management strategies. *Agronomy Journal, 99*(2) 501–510. https://doi.org/10.2134/agronj2006.0105

Meisinger, J. J., Schepers, J. S., & Raun, W. R. (2008). Crop nitrogen requirement and fertilization. In: Schepers, J. S., & Raun, W. R. (eds). Nitrogen in Agricultural Systems. Agronomy Monograph 49. ASA, CSSA, SSSA, Madison, WI. pp 563–612.

Melo Damian, J., De Castro Pias, O. H., Santi, A. L., Di Virgilio, N., Berghetti, J., Barbanti, L., & Martelli, R. (2016). Delineating management zones for precision agriculture applications: a case study on wheat in sub-tropical Brazil. *Italian Journal of Agronomy, 11*(3), 171–179. https://doi.org/10.4081/ija.2016.713

Monzon, J. P., P. A. Calviño, V. O. Sadras, J. B. Zubiaurre, and F. H. Andrade (2018). Precision agriculture based on crop physiological principles improves whole-farm yield and profit: A case study. *European Journal of Agronomy, 99*, 62–71. https://doi.org/10.1016/j.eja.2018.06.011

Moshia, M. E., Khosla, R., Davis J. G., Westfall D. G., & Doesken, K. (2015). Precision manure management on site-specific management zones: Topsoil quality and environmental impact. *Communications in Soil Science and Plant Analysis, 46*(2), 235–258. https://doi.org/10.1080/00103624.2014.967862

Ping, J. L., & Dobermann, A. (2005). Processing of yield map data. *Precision Agriculture 6*, 193–212. https://doi.org/10.1007/s11119-005-1035-2

Porter, P. M., Lauer, J. G., Huggins, D. R., Oplinger, E. S., & Crookston, R. K. (1998). Assessing spatial and temporal variability of corn and soybean yields. *Journal of Production Agriculture, 11*(3), 359–363. https://doi.org/10.2134/jpa1998.0359

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Santesteban, L. G., Guillaume, S., Royo, J. B., & Tisseyre, B. (2013). Are precision agriculture tools and methods relevant at the whole-vineyard scale? *Precision Agriculture, 14*, 2–17. https://doi.org/10.1007/s11119-012-9268-3

Sawyer, J. E. (1994). Concepts of variable rate technology with considerations for fertilizer application. *Journal of Production Agriculture, 7*(2), 195–201. https://doi.org/10.2134/jpa1994.0195

Smith, R. G., Menalled, F. D., & Robertson, G. P. (2007). Temporal Yield Variability under Conventional and Alternative Management Systems. *Agronomy Journal, 99*(6), 1629–1634. https://doi.org/10.2134/agronj2007.0096

Sudduth, K. A., & Drummond, S. T. (2007). Yield Editor. *Agronomy Journal*, *99*(6), 1471–1482. https://doi.org/10.2134/agronj2006.0326

Sudduth, K. A., Drummond, S. T., & Myers, D. B. (2012). Yield Editor 2.0: Software for automated removal of yield map errors. In 2012 Dallas, Texas, July 29-August 1, 2012 (p. 1). American Society of Agricultural and Biological Engineers. Retrieved April 21, 2021, from https://doi.org/10.13031/2013.41893.

USDA. (2020). United States Department of Agriculture - Crop Production Historical Track Records. Retrieved April 21, 2021, from https://www.nass.usda.gov/Publications/Todays_Reports/reports/croptr20.pdf

Vega, A., Córdoba, M., Castro-Franco, M., & Balzarini M. (2019) Protocol for automating error removal from yield maps. *Precision Agriculture, 20,* 1033–1044. https://doi.org/10.1007/s11119-018-09632-8

CHAPTER 3:

PROPOSED METHOD FOR STATISTICAL ANALYISOS OF ON-FARM SINGLE STRIP TREATMENT TRIALS

ABSTRACT

On-farm experimentation (OFE) has allowed farmers to improve crop and land management, as well as increase productivity and profitability over time. One of the most prevalent OFE designs is the randomized complete blocks design (RCBD) with field length-strips as individual plots, thereby facilitating use of farm equipment for implementation and harvest. This design requires advanced planning of block location and space, and has limited statistical power when only three to four replications are implemented. Yield monitor systems that collect yield information at one second intervals generate yield data at within-plot levels, allowing for development of more meaningful, easily implementable OFE designs. Here we explored statistical frameworks to quantify the effect of a single treatment strip using georeferenced yield monitor data, as well yield stability-based management zones derived from prior year yield data from the farm. Nitrogen-rich single treatment strips per field were implemented in 2018 and 2019 on three fields each on two farms in central New York. We examined two approaches, namely Least Squares and Generalized Least Squares, for estimating treatment effects and two approaches, estimation assuming independence versus spatial covariance, for estimating standard errors. The analysis showed that estimation of treatment effects using the Generalized Least Squares approach are unstable due to over emphasis on certain data points, while assuming independence also led to underestimation of standard errors. In conclusion, when estimating the treatment effect, the Least Squares approach should be used,

while spatial covariance should be assumed when estimating standard errors for evaluation of zone-based treatment effects using the single-strip per field design.

INTRODUCTION

Applied agricultural research has been traditionally conducted in research stations with findings presented to farmers by extension staff or staff from other development organizations (Mutsaers, 1997). On-farm experiments (OFE), however, allow for more seamless transfer of research findings, because the research is conducted in an environment relevant to the farm in terms of soil types, management, weather, etc., often resulting in more adoptable and sustainable solutions for farmers (Mutsaers, 1997; Petersen 1994). In the past decade, OFE partnerships between farmers and industry or university researchers have expanded as the approach has been shown to improve farmers' crop and land management with increased productivity (Kyveryga et al., 2019). However, a valid experimental design is required to ensure the statistical validity of the OFE outcome.

Blocking, randomization, and replication are employed to minimize external variability not resulting from the treatment that is being evaluated (Federer, 1955; Little & Hills, 1978). The most prevalent research design for OFE is the randomized complete blocks design (RCBD) with field-length strips as individual plots, which facilitates the use of farm equipment for implementation and harvest (Kyveryga et al., 2018; Piepho et al. 2011). In this design, field-length strips are experimental units (EUs), the smallest entities to which a treatment can be randomly assigned (Gotway Crawford et al. 1997). Multiple replications of the treatment and control strips are then placed within a field, effective blocking treatment pairs to minimize

variability within a block (Fisher, 1926; Kyveryga et al., 2018) (Figure 3.1). Analysis of variance (ANOVA) traditionally has been used to test the statistical significance of the treatment.

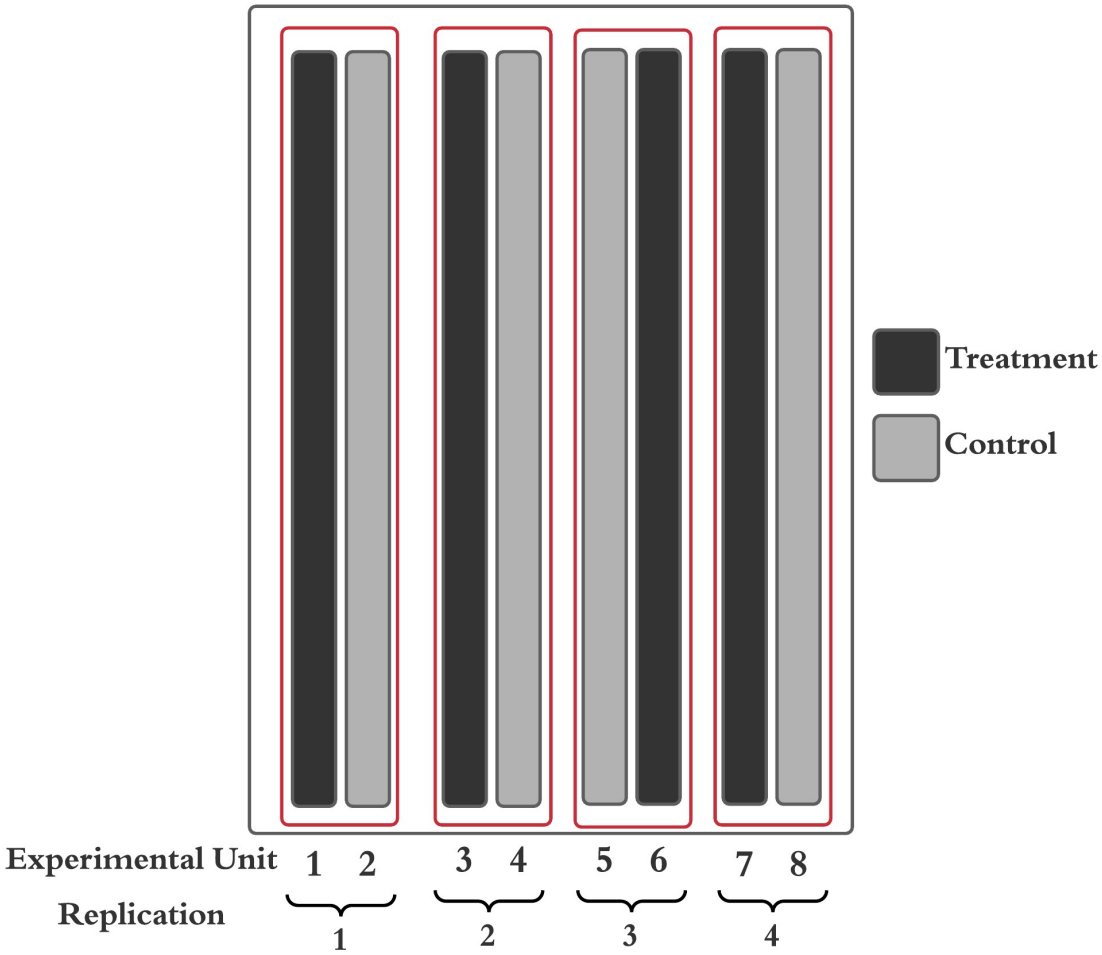

Figure 3.1: An example of a randomized compete block design (RCBD) to evaluate if a treatment impacted outcomes, such as yield. In this example there is one control and one treatment strip per block (replication) and the trial is replicated three times. When farm equipment permits, strips serve as experimental units (EUs), which typically cover the length of a field (Kyveryga et al., 2015).

