CLIMATIC VARIABILITY AND CHANGE IN THE MIDWESTERN UNITED STATES:
IMPLICATIONS FOR NITROGEN LEACHING IN AGRICULTURAL SYSTEMS

By

William James Baule

A DISSERTATION

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

Geography-Doctor of Philosophy

2022

ABSTRACT

CLIMATIC VARIABILITY AND CHANGE IN THE MIDWESTERN UNITED STATES:
IMPLICATIONS FOR NITROGEN LEACHING IN AGRICULTURAL SYSTEMS

By

William James Baule

How has the background climate of the Midwestern United States changed over recent
decades and how has this affected nitrate leaching? These are the core questions addressed in this
dissertation, through three self-contained studies focused on different aspects of the climate-
agriculture interface in the Midwestern United States. In Chapter 2, statistical methods are used
to quantify the solar radiation biases present in a widely used reanalysis-based
hydrometeorological dataset over space, implement statistical bias correction and interpolation to
address the spatial nature of this bias, and quantify the impacts of the solar radiation bias and
proposed correction on simulated maize yields and water stress. Correction of reanalysis solar
radiation alone brought simulated yield and water usage more in line with simulations forced
with in-situ solar radiation. Chapter 3 examines changes in precipitation, utilizing a unique
approach to station screening during the period 1951-2019 over a region encompassing the Great
Lakes and broader Midwestern regions, of the United States. A multiple tier procedure was
utilized to identify high quality input data series from the Global Historical Climatology
Network-Daily dataset. Temporal and spatial trends were analyzed for a broad range of related
annual and seasonal indicators ranging from accumulated totals and frequency of threshold
events to event duration and potential linkages with total precipitable water. Our analyses
confirm the results of previous studies while providing unique insights to data quality and
seasonality. The trends of the indicators in our study exhibited more cohesive spatial patterns and
temporal similarities when compared with studies with different quality control criteria,

illustrating the importance of quality control of observations in climatic studies and highlighting the complexity of the changing character of precipitation. In Chapter 4, System Approach to Land Use Sustainability, a process-based crop model was applied with gridded soil and meteorological data using a yield stability zone concept to simulate corn and soybean production in 14 Midwestern states at the sub-field scale during the 1989-2019 period. Five zones based on multi-year yield stability were simulated for each field at 30m x 30m resolution, with zones being relative to each individual field. Outputs were evaluated using a nitrogen balance approach to establish zone-specific statistical distributions of nitrate leaching across the 14 states, specifically highlighting periods with changing and highly variable precipitation. Results indicate that low stable, unstable hill tops, and unstable slope zones are associated with an outsized contribution to overall nitrate leaching and that unstable zones exhibit variable year-to-year response to weather tied to their position in the landscape. Spatial analysis of the results suggests leaching is tied to precipitation variability, water stress, and total precipitation amount. In aggregate, the chapters presented here highlight the interconnectedness of the soil-plant-atmosphere continuum to changes in hydrologic regime and sensitivity to the biases in the data used to conduct analyses, run models, and from which conclusions are drawn. The study findings shed light on the potential for improved management of agricultural fields and illustrate how process-based crop models can be useful for designing management practices to reduce environmental pollution and increase profits to producers.

ACKNOWLEDGEMENTS

First, I want to thank my advisor Dr. Jeff Andresen. Without his patience, guidance, and expertise this dissertation would not have been possible. Next, I thank my dissertation committee: Drs. Julie Winkler, Bruno Basso, and Amor Valeriano Ines for their support during this project. In addition, I would like to thank the faculty and staff of the Department of Geography, Environment, & Spatial Sciences at Michigan State University, the Great Lakes Integrated Sciences + Assessments, the Basso Lab students and staff, my family, fellow graduate students, and friends for their continued support and comradery.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO ABBREVIATIONS

ASOS	Automated Surface Observing System
BC/QM	Bias Correction and Quantile Mapped
CDD	Consecutive dry days
CDL	Cropland Data Layer
CENTURY	CENTURY soil organic matter model
CLU	Common Land Unit
COOP	Cooperative Observer Program
CRN	Climate Reference Network
CWD	Consecutive wet days
DAYCENT	Daily Version of CENTURY
DD	Number of days with consecutive days without measurable precipitation
DF	Distribution Fit
DJF	December January February
DNCD	Denitrification and DeComposition model
DON	Dissolved Organic Nitrogen
DOY	Day of Year
Drght_Fac	Drought Stress
DRN	Cumulative drainage
DSSAT	Decision Support System for Agrotechnology Transfer
EOA	Evapotranspiration
ET	Evapotranspiration

ET/PET	Ratio of October-September totals of SALUS simulated Evapotranspiration to Penman-Monteith grass reference Potential Evapotranspiration
ETCCDI	Expert Team on Climate Change Detection and Indices
GHCN	Global Historical Climatology Network
GHCN-D	Global Historical Climatology Network-Daily
GLISA	Great Lakes Integrated Sciences + Assessments
GWAD	Dry Grain Weight
HS	High Stable
IDW	Inverse Distance Weighting
JJA	June July August
kg/ha	kilogram per hectare
LS	Low Stable
m	meter
mm	Millimeter
MAM	March April May
MAXT GS	Annual Growing Season Mean High Temperature
MBN	Microbial Biomass
N	Nitrogen
N_Plant	Plant nitrogen uptake
NARR	North American Regional Reanalysis
NCAR	National Center for Atmospheric Research
NCEI	National Centers for Environmental Information
NCEP	National Center for Environmental Prediction