The RCBD with field strips as EUs, however, poses challenges for both farmers and scientists. For farmers, it requires planning and when yield from individual strips needs to be collected, it often hampers farm operations during planting and harvest, the busiest and most

labor intensive time of the year, discouraging farmers to conduct OFE (Griffin et al., 2004; Rudolph et al., 2016). The design and its analysis pose challenges for scientists as well, because fields may only allow for three to four blocks to be implemented, which gives limited statistical power if each strip is considered as being one EU (Piepho et al., 2011).

The arrival and more widespread adoption of yield monitoring systems now allow for yield data collection at a much higher spatial resolution (collection of data within strips). This allows for documentation of spatial variability caused by a variety of different reasons, including non-uniform distribution of soil properties, soil moisture, pest pressure, rooting depth, and other factors (Sawyer, 1994). Recent studies of corn (*Zea mays* L.) fields in New York have shown large variability in yield within fields (Kharel et al. 2019a; Cho et al. 2021). The existence of such variability is well recognized by both farmers and scientists (Fisher 1931). However, until the arrival of yield monitoring systems, the only way to deal with such variability in OFE was to conduct small plot research, where heterogeneity within fields is assumed to be small (Fisher, 1926; Oehlert, 2010), or on a carefully chosen field with the least amount of spatial variability, based on farmers' past experiences (Gomez & Gomez, 1984). Results from such trials cannot be extended to other fields, unless they have the same underlying conditions as the trial field (Gomez & Gomez, 1984). In addition, it is quite possible for spatial variability within a field–even when selecting smaller experimental units–to mask any treatment effects being tested in OFE (Panten et al., 2010; Kim et al. 2020).

A demonstration plot, where a field is split into either a treatment or control, is considerably easier to implement for farmers than a replicated trial with multiple strips in a field (Griffin et al., 2006; Lambert et al. 2004; Willers et al. 2008; Yan et al., 2002). This demonstration plot then results in response values (such as yield) for both the treatment and the

control (rest of the field), which is especially easy to establish when yield monitor data are used to average yield for the treated area versus the untreated control portion of the field (Griffin et al., 2006; Piepho et al., 2011). However, no statistical analyses can be performed with only two EUs, and the validity of any inferences based on demonstration plots is highly debated. While many (Cox, 2009; Piepho et al., 2011; Kyveryga et al., 2018; Johnson, 2006) argue that concepts such as randomization, replications, and blocking should not be ignored for appropriate statistical analysis, some argue that given the limited resources of farmers, demonstration plots still provide useful information for decision making (Mutsaers, 1997; Yan et al. 2002; Stroup et al. 1993).

Given the ease of implementation of single strip EUs, a couple of studies have compared different statistical models for analyzing single strip evaluations, while taking into account the intensity of yield data when collected with a yield monitor system (many data points per strip, rather than just one value for the entire strip). Rudolph et al. (2016) compared wheat (Triticum aestivum L.) yield in England using three strips varying in nitrogen (N) application (low, standard and high). The entire field received a standard amount of N, while two strips received either 60 kg N ha^{-1} more (high) or 60 kg N ha^{-1} less (low) than the standard amount. The study compared outputs from statistical frameworks with and without taking into account spatial correlation, concluding that standard errors on the treatment effects were underestimated when spatial correlation is not considered. Lawes and Bramley (2012) worked with farmers in Australia who wanted to analyze the effect of more versus less fertilizer for canola (*Brassicus napus* L.) and Barley (*Hordeum vulgare* L.). For two fields, two management zones (low yielding vs high yielding) were arbitrarily classified based on the farmer's intuition and past experience related to the field. For the 3rd field, three management zones were delineated, based on relative yield and electrical conductivity. ANOVA and ANOVA with spatial covariance, i.e.,

Spatial ANOVA, were compared for quantifying the effect of fertilizer per each zone. Contrary to Rudolph et al. (2016), Lawes & Bramley (2012) conclude that complex spatial analysis was not needed, as their approach produced similar results to the traditional statistical analysis.

While Lawes & Bramley accounted for spatial variability of yield within each field using zones, they did not account for temporal yield variability, an important factor for understanding results of on-farm experimentation (Kharel et al., 2019a). Temporal yield variability–heterogeneity in yield across multiple years at the field level–can be caused by a variety of factors, including weather, management, and topography (Andresen et al., 2001; Kravchenko et al., 2005; Mallory & Porter, 2007; Smith et al., 2007). With availability of yield data from past years, both spatial and temporal yield variability over time can be assessed (Kharel et al., 2019a; Cho et al., 2021). The main challenge remains whether we can design a statistically sound approach to treatment evaluation on farms, by farmers, given access to high resolution yield data, as well as quantification of both spatial and temporal yield variability.

Here we propose a statistical framework for estimating the treatment effect and standard error of the treatment effect (treatment versus control) based on yield monitor data from a single strip, on-farm, research trial, taking into account both spatial and temporal variability in yield to identify treatment signals, i.e., to reduce variability. It is important to note that we are differentiating the estimation of treatment effects and estimation of standard errors. For estimation of treatment effects, we will explore two frameworks, namely Least Squares (LS) and Generalized Least Squares with spatial covariance (GLS), while for estimation of standard errors, we will explore two frameworks, that of assuming independence (Independence) and assuming spatial covariance (Spatial).

MATERIALS AND METHODS

Yield Monitor Datasets

One dairy farm and one cash grain operation in Central New York participated in this study. Single-strip N treatment trials were conducted in three site-years, two sites in 2018 and one site in 2019, per farm (Table 3.1). Corn was harvested for grain by the cash grain operation and for silage by the dairy farm. Yield monitor data on the 6 site-years as well as the historic yield records of the farms (seven years of corn silage yield records for the dairy farm and 5 years of corn grain yield records for the cash grain operation) were used in the analysis. All yield data were collected using John Deere 3 (John Deere, Moline, IL, USA) systems.

Table 3.1: Crop type, year, field size, average yield production, information about the nitrogen strips, soil types, and distribution of management zones, delineated as suggested by Kharel et al. (2019a), were summarized.

	Unit	Grain Operation			Dairy Farm		
Field name		Field A	Field B	Field C	Field D	Field E	Field F
Crop type		Grain	Grain	Grain	Silage	Silage	Silage
Year		2018	2018	2019	2018	2018	2019
Field size	Ha	2.59	2.02	2.99	22.74	42.41	33.63
Average yield	Mg/ha	13.07	11.99	12.84	40.80	39.90	54.47
Nitrogen strip							
Source		UAN	UAN	UAN	Urea	Urea	Urea
Method		Injected	Injected	Injected	Broadcast	Broadcast	Broadcast
Rate	kg/ha	56	56	56	121	121	121
Width	m	9	9	9	24	24	24
Soil Type							
Most common		Honeoye (Fine-loamy, mixed, semiactive, mesic Glossic Hapludalfs)	Honeoye (Fine-loamy, mixed, semiactive, mesic Glossic Hapludalfs)	Lima (Fine-loamy, mixed, semiactive, mesic Oxyaquic Hapludalfs)	Ontario (Fine-loamy, mixed, active, mesic Glossic Hapludalfs)	Ovid (Fine-loamy, mixed, active, mesic Aeric Endo-aqualfs)	Honeoye (Fine-loamy, mixed, semiactive, mesic Glossic Hapludalfs)

Unit	Grain Operation			Dairy Farm		
Second most common	Lima (Fine-loamy, mixed, semiactive, mesic Oxyaquic Hapludalfs)	Lima (Fine-loamy, mixed, semiactive, mesic Oxyaquic Hapludalfs)	Kendaia (Fine-loamy, mixed, semiactive, nonacid, mesic Aeric Endo-aquepts)	Benson (Loamy-skeletal, mixed, active, mesic Lithic Eutrudepts)	Cazenovia (Fine-loamy, mixed, active, mesic Glossoboric Hapludalfs)	Ontario (Fine-loamy, mixed, active, mesic Glossic Hapludalfs)
Distribution of Management Zone						
Zone 1	76.09	35.99	78.12	11.46	24.82	28.56
Zone 2	23.91	0.00	0.00	88.54	38.78	0.00
Zone 3	0.00	30.81	21.88	0.00	6.99	5.76
Zone 4	0.00	33.20	0.00	0.00	29.41	65.68

Raw yield monitor data were cleaned to eliminate systematic and random errors (Dobermann & Ping, 2004; Vega et al., 2019; Kharel et al., 2019b). Pass overlap (driving over areas already harvested), yield estimates due to harvesting equipment slowing down or speeding up, and inconsistencies between sensor delays (flow and moisture delays), and the harvester location are examples of issues that contribute to errors in the raw yield monitor data (Kharel et al., 2019b). Data were transformed in AgLeader format and exported in .CSV format using SMS Advanced Software (Ag Leader Technology, Ames, IA, USA). Yield Editor (Sudduth et al., 2012; Sudduth & Drummond, 2007) was used for post-harvest data cleaning, as described in Kharel et al. (2018) and based on Kharel et al. (2019b). The yield cleaning protocol was used for all yield data, including seven years of data for the dairy farm and five years for the cash grain operation.

Zone Delineation

Yield stability-based management zones, described in Kharel et al. (2019a), were delineated for each field with data through the year prior to strip implementation. Cleaned historic yield monitor data were interpolated using kriging with the Matérn covariance function with 2 x 2 m resolution (Cho et al., 2021). Temporal average yield and standard deviation of yield were determined at the farm-level for farm-specific zone delineation (Kharel et al., 2019a). The multi-year maps of the fields where N strips would be applied were compared against the farm-level average yield and the standard deviation to classify yield pixels (2 x 2 m) as high-yielding or low-yielding (above or below the farm average) and stable or unstable (below or above the farm-level temporal standard deviation). Pixels in zones 1 and 4 represented stable yielding areas with high or low yields, respectively. Yields in zones 2 and 3 were variable over time with, on average, yield above the farm average in zone 2 and below average in zone 3 (Table 3.1). Along with yield at the trial year, geolocation provided by the yield monitor data and the N application, data points were assigned into four zones based on the past yield.

Single Strip Trial Design

For each site-year, a field-length single strip was placed where extra N was applied at or before planting. The placement of the single strip was chosen to intersect with a minimum of two zones per field. The width of the strip was 9 m for the fields at the cash grain operation (Fields A, B, and C) and 24 m for the fields at the dairy farm (Field D, E, and F). At the grain operation, the extra N was applied at a rate of 56 kg/ha as urea ammonium nitrate (UAN). At the dairy farm, the application rate was 121 kg/ha, broadcasted in the form of urea (Table 3.1). For statistical analyses, two field-length strips located 10 m on the left and the right side of the N

strip were identified as control strips (Figure 3.2). Strips were of field length and at least two harvester widths wide (9 m for Field A, B, and C; and 24 m for Field D, E, and F). Fields were managed uniformly by the farmers once the N strips were implemented.

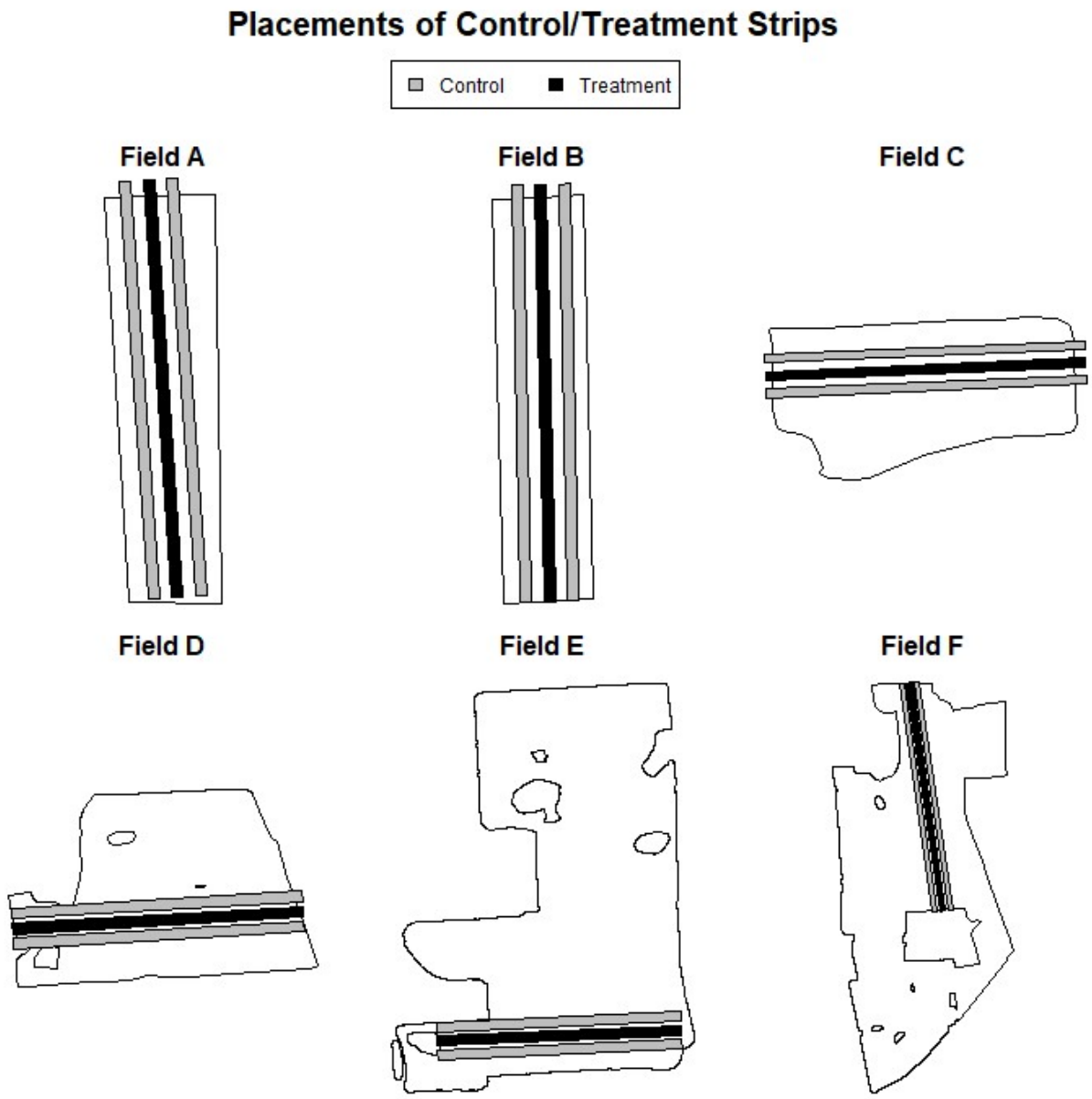

Figure 3.2: Placements of N treatment strips for six fields (black) surrounded on both sides by control strips (grey) for six fields on a cash grain operation (Fields A, B, C) and a dairy farm (Fields D, E, F).

Statistical Modelling

Two approaches were explored for estimating the treatment effects: the Least Squares estimation and the Generalized Least Squares estimation with spatial covariance. The Least Squares estimation and the Generalized Least Squares estimation of the treatment effects can be solved using the following formula:

$$Least\ Squares\ Estimation: (X^T X)^{-1} X^T y \tag{3.1}$$

$$Generalized\ Least\ Squares\ Estimation: (X^T \Sigma^{-1}\ X)^{-1} X^T \Sigma^{-1} y \tag{3.2}$$

where X is the design matrix containing data of the independent variables, y is a vector of the response, and Σ is the covariance matrix modeling the spatial dependence structure among data points. The Matérn covariance function was used (Cho et al. 2021). The parameters for the covariance function was estimated through maximum likelihood estimation using the GpGp package (Katzfuss & Guinness, 2019).

$$M_{LS} = (X^T X)^{-1} X^T \tag{3.3}$$

$$\mathrm{M_{GLS}} = (X^T \Sigma^{-1}\ X)^{-1} X^T \Sigma^{-1} \tag{3.4}$$

Both M_{LS} and $\mathrm{M_{GLS}}$ are P x N matrices where P is the number of explanatory variables and N is the number of data points. Multiplying M_{Ls} or $\mathrm{M_{GLS}}$ by the response vector, *y,* results in a vector of the beta coefficients with M elements. Standard errors, the level of uncertainties around treatment effects estimation, can be calculated by using the following general formula:

$$Standard\ Errors: \sqrt{MKM^T} \tag{3.5}$$

where M is either M_{LS} or M_{GLS}, depending on the treatment estimation approach, and K is some covariance matrix. Two approaches were explored for estimating the standard errors, namely assuming independence and assuming spatial covariance. If we assume independence, the covariance matrix K would be $\hat{\sigma}^2 I$, where $\hat{\sigma}^2 = \frac{1}{N}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2$, Y_i is the response, $\hat{Y}_i$ is the

prediction by the model, N is the number of data points and I is the identity matrix. If we take into account spatial covariance, K would be Σ as described above.

The estimation of treatment effects and standard errors often is treated as a single problem instead of two independent ones. In statistics, Least Squares estimation refers to the estimation of treatment effect via (3.1) and estimation of standard errors assuming independence, which is $\sqrt{\hat{\sigma}^2 M_{LS} I M_{LS}^T}$. In this study, however, we treated estimation of treatment effects and standard errors as two separate problems to closely analyze each approach. Four possible combinations of the treatment effects and standard errors estimations exist: (a) the Least Squares estimation with independence assumption for standard errors, (b) the Least Squares estimation with spatial covariance for standard errors, (c) the Generalized Least Squares estimation with independence assumption for standard errors, and (d) the Generalized Least Squares estimation with spatial covariance for standard errors. Approach (a) is referred to as LS 1, (b) is referred to as LS 2, and (d) is referred to as GLS from on. Of the four approaches, (c) was excluded from the analysis as it was deemed inferior to three other approaches, as explained under Results and Discussion, Model Diagnostics.