NDVI	Normalized Difference Vegetation Index
Nitrate_Bl	Nitrate below ground
NLC	Nitrate Leached
NLDAS-2	National Land Data Assimilation System version 2
PAW	Plant Available Water
PET	Potential Evapotranspiration
PET WAT YR	Water year (October-September) Penman-Monteith grass reference Potential Evapotranspiration
PRCP_GS	Growing Season (May-Sept.) Precipitation Accumulation (mm)
PRCPTOT	Annual total wet day precipitation
QC	Quality Control
r	Pearson's Correlation Coefficient
R^2	Coefficient of Determination
R1.26mm	Number of days with measurable precipitation
R10mm	Number of heavy precipitation days
R20mm	Number of very heavy precipitation days
R95pTOT	Precipitation on very wet days
R99pTOT	Precipitation on extremely wet days
RMSE	Root Mean Squared Error
ROF	Cumulative runoff
Rx1day	Maximum 1-day precipitation
Rx5day	Maximum 5-day precipitation
S	Slope
SALUS	System Approach to Land Use Sustainability

SDII	Simple daily intensity index
SOM	Soil Organic Matter
SON	Soil Organic Nitrogen
SON	September October November
TOTPRCPWAT	Total Precipitable Water
USA	United States of America
USD	Unstable Depression
USDA	Unites States Department of Agriculture
USH	Unstable Hilltop
USHCN	United States Historical Climate Network
USO	Unstable Other
WW	Number of days with consecutive days with measurable precipitation
WWI	Weather Whiplash Index
YSZ	Yield Stability Zone

CHAPTER 1. INTRODUCTION

Research Context

Climate change, climate variability, and the underlying quality of observed climate data are all critical to our understanding and interpretation of climatic phenomena and their impacts. This is particularly true when considering processes that use climate information as inputs to models or in decision making processes (Briley et al. 2015; Winkler 2016). The quality and completeness of the data available, the methods by which they were observed/generated, and inherent nature of the climatic variables under consideration can all influence the conclusions drawn from analysis. This extends from primary atmospheric variables of temperature and precipitation to less frequently observed data such as daily solar radiation. The nature of each individual variable presents unique challenges to observation/quality-control.

One of the climatic variables with far fewer available observations over the period of instrumental records is solar radiation (Perdinan et al. 2020; Kiefer et al. 2019; Slater 2016). It is not regularly included in routine measurements taken by the two largest in-situ climate and observing networks in the USA (i.e. the Cooperative Observer Program (COOP) and the Automated Surface Observing System (ASOS)), and until recently, with the advent of the Climate Reference Network (CRN), was restricted to specialty networks such as those focused on agricultural or emergency management (Mahmood et al. 2017; Diamond et al. 2013). Given the paucity of in-situ observations over space and time, remotely sensed, reanalysis based, or modeled solar observation values are frequently used in-lieu of observations. The biases introduced by each of these methods are different, with reanalysis displaying some of the largest biases (Perdinan et al. 2020). Reanalyses are desirable for modeling applications as their outputs are usually serially complete, both spatially and temporally, and are informed by in-situ observations and representations of physical processes. However, due to the sparse nature of the

observational record of solar radiation, both spatially and temporally, many open questions

remain regarding biases and their treatment in modeled solar radiation values.

Changes in total annual precipitation and the intensity and frequency of heavy

precipitation have occurred across many areas of the United States and around the world during

the twentieth century and early twenty-first centuries (Baule et al., 2022; Konrad 2001; Pryor et

al. 2009; Pryor 2009; Walsh et al. 2014; Westra et al. 2013). These trends are directly associated

with many impacts including increases in crop damage and large financial losses related to the

resultant flooding (Rosenzweig et al. 2002; Pielke et al. 2000). Establishing a quantitative, cause

and effect relationship between changing precipitation and crop-losses over time has proven

difficult due to heterogeneities in agricultural practices (i.e. fertilizer form/rate) and crop genetics

over several decades (Rosenzweig et al. 2002). Loss of nutrients, particularly nitrogen, is one

such risk related to heavy precipitation and can represent a substantial financial liability to

producers depending on crop fertility requirements, fertilizer application rates and operation size

(Robertson et al. 2013).

Precipitation patterns in the Midwestern region of North America have exhibited

substantial changes in recent decades, particularly the intensity and timing of heavy precipitation

events. However, there has been substantially less documentation in the literature on moderate

and light accumulation events (e.g. Roque-Malo and Kumar 2017), especially during the cooler

seasons. The recently observed changes in the precipitation regime of the Midwest have also

resulted in changes in the nitrogen cycle, as precipitation and soil moisture act as the primary

physical drivers of nitrogen cycling in terrestrial ecosystems. (Bernot et al. 2006; Galloway et al.

2004). Changes in the timing, intensity, and form of precipitation have the potential to

substantially change the cycling of nitrogen fertilizer in the location where it's applied (e.g.

mineralization, immobilization, uptake, leaching, denitrification, etc.) and in sources downstream of the agricultural fields from where it is transported via surface runoff and leaching through the soil profile (Kalkhoff et al. 2016). This is particularly true during times of the year prior to the establishment of the crop before it can be taken up by the crop, during heavy precipitation events, or in fields where nitrogen fertilizers have been applied in excess of the needs of the planted crop (Di and Cameron 2002). All three factors mentioned are important in the Midwest, however the excessive use of nitrogen fertilizers is the most frequently discussed of the three factors in the literature and the target of management practices aimed at reducing nitrate pollution of the surrounding environment. It has been shown through extensive field trials in the state of Michigan that the uptake of applied nitrogen in grain crops on average is only about 50% of the total amount applied to the field (Robertson 1997; Syswerda et al. 2012). Globally the amount of nitrogen lost to the environment is estimated to be higher, near 2/3^{rds} of the total applied (Liu et al. 2010). The result of excessive nitrogen application is the potential loss of these nutrients to the larger environment, often through leaching into groundwater or transported to surface waters through drainage tile (Basso and Ritchie 2005). In addition to heavy precipitation, excessive irrigation is also a major contributor to the leaching of excess to nitrogen in areas where irrigation is prevalent (Letey and Vaughan 2013), although this plays less of a role in the Midwest where the majority of crops are produced under rainfed conditions (Green et al., 2018). Excess nitrogen has been shown to interfere with various ecosystem processes downstream of agricultural lands under cultivation and has potential as a significant pollutant of ground and surface waters, especially in the form of nitrate (NO_3^-) (Lin et al. 2001).

Given the observed and projected changes in precipitation and the intricate linkages of the climate system to nitrogen cycling, Bowles et al. (2018) outlined several potential impacts of

drier summers with more consecutive dry days, more intense rainfall when rain does occur

(which is projected for much of the Midwest in the future (Hayhoe et al., 2018)), and low soil

moisture on nitrogen cycling in the central United States. Soil moisture impacts microbial redox

reactions that transform nitrogen (Bowles et al. 2018), and also affects the movement of

substrates and products during decomposition Low soil moisture reduces overall microbial

activity, and slows organic nitrogen breakdown, nitrogen mineralization and especially

nitrification, denitrification, and nitrate leaching. Low soil moisture also affects microbial uptake

of inorganic nitrogen and reduces plant uptake as growth and nitrogen transport to roots both

slow. Conversely, as soil becomes saturated and oxygen availability is reduced, denitrification

accelerates losses of nitrous oxide. Leaching losses are also increased when saturated soils allow

the movement of nitrate below the root zone, ultimately polluting the ground or surface water.

Thus, rainfall and soil water dynamics directly influence N transformations critical for crop

nitrogen availability, and also drive the process of leaching and denitrification that release

nitrogen into surrounding ecosystems.

   Changing climatic conditions and weather variability have been shown to have impacts

on crop yields and the fate of nutrients at the regional and field scales in different regions across

the globe both in observational (e.g. Zhou and Butterbach-Bahl 2014) and simulation studies

(e.g. Congreves et al. 2016). In these studies, individual fields have often been treated as

homogenous land units, and management has been uniform across the field. However,

substantial variability occurs at sub-field scales, and management can be tailored to specific

zones within a field based on a variety of methods (e.g. historical yield, soil type, position in the

landscape, yield stability). Knowledge and management of yield stability zones are one such

methodological approach that incorporates georeferenced grain yield data from yield monitors or

remotely sensed imagery and delineates zones based on inter-annual variability of crop yields (Basso et al. 2019; Jin et al. 2019; Maestrini and Basso 2018a,b; Basso et al. 2007). Once identified, these zones can be utilized to prescriptively manage that field to optimize yields, profits, and reduce negative environmental effects from excess fertilization. These zones can also be used to better understand the influence of weather variability and climate on different crops in non-homogenous zones at the sub-field scale (Maestrini and Basso 2018a). Unanswered questions remain regarding the behavior of these zones over time in terms of nitrogen related variables and how the zones respond to weather/climate variability.

Dissertation Focus and Organization

The goal of this work is to address questions regarding the spatial and temporal nature of climatic data quality, climatic trends/variability, and how this affects simulating the impacts of climate on nitrogen loss in a diverse regional landscape. Chapters 2 through 4 are self-contained studies, each with their own unique but interrelated questions. The following research questions are addressed in Chapters 2 through 4.

(1) *Is it possible to correct a known spatially-variable bias in modeled solar radiation values, given a paucity of observational data?*

(2) *How has precipitation changed across the Midwestern and Great Lakes Regions of the United States from 1951-2019, and how does data quality affect interpretation of hydroclimatic trends?*

(3) *Given background hydroclimatic trends, how has sub-field nitrogen loss associated with corn production responded across the Midwestern United States from 1989-2019?*

In Chapter 2, statistical methods are used to quanitfy the solar radiation biases across space present in a widely-used reanlysis-based hydrometeorological dataset, implement statistical bias correction and interpolation to address the spatial nature of this bias, and quantify the impacts of bias and the proposed correction to solar radiation on simulated maize yields and water stress. Impacts of the bias correction are illustrated using SALUS at multiple locations across the study region. Correction of the reanalysis solar radiation alone brought simulated yield and water usage more in line with simulations forced with in-situ solar radiation, although the effect of differences in temperature between in-situ observations and reanalysis has substantial effect of simulated yields and water usage.

Chapter 3 identifies and examines changes in precipitation, utilizing a unique approach to station screening during the period 1951-2019 over a region encompassing the Great Lakes and broader Midwestern region of the United States. A multiple tier procedure was utilized to identify high quality input data series from the Global Historical Climatology Network-Daily dataset. Annual and seasonal time series of precipitation indicators were calculated and subjected to breakpoint analysis as further quality control. Temporal and spatial trends were analyzed for a broad range of related indicators ranging from accumulated totals and frequencies of threshold events to event duration and linkages of trends in precipitation indicators with trends in total precipitable water. Results indicate that precipitation has generally increased across the region in terms of magnitude, although there is substantial variation across the study domain in significance and magnitude of trends by indicator. Trends were spatially most consistent across eastern areas of the study domain while relatively greater site to site variability in precipitation and trends was observed across northern and western portions. Trends in the seasonal indicators were generally fewer in number and less spatially coherent. The greatest numbers and most of

the significant trends of the seasonal indicators occurred in the fall with the fewest in the spring. Correlation of indicator trends with annual and seasonal trends of mean total precipitable water suggests weak correlations annually and moderate correlations at the seasonal level. Our results confirm the results of previous studies while providing unique insights to data quality and seasonality. The trends of the indicators in our study generally exhibited more cohesive spatial patterns and temporal similarities when compared with studies with different quality control criteria, illustrating the importance of quality control of observations in climatic studies and highlight the complexity of the changing character of precipitation.

In Chapter 4, the SALUS (System Approach to Land Use Sustainability) process-based crop model was applied with gridded soil and meteorological data using a yield stability zone concept (Basso et al. 2019) to simulate corn and soybean production in 14 Midwestern states at the sub-field scale during the 1989-2019 period. Five zones based on multi-year yield stability were simulated for each field at ~30m x 30m resolution. Individual fields were delineated by Common Land Unit (CLU). Outputs were evaluated using a nitrogen balance approach to establish zone-specific statistical distributions of nitrate leaching across the 14 states, specifically highlighting periods with changing and highly variable precipitation. Results indicate that areas of the field with low and/or unstable yields are associated with an outsized contribution to overall nitrate leaching and that unstable zones exhibit variable year-to-year response to weather tied to their position in the landscape. Spatial analysis of the results suggests leaching is tied to precipitation variability, water stress, and total precipitation amount. Three tillage regimes were simulated, with the increase in nitrogen cycling increasing with the depth of tillage. When constrained to individual soil type, relationships to water stress, precipitation, nitrogen uptake, soil nitrogen, and leaching become clearer. At larger spatial scales inherent variability tied to soil

type, soil organic matter (SOM), and intensity of production dominate the aggregate responses in nitrate leaching. The regional analysis conducted here, focused on the relative behavior of the yield stability zones, gives abundant directions for further analysis to potentially shed light on the role of Plant Available Water (PAW) and SOM in regional nitrogen dynamics in relation to climate variability.