Explanatory Variables Selection & Model Fitting

It is reasonable to believe that yield data measured with a yield monitor system are both spatially and temporally correlated. Yield estimates that are located closer together tend to have similar values when compared to points that are further apart. A location that is historically high yielding over the past years and has consistently been that way, thus, is likely to exhibit high yield in the following years. Exploratory data analysis on the spatial and temporal yield

distribution within the field was performed to appropriately control for such spatial autocorrelation when estimating the effect of N treatment.

To account for intra-field spatial yield variation, relative latitude and longitude and their linear interaction were treated as fixed effects. Relative latitude and longitude were calculated by subtracting the average value of each, i.e., the goal was to anonymize the location of the field. By adding these factors as fixed effects, we can possibly control for linear yield trend across the field due to topographic factors within the field. Spatial yield variation was also considered under the statistical modeling. LS 2 takes into account spatial covariance structure, Σ, when estimating standard errors while GLS use it when estimating both treatment effects and standard errors estimation. On the other hand, LS 1 assumes that each data point is independent when estimating both the treatment effects and the standard errors.

Temporal yield variability, or heterogeneity in yield across multiple years, also needed to be accounted for in order to accurately examine the effect of the treatment and its standard error. Temporal average yield, estimated by averaging over the past years at the given location, was treated as a fixed effect to control for the historic yield level of the field. In addition, management zones (Kharel et al. 2019a) and its interactions with N treatment were treated as fixed effects to control for temporal yield level and its variation at the farm level. The interaction term between four management zones and N treatment was added because yield response to N was expected to be differ among management zones. Thus, the following model was fitted via three approaches as mentioned above (LS 1, LS 2, and GLS), using the R base package (R Core Team, 2019) and GpGp (Guinness & Katzfuss, 2019):

$$Yield \sim Zone + Nitrogen{:}Zone + AverageYield + Latitude * Longitude \qquad (3.6)$$

where *Yield* is the response variable representing the yield at the site-year during the trial, *Zone* is the management zone (Zone 1, 2, 3, or 4), *Nitrogen* is the binary variable representing whether the data point was either located in the N strip (Nitrogen = 1) or one of the two control strips (Nitrogen = 0), *AverageYield* is the temporal average yield, calculated based on the past yield data at the given location, and *Latitude* and *Longitude* are the normalized geolocation of the data points. A design matrix, *X*, and the response vector, *y*, was formed based on the formula (3.6) and were used to estimate the treatment effects and their standard errors (see Statistical Modelling under Materials and Methods). Treatment effects and standard errors estimated per LS 1, LS 2, and GLS were first compared.

Model Output and Diagnostics

As noted, treatment effects for extra N per management zone were estimated using two approaches: the Least Squares approach, (3.1) and the Generalized Least Squares approach (3.2). The treatment effects in both cases can be represented via a linear combination of the responses. The linear combination depends on the design matrix, *X,* thus, M_{LS}, (3.3) for the Least Squares approach and $\mathrm{M_{GLS}}$, (3.4), for the Generalized Least Squares approach. Because each row of M_{LS} and M_{GLS} contains coefficients of the linear combination of the responses, observing M_{Ls} and $\mathrm{M_{GLS}}$ allows us to analyze how each data points contribute to the estimation of the treatment effects. Thus, the spatial distributions of M_{Ls} and $\mathrm{M_{GLS}}$ were analyzed to determine the better approach in estimating the treatment effects of extra nitrogen per management zone.

The impact of the number of observations on the estimation of treatment effects and standard errors also was analyzed. The number of data points should not significantly affect treatment effects nor standard errors, given a sufficient number of data points to start with. In this

analysis, we randomly selected 60%, 70%, 80% and 90% of the data for a site-year. The treatment effects using the two approaches (Least Squares and Generalized Least Squares), and the standard errors using the two approaches (assuming independence and spatial covariance) were estimated for each subset. This process was repeated 100 times to generate distributions which were analyzed in detail to determine the appropriateness of each approach.

RESULTS AND DISCUSSION

Yield in Control and Nitrogen Strips

For the grain fields, the average yield in the treatment strip was 0.22 and 0.26 Mg/ha higher for Field B and C, respectively, while for Field A the average yield in the treatment strip was 0.12 Mg/ha lower. For the silage fields, the treatment strip was higher yielding on average for field D and E (2.38 and 5.20 Mg/ha higher, respectively), while for Field F, the control strips exhibited higher yields than the control strips by 2.32 Mg/ha (Figure 3.3). The overall yield distributions were similar across fields, with a standard deviation between 0.5-1 Mg/ha for grain fields (Field A, B, and C) and between 10-13 Mg/ha for the silage fields (Field D, E, and F).

The differences in yield between the N strips and their controls varied greatly among zones (Figure 3.4). For example, while on average the control strips were higher yielding than the N treatment for Field A, analysis per zone showed a positive response of 0.25 Mg/ha for Zone 2, while in Zone 1, yield in the N strip was 0.25 Mg/ha lower than in the control. Similarly, for Field C, the yield in the N strip was higher across the field, primarily because of a large yield increase in Zone 3 (1.66 Mg/ha), while average yield in Zone 1 was 0.11 Mg/ha lower than in the control. For Field B, D, E, and F, the difference in average yield between treatment and control per zone was consistent with differences measured at the field level. For Field B, D, and E,

where the treatment strips were higher-yielding than the control strips on average, the yield from treatment strips were higher-yielding than the control strips per zone as well. Field F exhibited lower yield in the N strip, consistent for all zones in the field. Conclusion as to whether these differences are statistically significant were analyzed through various statistical frameworks in the following sections.

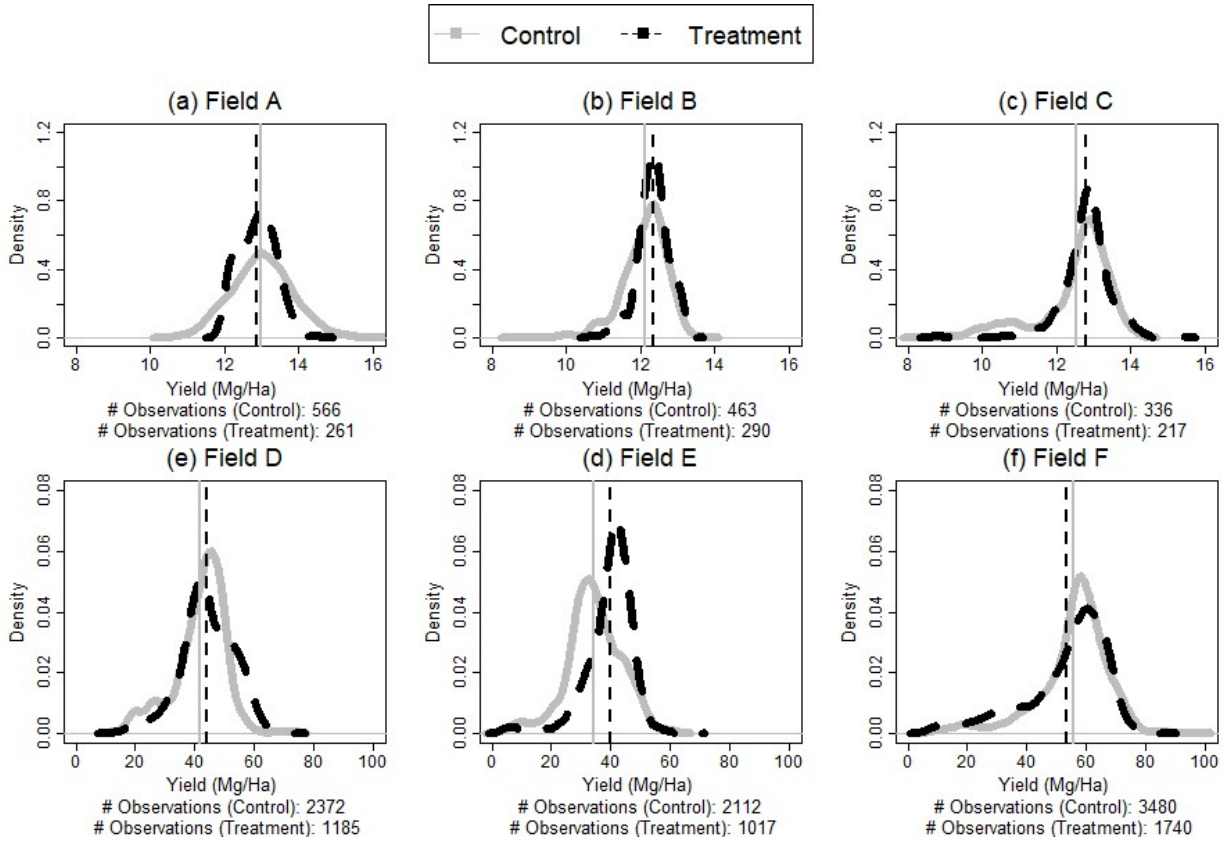

Figure 3.3: Density plot representing yield distributions of a N treatment strip (higher N) and the neighboring control strips. The dotted black and the solid grey vertical lines represent the average yield for the N treatment strip and the control strips, respectively.