Chapter 5 summarizes and discusses these results from the three preceding chapters, offers directions for future research, and highlights the intellectual merit and broader impacts of this research, both to the research community and society at large.

REFERENCES

REFERENCES

Basso, B., and J. T. Ritchie, 2005: Impact of compost, manure and inorganic fertilizer on nitrate leaching and yield for a 6-year maize–alfalfa rotation in Michigan. *Agric. Ecosyst. Environ.*, **108**, 329–341, doi:10.1016/j.agee.2005.01.011. http://www.sciencedirect.com/science/article/pii/S0167880905000435.

——, M. Bertocco, L. Sartori, and E. C. Martin, 2007: Analyzing the effects of climate variability on spatial pattern of yield in a maize–wheat–soybean rotation. *Eur. J. Agron.*, **26**, 82–91, doi:10.1016/J.EJA.2006.08.008. https://www.sciencedirect.com/science/article/pii/S1161030106001067 (Accessed September 25, 2019).

——, G. Shuai, J. Zhang, and G. P. Robertson, 2019: Yield stability analysis reveals sources of large-scale nitrogen loss from the US Midwest. *Sci. Rep.*, **9**, 5774, doi:10.1038/s41598-019-42271-1. http://www.nature.com/articles/s41598-019-42271-1 (Accessed September 25, 2019).

Baule, W.J., Andresen, J.A. and Winkler, J.A., 2022: Trends in Quality Controlled Precipitation Indicators in the United States Midwest and Great Lakes Region. *Frontiers in Water*, p.8. doi: 10.3389/frwa.2022.817342

Bernot, M. J., J. L. Tank, T. V Royer, and M. B. David, 2006: Nutrient uptake in streams draining agricultural catchments of the midwestern United States. *Freshw. Biol.*, **51**, 499–509, doi:10.1111/j.1365-2427.2006.01508.x. http://onlinelibrary.wiley.com/doi/10.1111/j.1365-2427.2006.01508.x/abstract.

Briley, L., D. Brown, and S.E. Kalafatis, 2015: Overcoming barriers during the co-production of climate information for decision-making. Clim Risk Mngmt, **9**, 41-49, https://doi.org/10.1016/j.crm.2015.04.004

Bowles, T. M., S. S. Atallah, E. E. Campbell, A. C. M. Gaudin, W. R. Wieder, and A. S. Grandy, 2018: Addressing agricultural nitrogen losses in a changing climate. *Nat. Sustain.*, **1**, 399–408, doi:10.1038/s41893-018-0106-0. http://www.nature.com/articles/s41893-018-0106-0 (Accessed September 25, 2019).

Congreves, K. A., B. Dutta, B. B. Grant, W. N. Smith, R. L. Desjardins, and C. Wagner-Riddle, 2016: How does climate variability influence nitrogen loss in temperate agroecosystems under contrasting management systems? *Agric. Ecosyst. Environ.*, **227**, 33–41, doi:10.1016/J.AGEE.2016.04.025. https://www.sciencedirect.com/science/article/pii/S0167880916302286#bib0150 (Accessed April 29, 2019).

Di, H. J., and K. C. Cameron, 2002: Nitrate leaching in temperate agroecosystems: sources, factors and mitigating strategies. *Nutr. Cycl. Agroecosystems*, **64**, 237–256, doi:10.1023/A:1021471531188. http://link.springer.com/10.1023/A:1021471531188 (Accessed September 25, 2019).

Diamond, H.J., T.R. Karl, M.A. Palecki, C.B. Baker, J.E. Bell, R.D. Leeper, D.R. Easterling, J.H. Lawrimore, T.P. Meyers, M.R. Hlfert, G. Goodge, and P.W. Thorne, 2013: U.S. climate reference network after one decade of operations: status and assessment. *Bull. American Meteorological Soc.*, **94**, 485-498, doi: 10.1175/BAMS-D-12-00170.1

Galloway, J. N., and Coauthors, 2004: Nitrogen Cycles: Past, Present, and Future. *Biogeochemistry*, **70**, 153–226, doi:10.1007/s10533-004-0370-0. https://link.springer.com/article/10.1007/s10533-004-0370-0.

Hayhoe, K., and Coauthors, 2018: *Chapter 2 : Our Changing Climate. Impacts, Risks, and Adaptation in the United States: The Fourth National Climate Assessment, Volume II*. Washington, DC, https://nca2018.globalchange.gov/chapter/2/ (Accessed September 25, 2019)

Jin, Z., S. V. Archontoulis, and D. B. Lobell, 2019: How much will precision nitrogen management pay off? An evaluation based on simulating thousands of corn fields over the US Corn-Belt. *F. Crop. Res.*, **240**, 12–22, doi:10.1016/J.FCR.2019.04.013. https://www.sciencedirect.com/science/article/pii/S0378429018314369 (Accessed September 25, 2019).

Kalkhoff, S. J., L. E. Hubbard, M. D. Tomer, and D. E. James, 2016: Effect of variable annual precipitation and nutrient input on nitrogen and phosphorus transport from two Midwestern agricultural watersheds. *Sci. Total Environ.*, **559**, 53–62, doi:10.1016/j.scitotenv.2016.03.127. http://www.sciencedirect.com/science/article/pii/S004896971630554X.