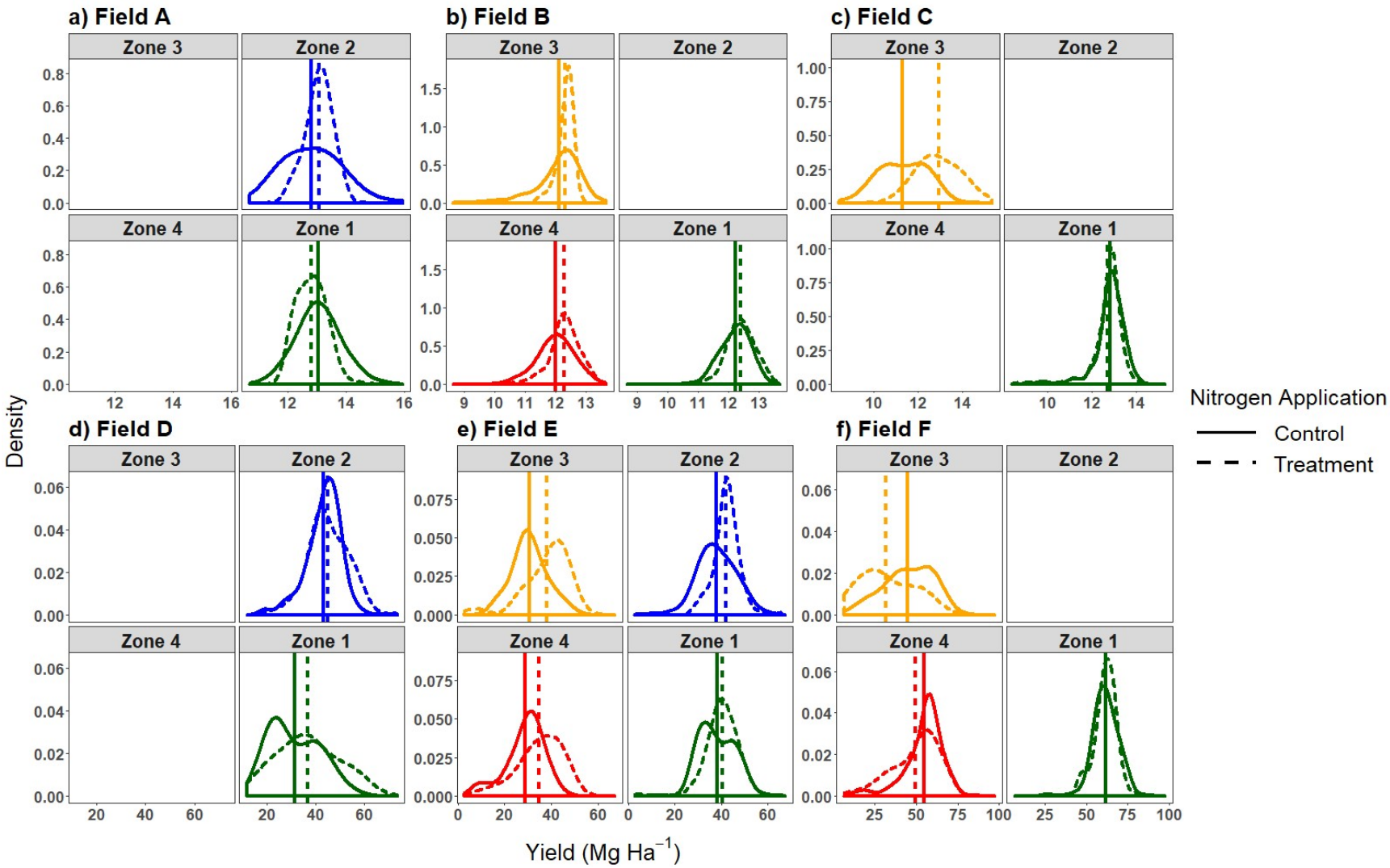

Figure 3.4: Density plots representing corn yield distribution of N treatment and control strips in individual fields with up to four yield stability zones, as per Kharel et al. (2019). The dotted black and the solid grey vertical line represent the average yield for the treatment and control strip, respectively.

Spatio-temporal Yield Variation

Yield, as well as spatial distribution of yield, varied greatly within each field (Figure 3.5). The standard deviation of yield, a measure of spatial yield variation, was between 0.5-0.8 Mg/ha for the Grain fields (Field A, B, and C) and 8-12 Mg/ha for the silage fields (Field D, E, and F). As expected, yields were spatially correlated, since data points that are spatially closer together have similar values. However, some data points had much higher or lower yield estimates than neighboring points. Presence of spatial autocorrelation among data points strongly suggested that the Least Squares approach for estimating the treatment effect and independence assumption for estimating the standard errors are inadequate.

Linear relationships between spatial coordinates and yield were present in Field A. Data points on the left side of the field tended to be higher-yielding than the data points on the right side of the field (Figure 3.5a). Because of a risk of such within-field trends, spatial coordinates and its linear interaction were added as fixed effects in the overall statistical model.

Spatial distribution of temporal average yields, calculated using the yield data prior to the trial, showed similar yield distribution when compared against the spatial distribution of yield during the trial year (Figure 3.5a and 3.5b). For example, the middle strip in Field C was higher-yielding than the rest of the field, both historically (Figure 3.5b) and in the trial year. It is important to note that the middle strip was the location where the extra nitrogen was applied during the trial year. Temporal average yield was included as fixed effects in the overall statistical model, in order to control for the historic level of production at the field level.

Within the same farm, field levels differed (Figure 5b). For silage fields, Field A and C were high-yielding, while Field B had a wide range in yield within the field. For grain fields, Field D was high-yielding, Field E had yields similar to the average yield production of the farm,

and Field F was lower-yielding. Yield temporal variation was also noticeably different, especially for grain fields (Figure 5c). For all 3 silage fields, temporal yield variation was generally low across the field. For grain fields, Field D had high temporal yield variation, while Field F had low yield variation. The temporal average yield and its variation at the farm-level were captured by the four yield-stability based management zones, as described by Kharel et al. 2019 (Figure 5d). Such management zones can be used to account for both the average level of production and its variation across time at the farm-level. Thus, management zones, as well as its interaction with N treatment, were added as fixed effects in the overall statistical model.

Model Outputs

Treatment effects per zone for each field were noticeably different, depending on the method used for the estimation (LS 1 & 2 vs GLS; Table 3.2). For example, the N effect for zone 1 for Field A via LS was -0.23 Mg/ha, meaning that addition of N decreased yield by 0.23 Mg/ha for Zone 1. However, it was 0.45 Mg/ha when GLS approach was used, suggesting higher yield for the N strip. The treatment effects of LS 1 and LS 2 were both estimated using the Least Squares approach, thus resulting with the same outputs. Standard errors estimated assuming independence (LS) and spatial correlation (LS 2 and GLS) also a showed noticeable difference. For all fields, standard errors assuming spatial covariance (LS 2 and GLS) were greater than those estimated assuming independence (LS). The outputs show that standard errors estimated with independence assumption are smaller than standard errors generated while taking spatial covariance into account. Assuming spatial covariance (LS 2 and GLS) provided more conservative estimation of standard errors, as yields are expected to exhibit autocorrelation.

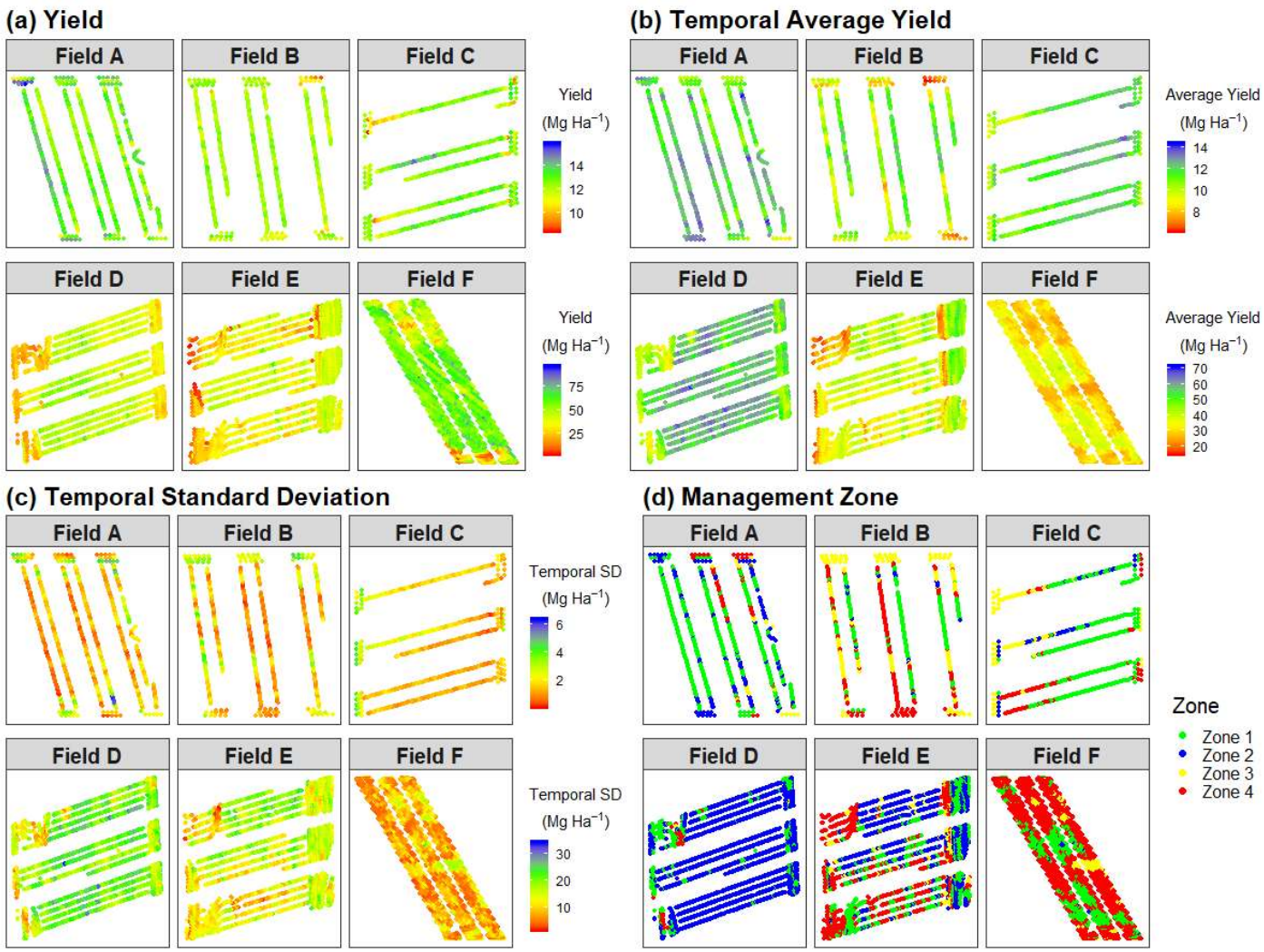

Figure 3.5: Spatial distribution of yield (Figure 3.5a), historic average yield (Figure 3.5b), temporal standard deviation (Figure 3.5c) and management zones (Figure 3.5d), as estimated based on previous years of yield produced for six fields (A through F). Yield stability management zones were delineated as described in Kharel et al. (2019). Zone 1 represents high-yielding-stable, Zone 2 represents high-yielding-unstable, Zone 3 represents low-yielding-unstable, and Zone 4 represents low-yielding-stable areas of the field.

Table 3.2: A regression model with the yield as the response variable, temporal average yield, normalized latitude, longitude, linear interaction between latitude and longitude, four corn management zones delineated as suggested by Kharel et al. (2019) and their interaction with N treatment as the explanatory variables were summarized for six fields (Field A, B, C, D, E, and F). For LS 1, the Least Squares was used for treatment effects and independence was assumed for standard errors. For LS 2, the Least Squares was used for treatment effects and spatial covariance was used for standard errors. For GLS, the Generalized Least Squares was used for treatment effects and spatial covariance was used for standard errors.

Field/Model		Intercept		Average Yield		Latitude		Longitude		Latitude:Longitude		Zone 2		Zone 3		Zone 4		Zone 1:N		Zone 2:N		Zone 3:N		Zone 4:N	
		Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
Field A	LS 1	14.95	0.50	-0.15	0.04	0.00	0.00	-0.01	0.00	0.00	0.00	-0.15	0.07	-	-	-	-	-0.23	0.06	0.16	0.11	-	-	-	-
	LS 2	14.95	1.85	-0.15	0.15	0.00	0.00	-0.01	0.01	0.00	0.00	-0.15	0.27	-	-	-	-	-0.23	0.24	0.16	0.34	-	-	-	-
	GLS	14.08	0.73	-0.05	0.05	0.00	0.00	-0.03	0.01	0.00	0.00	-0.10	0.05	-	-	-	-	0.45	0.16	0.52	0.18	-	-	-	-
Field B	LS 1	10.02	0.30	0.19	0.03	0.00	0.00	-0.01	0.00	0.00	0.00	-	-	0.30	0.08	0.01	0.07	0.20	0.06	-	-	0.36	0.09	0.38	0.07
	LS 2	10.02	0.78	0.19	0.07	0.00	0.00	-0.01	0.01	0.00	0.00	-	-	0.30	0.22	0.01	0.17	0.20	0.18	-	-	0.36	0.26	0.38	0.17
	GLS	11.17	0.36	0.08	0.03	0.00	0.00	-0.01	0.01	0.00	0.00	-	-	-0.05	0.07	-0.03	0.06	0.14	0.13	-	-	0.35	0.16	0.24	0.13
Field C	LS 1	6.65	0.84	0.51	0.07	-0.02	0.00	0.00	0.00	0.00	0.00	-	-	-0.54	0.18	-	-	-0.16	0.07	-	-	1.13	0.14	-	-
	LS 2	6.65	2.69	0.51	0.23	-0.02	0.01	0.00	0.00	0.00	0.00	-	-	-0.54	0.63	-	-	-0.16	0.24	-	-	1.13	0.41	-	-
	GLS	8.77	1.10	0.32	0.09	-0.02	0.01	0.00	0.00	0.00	0.00	-	-	-0.21	0.20	-	-	0.00	0.18	-	-	0.93	0.24	-	-
Field D	LS 1	-6.65	1.04	0.86	0.02	-0.04	0.00	0.00	0.00	0.00	0.00	0.81	0.52	-	-	-	-	0.38	0.74	1.62	0.25	-	-	-	-
	LS 2	-6.65	10.19	0.86	0.19	-0.04	0.05	0.00	0.01	0.00	0.00	0.81	2.26	-	-	-	-	0.38	3.31	1.62	1.78	-	-	-	-
	GLS	35.04	3.57	0.08	0.03	-0.07	0.05	0.01	0.01	0.00	0.00	0.87	0.29	-	-	-	-	0.53	1.05	0.07	0.97	-	-	-	-
Field E	LS 1	13.77	1.35	0.59	0.03	-0.02	0.01	0.01	0.00	0.00	0.00	-0.99	0.46	-2.10	0.78	-1.88	0.54	3.48	0.56	4.12	0.44	7.51	1.01	4.70	0.56
	LS 2	13.77	6.33	0.59	0.16	-0.02	0.04	0.01	0.01	0.00	0.00	-0.99	1.84	-2.10	2.00	-1.88	2.01	3.48	2.22	4.12	2.37	7.51	2.98	4.70	2.36
	GLS	28.09	2.43	0.20	0.04	0.05	0.04	0.02	0.01	0.00	0.00	0.33	0.33	-0.15	0.50	-0.46	0.39	1.01	1.22	1.00	1.22	1.89	1.32	1.68	1.24
Field F	LS 1	-4.54	2.06	1.86	0.05	0.01	0.00	0.07	0.00	0.00	0.00	-	-	3.38	1.04	4.16	0.53	-3.29	0.52	-	-	-9.70	1.35	-4.46	0.37
	LS 2	-4.54	18.32	1.86	0.54	0.01	0.02	0.07	0.07	0.00	0.00	-	-	3.38	6.35	4.16	3.32	-3.29	3.14	-	-	-9.70	6.54	-4.46	2.58
	GLS	34.41	9.17	0.32	0.06	0.04	0.02	0.13	0.08	0.00	0.00	-	-	-0.91	0.57	0.26	0.27	-0.42	1.14	-	-	0.18	1.43	-0.57	1.10

Model Diagnostics

Compared to coefficients for the Generalized Least Squares approach, $(X^T\Sigma^{-1}X)^{-1}X^T\Sigma^{-1}$, the coefficients for the Least Squares approach, $(X^TX)^{-1}X^T$, showed relatively uniform values across the field when compared to coefficients for the Generalized Least Squares approach (Figure 3.6). The coefficient for the Generalized Least Squares approach presented unusual outliers across the field. These outliers were attributed to sharp changes between explanatory variables that are spatially close in proximity. Unlike the Least Squares approach, which puts equal weight across data points due to the independence assumption, GLS puts more weight on two data points that are close in proximity, but have a drastic difference in their explanatory variables (i.e., management zones). As we observed in Figure 3.5, there are many data points that are close in distance, but are different management zones.

The problem of estimating the treatment effect with a categorical explanatory variable, while also accounting for spatial covariance structure, was noticed by Griffin et al. (2006). The author called it the "neighboring observation problem" and pointed out that observations that are located on the edge of different treatments influence the power in estimating the treatment differences. To bypass this problem, the Least Squares approach is proposed for estimating the treatment effects per management zone.

The number of observations included in the analyses (a proxy for size of the field and the treatment strip) affected the estimation of both the treatment effects and the standard errors. For both treatment effects and standard errors, the estimation became less variable, as suggested by the narrower shape of distributions as the number of data points increase (Figure 3.6). The effect of number of data points on the estimation of treatment effects varied between the Least Squares and Generalized Least Squares approach. The center of the distribution (mean or median) stayed

the same, regardless of the number of data points for the Least Squares approach, but this was not the case for GLS (Figure 9a, c, and d). The distribution varied depending on how many data points were included for the Generalized Least Squares approach. This is likely due to outliers in coefficients, as noted in Figure 3.6. Inclusion of certain data points would heavily influence the estimation of the treatment effects for the GLS approach. Based on these observations, the Least Squares estimation for the treatment effect is more appropriate when compared against the Generalized Least Squares estimation, as LS is more robust in its estimation, regardless of the number of observations.