Kiefer, M.T., J.A. Andresen, D. Doubler, A. Pollyea, 2019: Development of a gridded reference evapotranspiration dataset for the Great Lakes region. *J Hydrology: Regional Studies*, **24**, 100606, doi: 10.1016/j.ejrh.2019.100606

Konrad, C. E., 2001: The Most Extreme Precipitation Events over the Eastern United States from 1950 to 1996: Considerations of Scale. *J. Hydrometeorol.*, **2**, 309–325, doi:10.1175/1525-7541(2001)002<0309:TMEPEO>2.0.CO;2. http://journals.ametsoc.org/doi/abs/10.1175/1525-7541%282001%29002%3C0309%3ATMEPEO%3E2.0.CO%3B2 (Accessed September 25, 2019).

Letey, J., and P. Vaughan, 2013: Soil type, crop and irrigation technique affect nitrogen leaching to groundwater. *Calif. Agric.*, **67**, 231–241.

Lin, B.-L., A. Sakoda, R. Shibasaki, and M. Suzuki, 2001: A Modelling Approach to Global Nitrate Leaching Caused by Anthropogenic Fertilisation. *Water Res.*, **35**, 1961–1968, doi:10.1016/S0043-1354(00)00484-X. http://www.sciencedirect.com/science/article/pii/S004313540000484X.

Liu, H. L., and Coauthors, 2011: Using the DSSAT-CERES-Maize model to simulate crop yield and nitrogen cycling in fields under long-term continuous maize production. *Nutr. Cycl. Agroecosystems*, **89**, 313–328, doi:10.1007/s10705-010-9396-y. http://link.springer.com/10.1007/s10705-010-9396-y (Accessed September 25, 2019).

Maestrini, B., and B. Basso, 2018a: Drivers of within-field spatial and temporal variability of crop yield across the US Midwest. *Sci. Rep.*, **8**, 14833, doi:10.1038/s41598-018-32779-3. http://www.nature.com/articles/s41598-018-32779-3 (Accessed September 25, 2019).

——, and ——, 2018b: Predicting spatial patterns of within-field crop yield variability. *F. Crop. Res.*, **219**, 106–112, doi:10.1016/j.fcr.2018.01.028. https://linkinghub.elsevier.com/retrieve/pii/S0378429017315435 (Accessed September 25, 2019).

Mahmood, R., R. Boyles, K. Brinson, C. Fiebrich, S. Foster, K. Hubbard, D. Robinson, J. Andresen, and D. Leathers, 2017: Mesonets: mesoscale weather and climate observations for the United States. *Bull. American Meteorology Soc.*, **98,** 1349-1361, doi: 10.1175/BAMS-D-15-00258.1

Perdinan, J.A. Winkler, and J.A. Andresen, 2020: Evaluation of multiple approaches to estimate daily solar radiation for input to crop process models. *Atmosphere*, **12,** 8, doi: 10.3390/atmos12010008

Pielke, R. A., Jr., and M. W. Downton, 2000: Precipitation and damaging floods: Trends in the United States, 1932–1997. *J. Clim*, 3625–3637. /Users/bjbaule/Library/Application Support/Zotero/Profiles/fo62ow39.default/zotero/storage/JFH5VDSS/summary.html.

Powell, E. J., and B. D. Keim, 2015: Trends in Daily Temperature and Precipitation Extremes for the Southeastern United States: 1948–2012. *J. Clim.*, **28**, 1592–1612, doi:10.1175/JCLI-D-14-00410.1. http://journals.ametsoc.org/doi/10.1175/JCLI-D-14-00410.1 (Accessed September 25, 2019).

Pryor, S. C., 2009: *Understanding climate change : climate variability, predictability, and change in the midwestern United States*. Indiana University Press, 296 pp. http://catalog.lib.msu.edu/search~/X?search=Understanding+climate+change+%3A+climate+variability%2C+predictability%2C+and+change+in+the+midwestern+United+States (Accessed September 25, 2019).

Pryor, S. C., J. A. Howe, and K. E. Kunkel, 2009: How spatially coherent and statistically robust are temporal changes in extreme precipitation in the contiguous USA? *Int. J. Climatol.*, **29**, 31–45, doi:10.1002/joc.1696. http://onlinelibrary.wiley.com/doi/10.1002/joc.1696/abstract.

Robertson, G. P., T. W. Bruulsema, R. J. Gehl, D. Kanter, D. L. Mauzerall, C. A. Rotz, and C. O. Williams, 2013: Nitrogen–climate interactions in US agriculture. *Biogeochemistry*, **114**, 41–70, doi:10.1007/s10533-012-9802-4. http://link.springer.com/10.1007/s10533-012-9802-4 (Accessed September 25, 2019).

Roque-Malo, S., and P. Kumar, 2017: Patterns of change in high frequency precipitation variability over North America. *Sci. Rep.*, **7**, 10853, doi:10.1038/s41598-017-10827-8. http://www.ncbi.nlm.nih.gov/pubmed/28924195 (Accessed April 27, 2018).

Rosenzweig, C., F. N. Tubiello, R. Goldberg, E. Mills, and J. Bloomfield, 2002: Increased crop damage in the US from excess precipitation under climate change. *Glob. Environ. Chang.*, **12**, 197–202, doi:10.1016/S0959-3780(02)00008-0. http://www.sciencedirect.com/science/article/pii/S0959378002000080.

Slater, A.G., 2016: Surface solar radiation in North America: a comparison of observations, reanalyses, satellite, and derived products. *J. Hydrometeorology*, **17**, 401-420, doi: 10.1175/JHM-D-15-0087.1

Syswerda, S. P., B. Basso, S. K. Hamilton, J. B. Tausig, and G. P. Robertson, 2012: Long-term nitrate loss along an agricultural intensity gradient in the Upper Midwest USA. *Agric. Ecosyst. Environ.*, **149**, 10–19, doi:10.1016/j.agee.2011.12.007. http://www.sciencedirect.com/science/article/pii/S0167880911004336.

Walsh, J., and Coauthors, 2014: Ch. 2: Our Changing Climate. Climate Change Impacts in the United States: The Third National Climate Assessment. doi:10.7930/J0KW5CXT. https://doi.org/10.7930/J0KW5CXT.

Winkler, J.A., 2016: Embracing complexity and uncertainty. *Annals of the American Association of Geographers*, **106**, 1418-1433, doi: 10.1080/24694452.2016.1207973

Westra, S., L. V. Alexander, F. W. Zwiers, S. Westra, L. V. Alexander, and F. W. Zwiers, 2013: Global Increasing Trends in Annual Maximum Daily Precipitation. *J. Clim.*, **26**, 3904–3918, doi:10.1175/JCLI-D-12-00502.1. http://journals.ametsoc.org/doi/abs/10.1175/JCLI-D-12-00502.1 (Accessed September 25, 2019).

Zhou, M., and K. Butterbach-Bahl, 2014: Assessment of nitrate leaching loss on a yield-scaled basis from maize and wheat cropping systems. *Plant Soil*, **374**, 977–991, doi:10.1007/s11104-013-1876-9. http://link.springer.com/10.1007/s11104-013-1876-9 (Accessed September 25, 2019).

CHAPTER 2. SITE DEPENDENT BIAS CORRECTION OF GROWING SEASON SOLAR RADIATION IN THE MIDWESTERN UNITED STATES

Abstract

Daily solar radiation data are an essential input to many modeling applications in the agricultural and ecological sectors. Aside from specialized networks it is often infrequently measured directly. Reanalysis data are often used in lieu of in-situ observations, despite known biases that vary across space. This study presents a method to correct mean and quantile specific biases in growing season solar radiation data from reanalysis, based on available in-situ observations and interpolated the correction to reanalysis solar radiation data across space. In the dataset implemented, biases were evident in both the mean and variance of the reanalysis solar radiation prior to correction. Following correction, reanalysis data displayed a better fit with in-situ observations. Impacts of the bias correction are illustrated using a process-based crop simulation model at multiple locations across the study region. Correction of reanalysis solar radiation alone brought simulated yield and water usage more in line with simulations forced with in-situ solar radiation, though the effect of differences in temperature between in-situ observations and reanalysis has substantial effect of simulated yields and water usage.

Introduction

Solar irradiance is one of the key components of the surface energy balance and credible solar radiation data are essential for many activities in our current society. One example is in agriculture, where water availability and usage are major constraints in crop production in many areas and assessing these often requires solar radiation data. Accurate solar radiation data are key for activities related to crop water availability such as irrigation scheduling and are key in both humid and arid environments (Ramírez et al. 2005). In model-based applications where solar radiation data is required, biases may lead to greater uncertainty and potential errors in

simulation of potential evapotranspiration (PET), evapotranspiration (ET), and crop biomass, among other modeled outputs.

In lieu of in-situ observations of solar radiation, which are not typically included in most federally-managed observing networks over much of the observed meteorological record (Perdinan et al. 2020), gridded observational and reanalysis datasets have become very useful in operations that require solar radiation data due to their spatial and temporal completeness. There have been numerous studies evaluating gridded observational and reanalysis datasets for biases. The majority of these (e.g. Briley et al. 2021; Cucchi et al. 2020; Briley et al. 2017; Dee et al. 2011) considered the principal variables of air temperature and precipitation and examined various methods for correcting the uncovered biases. Far fewer studies have examined the biases present in lesser observed variables such as solar radiation, or their suitability as inputs for impact models (Perdinan et al. 2020; Kiefer et al. 2019; Slater 2016; Weedon et al. 2011; Battisti et al. 2019). In many cases (e.g. Weedon et al. 2011, Kiefer et al. 2019), studies attempted to correct the biases inherent in these datasets. The corrective methods developed and employed for solar radiation data series vary in complexity and by the study area examined and include approaches ranging from simple factor scaling (Cosgrove et al. 2003; Ngo-Duc et al. 2005) to more complex regression-based methods. The more complex methods typically involve a regression based empirical model trained on a proxy variable or a more complex scaling method. Examples of proxy variables implemented include: observed cloud cover (Weedon et al. 2011), a multi-variate atmospheric driver approach (Wei et al. 2014), scaling to over/under estimation of cloudiness with clear sky radiation. Kiefer et al. (2019) used regression and hourly in-situ observations to correct hourly gridded data. These studies document different radiation biases, which are dataset dependent.

A number of recent studies have examined the impacts of historical climate variability and change on the spatial and temporal distributions of crop production across the Midwestern United States (Hunt et al. 2020; Jin et al. 2019; Zhang and Villarini 2019; Maestrini and Basso 2018). Biases in temperature and precipitation and their impacts on crop production have been well studied (Mourtzinis et al. 2017; Liu et al. 2020). However, few studies to date have directly considered the impact of bias in gridded solar radiation on crop production in the Midwest. Examining these biases and their impacts are essential as these data are often used to drive impact models used by decision-makers. Perdinan et al. (2020) evaluated multiple approaches to estimate solar radiation and their impacts on simulated maize and soybean yields, leaf area index (LAI), and crop evapotranspiration from process-based crop models for one location in Wisconsin. The authors found little difference in simulated yields between the datasets for maize, while soybeans showed significant differences. However, substantial differences in daily values within the maize simulations were apparent with different solar estimates. Overall, smaller biases were observed at Hancock, WI in satellite derived, mechanistic models, and empirical estimates of solar radiation, while reanalysis were found to be biased high in most months which supports findings from Slater (2016).

Slater (2016) highlighted the spatial variability in the bias in multiple gridded solar radiation datasets across the Midwest. Daily biases in the gridded solar observations varied across the region in both sign and magnitude. Kiefer et al. (2019) noted, in a sub-daily analysis of ET for the Great Lakes region, that NLDAS-2 was biased high during the peak daytime hours for solar radiation and developed hourly correction procedures. Both studies illustrate the complicated spatial and temporal nature of biases in the region and attempts to correct any biases. Studies from other regions (e.g. Battisti et al. 2019) indicate that gridded solar

18

observations may be appropriate for ingestion in process-based crop simulations and other impact models but again biases inherent in gridded datasets contribute to errors in modeled outputs and increase uncertainty in predictions.