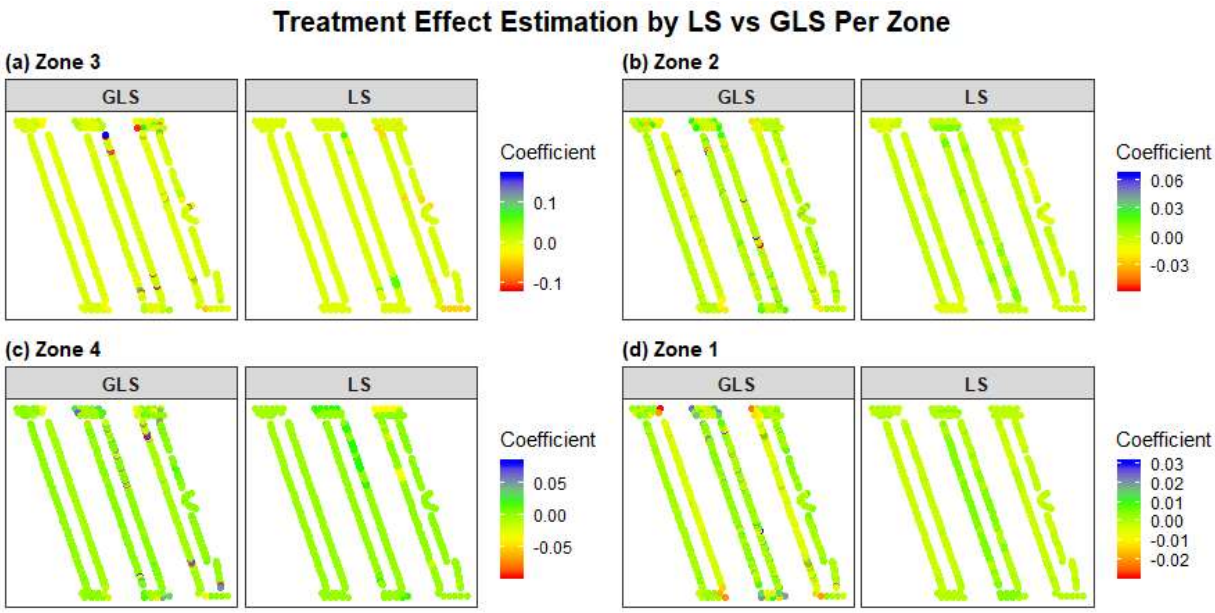

Figure 3.6: Spatial distribution of coefficients for estimating treatment effect of N treatment per yield-stability based management zone (Zone 1, 2, 3, and 4) via Least Squares method (LS) and Generalized Least Squares (GLS) method for Field A. Coefficients for LS were calculated by solving $(X^T X)^{-1} X^T$ and for GLS by solving $(X^T \Sigma^{-1} X)^{-1} X^T \Sigma^{-1}$, where X represents the design matrix and Σ the spatial covariance.

The number of observations also affected the estimation of standard errors. As noted in Table 3.2, standard errors assuming spatial covariance were higher, therefore, were more

conservative than the standard errors estimated assuming independence. In both cases, increasing the number of data points caused the distributions to become narrower and led to smaller estimation of the standard errors (Figure 3.7). However, the degree to which the standard errors decrease was more drastic when independence was assumed. The number of data points used did not affect the estimation of standard errors as much when spatial covariance was used. Increasing the number of data points should not affect the estimation of the standard errors once a sufficient number of data points exist. Based on these observations, accounting for spatial covariance is essential when estimating standard errors.

Rudolph et al. (2016) compared outputs from statistical frameworks with and without taking into account spatial correlation and concluded that standard errors on the treatment effects were underestimated when spatial correlation is not considered; our results corroborated this finding. For all 6 fields, standard errors estimated with spatial covariance were distinctively larger than standard errors estimated assuming independence. However, our results contradicted those by Lawes & Bramley (2012), who concluded that pairwise comparison between data points, without taking into account spatial covariance, and linear model estimation with spatial covariance, do not produce significantly different results. However, our analysis suggests that the effects of nitrogen are drastically different whether the Least Squares approach, a model that assumed independence, or the Generalized Least Squares approach, a model that takes into account spatial covariance, are used. The noticeable difference between our analysis and Lawes & Bramley (2012) is the use of yield stability-based management zones. We analyzed the impact of nitrogen per yield stability-based management zones, delineated according to Kharel et al. (2019), which provides information about both the temporal average yield and its variation in relation to the farm-level temporal average yield and its variation. While Lawes & Bramley

(2012) use management zones, they do not take into account spatial yield trends nor yield variability, which could be caused by a variety of factors, including the weather, management, and topography (Andresen et al., 2001; Kravchenko et al., 2005; Mallory & Porter, 2007; Smith et al., 2007). Temporal and spatial yield variation within a field is known to be high in the New York State and it is crucial to take into account those variabilities to appropriately estimate the effect of set treatment from an OFE.

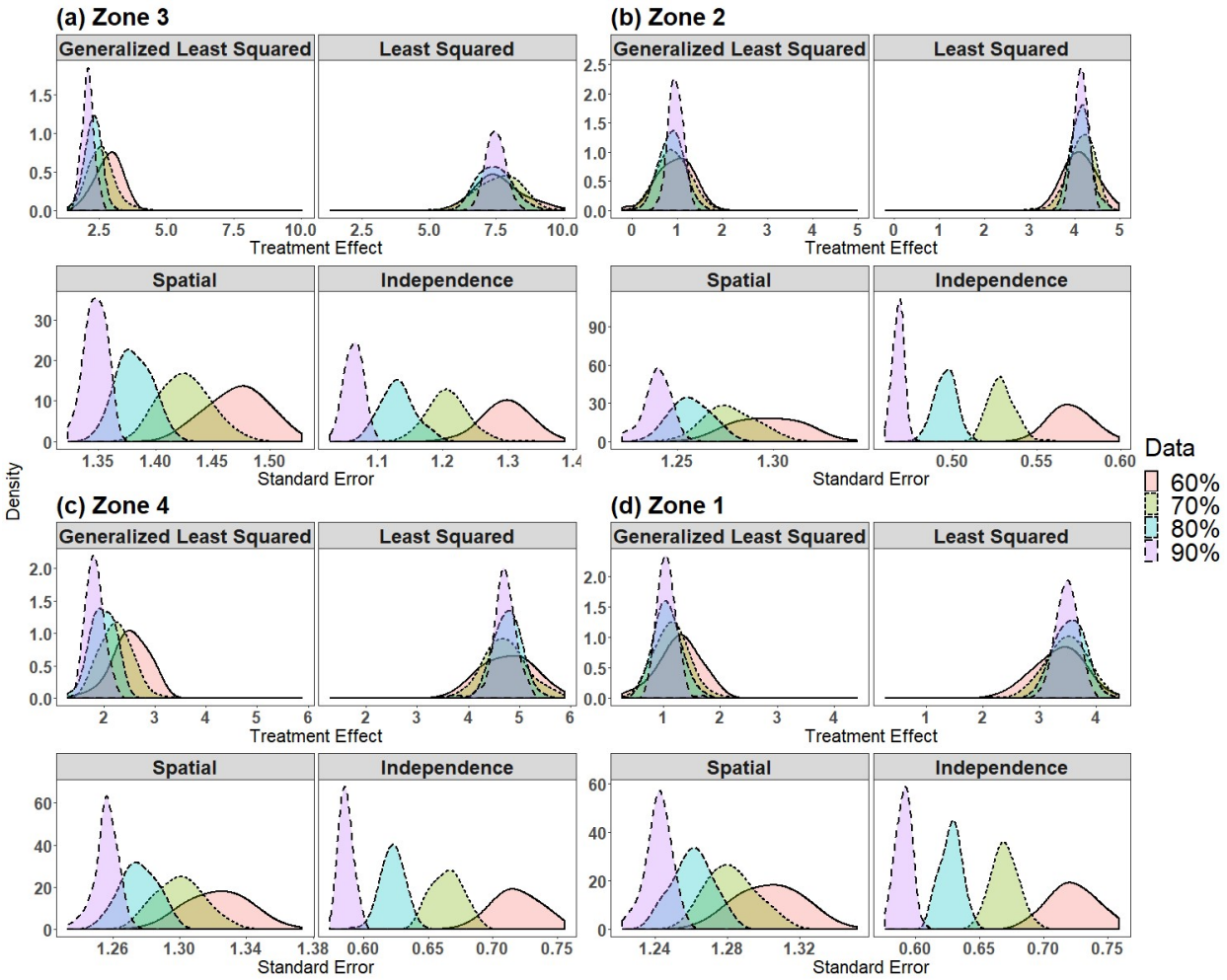

Figure 3.7: The effect of N treatment and standard errors of Zones 1(d), Zone 2(b), Zone 3(a), and Zone 4(c) for Field E was estimated using 60%, 70%, 80%, and 90% of randomly sampled data, 100 times each, using the Generalized Least Squares approach and the Least Squares approach, assuming spatial covariance versus independence.

CONCLUSIONS

It remains obvious, from a practical perspective, that for farmers interested in on-farm evaluation of a change in management, a single strip field evaluation is much easier to implement that multi-strip trials. Here, we explored approaches for estimating the effect of the treatment and uncertainty around that estimation (standard error) for a single strip treatment trial. The results of this study show that when yield is documented using yield monitor systems that collect yield data per second, variability in a field can be quantified using management zones. With high resolution zone maps and spatial covariance, treatment effects from a single strip treatment trial can be effectively estimated. Thus, we propose the use of LS 2, i.e., estimation of treatment effects using the Least Squares approach and standard errors with spatial covariance, for analyzing whether a specific management change is a benefit in terms of yield or other desired outcomes. Through this approach, we may appropriately estimate the effect of treatment and standard errors from any un-replicated, single-strip OFE, based on historic yield monitor data and yield stability-based management zones. Moreover, it allows for more field trial data to be added over time for better understanding of drivers for outcomes such as yield and the need for site-specific management in unstable yielding zones.