The primary goal of this study is to further develop the existing body of work on solar radiation bias correction across the Midwest and demonstrate the potential impacts of such work on process-based crop model simulations. In particular we focus on correcting the known solar radiation bias during the growing season in NLDAS-derived solar radiation products (e.g. Abatzoglou, 2013) and develop a flexible, individual-variable, site-based method for correcting this behavior in gridded solar radiation datasets that can be applied to other regions and other datasets. In addition to the correction method outlined below, a process-based crop model simulation was implemented to illustrate the impacts of different combinations of weather data and known biases on simulated crop production and water use across the Midwestern United States.

Methodology

This study primarily builds upon the methodology developed by Kiefer et al. (2019) that focused on the bias-correction of hourly downward solar radiation data during the April-September warm season period for the Great Lakes region. Kiefer et al. (2019) implements a "one-size fits all" bias/variance correction approach based on pooled downward shortwave solar radiation data from the Climate Reference Network [CRN] (Hubbard et al. 2005). Given the complexities of hourly solar radiation data and correcting its behavior, one uniform correction for the region was deemed appropriate by the authors. However, for most model-based applications, daily data are often required, as opposed to hourly. Several modifications of the methods outlined in Kiefer et al. (2019) were needed to develop a similar method for daily data

19

that could account for spatial differences in the behavior of solar radiation biases in our larger study region.

For this study, we obtained CRN downward solar radiation data from twenty-six individual observing sites (Figure 2.1) to evaluate the biases of the NLDAS-derived gridMET solar radiation data (Abatzoglou 2013). Fewer locations from the CRN network were available across the study domain for evaluation prior to 2008 and after 2015 due to a number of station closures and moves, so the years from 2008-2015 were chosen as the period of record due to the relatively higher spatial and temporal coverage of the network during that time.
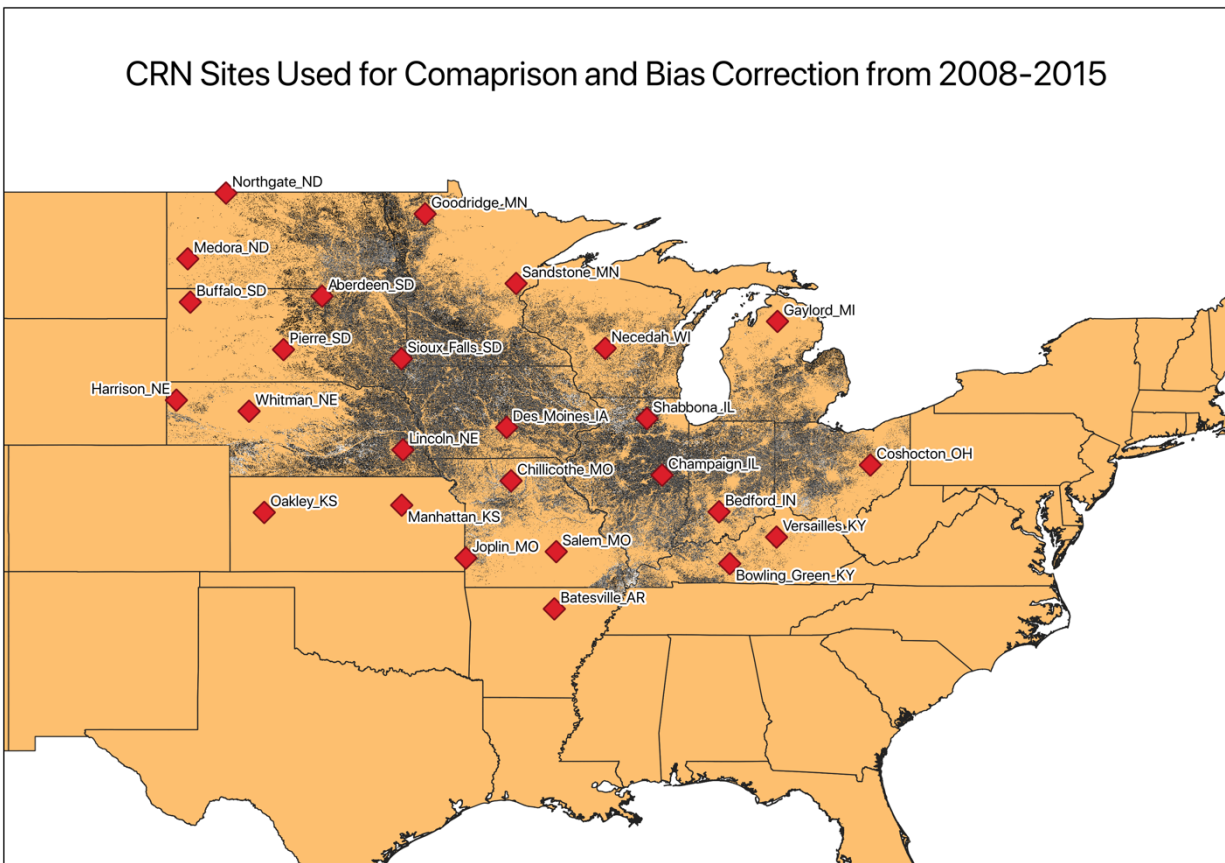
Figure 2.1 Location of CRN stations with acceptable temporal coverage, including solar radiation data from 2008-2015 over the study region. Sites are overlaid on map of thirty-meter pixels, identifying locations, with at least 3 years of corn or soybean production during 2008-2018 (NASS CDL).