REFERENCES

Andresen, J. A., Alagarswamy, G., Rotz, C. A., Ritchie, J. T., & LeBaron, A.W. (2001). Weather impacts on maize, soybean, and alfalfa production in the Great Lakes region, 1895–1996. Agronomy Journal, 93(5), 1059–1070. https://doi.org/10.2134/agronj2001.9351059x

Cho, J. B., Guinness, J., Kharel, T. P., Sunoj, S., Kharel, D., Oware, E. K., Aardt, J. A. V., & Ketterings, Q. M. (2021). Spatial estimation methods for mapping corn silage and grain yield monitor data. *Precision Agriculture*. https://doi.org/10.1007/s11119-021-09793-z

Cox, D. R. (2009) Randomization in the design of experiments. *International Statistical Review 77*, 415–429. https://doi.org/10.1111/j.1751-5823.2009.00084.x

Dobermann, A., & Ping, J. L. (2004). Geostatistical integration of yield monitor data and remote sensing improves yield maps. *Agronomy Journal*, *96*(1), 285–297. https://doi.org/10.2134/agronj2004.0285

Federer, W. T. (1955). *Experimental design*. New York, USA. The MacMillan Company.

Fisher, R. A. (1926). The Arrangement of Field Experiments. *Journal of the Ministry of Agriculture of Great Britain, 33*, 503-513.

Fisher, R. A. (1931). Principles of plot experimentation in relation to the statistical interpretation of the results. *Rothamsted Conferences, 13: The Technique of Field Experiments held at Rothamsted on 7 May, 1931 under the Chairmanship of Sir A. D. Hall,* 11-13. Rothamsted Experimental Station, Harpenden.

Gomez, K. A., & Gomez, A. A. (1984). *Statistical procedures for agricultural research.* Wiley, New York, USA.

Gotway Crawford, C. A., Bullock, D., Pierce, F., Stroup, W., Hergert, G. & Eskridge, K. (1997). Experimental Design Issues and Statistical Evaluation Techniques for Site-Specific Management. *In The State of Site Specific Management for Agriculture, 14.* https://doi/org /10.2134/1997.stateofsitespecific.c14

Griffin, T. W., Florax, R. J. G. M., & Lowenberg-DeBoer, J. M. (2006). *Field-Scale* Experimental Designs and Spatial Econometric Methods for Precision Farming: Strip-

Trial Designs for Rice Production Decision Making [Paper presentation]. *2006 Annual Meeting Southern Agricultural Economics Association*, Orlando, FL, United States. https://doi/org/10.22004/ag.econ.35367

Griffin, T. W., Lambert, D. M., & Lowenberg-DeBoer, J. M. (2004). Testing appropriate on-farm trial designs and statistical methods for precision farming: a simulation approach. *Proceedings of the 7th International Conference on Precision Agriculture and Other Precision Resources Management*, *USA*, 1733-1748.

Guinness, J., & Katzfuss, M. (2019) *GpGp: Fast Gaussian process computation using Vecchia's approximation.* R package version 0.3.2. Retrieved April 21, 2021, from https://CRAN.R-project.org/package=GpGp

Johnson, D. H. (2006). The many faces of replication. *Crop Science, 46*(6), 2486–2491. https://doi.org/10.2135/cropsci2006.04.0277

Kharel, T. P., Marema, A., Czymmek, K. J., Oware, E. K., & Ketterings, Q. M. (2019a). Combining spatial and temporal corn silage yield variability for management zone development. *Agronomy Journal, 111*(6), 2703–2711. https://doi.org/10.2134/agronj2019.02.0079

Kharel, T. P., Swink, S. N., Maresma, A., Youngerman, C., Kharel D., Czymmek, K. J. & Ketterings Q. M. (2019b). Yield monitor data cleaning is essential for accurate corn grain/silage yield determination. *Agronomy Journal 111*(2), 509–516. https://doi.org/10.2134/agronj2018.05.0317

Kharel, T., Swink, S. N., Youngerman, C., Maresma, A., Czymmek, K. J., Ketterings, Q. M., Kyveryga, P., Lory, J., Musket, T. A., & Hubbard, V. (2018). *Processing/cleaning corn silage and grain yield monitor data for standardized yield maps across farms, fields, and*

years. Cornell University, Nutrient Management Spear Program, Department of Animal Science, Ithaca, NY, USA. Retrieved June 9, 2021, from http://nmsp.cals.cornell.edu/publications/extension/ProtocolYieldMonitorDataProcessing1_15_2020.pdf

Kim, S., Daughtry, C., Russ, A., Pedrera-Parrilla, A., & Pachepsky, Y. (2020). Analysis of Spatiotemporal Variability of Corn Yields Using Empirical Orthogonal Functions. *Water, 12*(12), 3339. https://doi.org/10.3390/w12123339

Kravchenko, A. N., Robertson, G.P., Thelen, K. D., & Harwood, R. R. (2005). Management, topographical, and weather effects on spatial variability of crop grain yields. Agronomy Journal, 97(2), 514–523. https://doi.org/10.2134/agronj2005.0514

Kyveryga, P. M. (2019). On-Farm Research: Experimental Approaches, Analytical Frameworks, Case Studies, and Impact. *Agronomy Journal, 111*(6), 2633-2635. https://doi.org/10.2134/agronj2019.11.0001

Kyveryga, P. M., Mueller, T. A., & Mueller, D. M. (2018). On-Farm Replicated Strip Trials. D. Kent Shannon, D. E. Clay, & N. R. Kitchen (Eds). *Precision Agriculture Basics*. ASA, CSSA, SSSA, Madison, WI.

Kyveryga, P. M., Mueller, T., Paul, N., Arp, A., & Reeg, P. (2015). *Guide to On-Farm Replicated Strip Trials.* Iowa Soybean Association, IA, USA. Retrieved June 7, 2021, from https://www.researchgate.net/publication/315664663_Guide_to_On-Farm_Replicated_Strip_Trials

Lambert D. M., Lowenberg-Deboer, J., Bongiovanni, R. (2004). A comparison of four spatial regression models for yield monitor data: A case study from Argentina. *Precision Agriculture, 5*, 579–600. https://doi.org/10.1007/s11119-004- 6344-3

Lawes, R. A., & Bramley, R. G. V. (2012). A Simple Method for the Analysis of On-Farm Strip Trials. *Agronomy Journal, 104*(2), 371 – 377. https://doi.org/10.2134/agronj2011.0155

Little, T. M., & Hills, F. J. (1978). *Agricultural Experimentation: Design and Analysis,* John Wiley & Sons Ltd.

Mallory, E. B., & Porter, G. A. (2007). Potato yield stability under contrasting soil management strategies. *Agronomy Journal, 99*(2) 501–510. https://doi.org/10.2134/agronj2006.0105

Mutsaers, H. J. W. (1997) *A Field Guide for On-farm Experimentation.* International Institute of Tropical Agriculture.

Oehlert, G. W. (2010). *A First Course in the Design and Analysis of Experiments*. University of Minnesota Digital Conservancy.

Panten, K., Bramley, R. G. V., Lark, R. M., & Bishop T. F. A. (2010). Enhancing the value of field experimentation through whole-of-block designs. *Precision Agriculture 11*, 198–213. https://doi.org/10.1007/s11119-009-9128-y

Petersen, R. G. (1994). *Agricultural field experiments. Design and analysis.* Marcel Dekker Inc.

Piepho, H., Richter, C., Spilke, J., Hartung, K., Kunick, A., & Thöle, H. (2011). Statistical aspects of on-farm experimentation. *Crop & Pasture Science, 62*, 721-735. https://doi.org/10.1071/CP11175

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Rudolph, S., Marchant, P., Gillingham, V., Kindred, D., & Sylvester-Bradley, R. (2016*). Spatial Discontinuity Analysis, a novel geostatistical algorithm for on-farm experimentation.* Paper presented at the Proceedings of the 13th International Conference on Precision Agriculture.

Sawyer, J. E. (1994). Concepts of variable rate technology with considerations for fertilizer application. *Journal of Production Agriculture, 7*(2), 195–201. https://doi.org/10.2134/jpa1994.0195

Stroup, W. W., Hildebrand, P. E. & Francis, C. A. (1993). Farmer Participation for More Effective Research in Sustainable Agriculture. *Technologies for Sustainable Agriculture in the Tropics*, *56*. https://doi.org/10.2134/asaspecpub56.c12

Sudduth, K. A., & Drummond, S. T. (2007). Yield Editor. *Agronomy Journal*, *99*(6), 1471–1482. https://doi.org/10.2134/agronj2006.0326

Sudduth, K. A., Drummond, S. T., & Myers, D. B. (2012). Yield Editor 2.0: Software for automated removal of yield map errors. In 2012 Dallas, Texas, July 29-August 1, 2012 (p. 1). American Society of Agricultural and Biological Engineers. Retrieved April 21, 2021, from https://doi.org/10.13031/2013.41893

Vega, A., Córdoba, M., Castro-Franco, M., & Balzarini M. (2019) Protocol for automating error removal from yield maps. *Precision Agriculture, 20,* 1033–1044. https://doi.org/10.1007/s11119-018-09632-8

Willers, J. L., Milliken, G. A., Jenkins, J. N., O'Hara, C. G., Gerard, P. D., Reynolds D. B., Boykin D. L., Good P. V., & Hood K. B. (2008). Defining the experimental unit for the design and analysis of site-specific experiments in commercial cotton fields. *Agricultural Systems, 96*, 237–249. doi:10.1016/j.agsy.2007.09.003

Yan, W., Hunt, L. A., Johnson, P., Stewart, G., Lu, X. (2002). On-farm strip trials vs. replicated performance trials for cultivar evaluation. *Crop Science, 42*(2), 385–392. https://doi.org/10.2135/cropsci2002.0385