CRN is widely considered to be one of the highest quality and most spatially consistent climate observing networks for solar radiation in the United States. Even though the data are generally high quality, we noticed some exceptionally large values that are beyond the range of plausible values at that location/time of year on individual days. Given this, we decided a quality control measure, in addition to those performed by National Centers for Environmental Information (NCEI), was needed. We calculated estimated clear-sky solar radiation based on Allen et al. (1998) as follows:

$$R_s = (0.75 + 2 * 10^{-5} * E) * R_a \qquad (1)$$

where Rs is estimated incoming clear-sky solar radiation (MJm^{-2}d^{-1}); E is elevation above sea level (m), and R$_a$ is calculated extraterrestrial solar radiation (MJm^{-2}d^{-1}). If a daily observation exceeded the estimated clear-sky value by more than 5%, the observed value was replaced with the estimated clear-sky radiation value for that day from Equation 1.

Bias Correction

Following the filtering/QC procedure outlined above, we then evaluated and corrected the growing season mean bias and the variance of the tails. The procedure outlined here is a modification on the method developed by Kiefer et al. (2019) that uses only observed solar radiation data to correct the gridded data, rather than multiple in-situ or remotely sensed data (e.g. Slater 2016). Our method differs significantly from Kiefer et al. (2019) in that we are correcting daily (as opposed to hourly) data and our methods are site-dependent as opposed to a "one-size-fits-all" approach for correction.

To determine the mean bias of the gridded solar radiation data relative to the in-situ solar radiation data, we calculated 2008-2013 average values of solar radiation for each day of the growing season (April-September). Following Kiefer et al. (2019), we estimated two linear

regression equations for each CRN site and their associated gridMET pixel value to yield a series

corrected for mean bias:

$$S_o = a + b * S_n \qquad (2)$$

where S_o is the eight-year day of year mean solar radiation for an individual station, a is the

intercept coefficient, b is the slope of the regression line, and S_n is the gridded eight-year day of

year mean solar radiation from the gridMET dataset. Using the a and b coefficients, we then

applied the following regression equation to yield a mean bias corrected gridded solar radiation

data series:

$$S_{bc} = a + b * S_n \qquad (3)$$

where a, b, and S_n are the same as equation 2, and S_{bc} is the resulting mean bias-corrected

gridMET solar radiation data. This correction was only applied to growing season dates, as at the

latitudes in the study region, most of the solar radiation is received during the growing season

and biases appeared concentrated during the growing season.

Distribution Fit Correction

Following the mean bias correction, another step is required to correct the overestimation

of values at the lower end of the growing season distribution and the underestimation of high

values at the higher end of the distribution following the mean bias-correction. This bias is

evident in the raw gridMET data and is further amplified by the mean bias correction. To better

fit the distribution of gridded solar radiation values, a distribution fitting method was developed

and implemented to correct this behavior in gridded solar radiation values. From the bias

corrected DOY gridMET solar means and the in-situ CRN means, the differences were binned

into percentiles (n) spanning the 2.5[th], 5th, 10th, 20th, 30th, 40th, 50th, 60th, 70th, 80th, 90th, 95th, 97.5[th]

percentiles. This can be difference can be expressed as a ratio for each percentile bin:

$$\frac{P_{n\,gridmet}}{P_{n\,CRN}} \qquad (4)$$

where $P_{n\,gridmet}$ is the individual percentile value from the bias corrected gridmet dataset and $P_{n\,CRN}$ is the individual percentile bin from the in-situ CRN data. The difference was noted and applied as a multiplier to the daily gridded values that fell within each percentile bin for the gridded dataset, yielding a dataset that is corrected for both mean bias and the over/underestimation behavior evident in gridMET solar radiation data at the site level.

IDW Interpolation

To increase the utility of this method across space we used an Inverse Distance Weighting (IDW) scheme in ArcGIS to develop a 4-km gridded surface to interpolate the regression coefficients and percentile correction ratios across space. IDW methods often require a priori knowledge of the characteristics of the data to make appropriate decisions on parameters. In the case of ArcGIS, the Power/Distance Weighting Parameter is key in determining the characteristics of the resulting surface. Weights between 1 and 6, incremented by one, were tested and based on the relative sparseness of the CRN stations relative to the grid spacing a Power/Distance Weighting of 4 was selected (Figure 2.2).
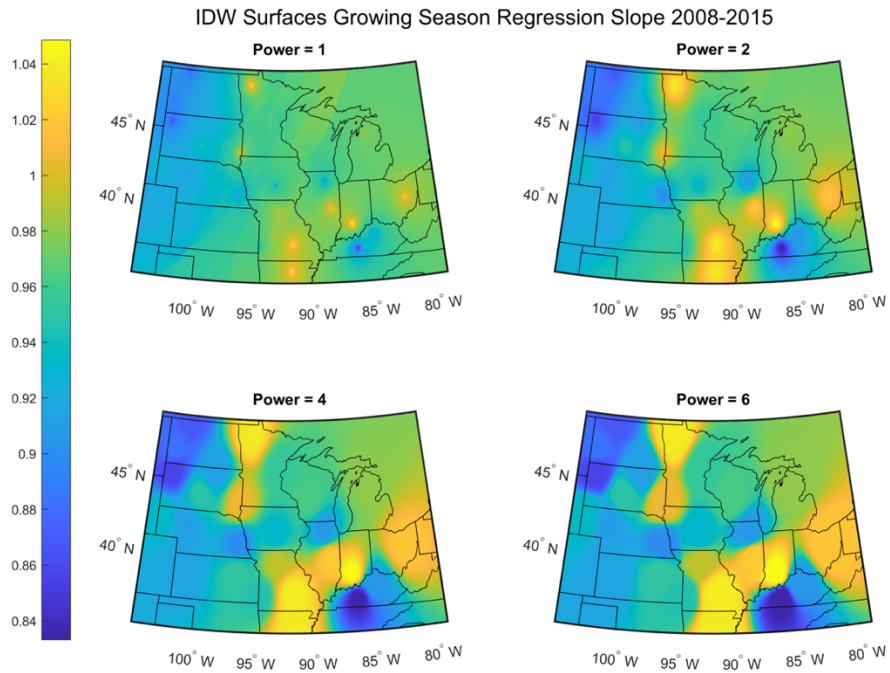
Figure 2.2 IDW surfaces of the regression slope in the bias correction step with different distance weight/power values.

Kiefer et al. (2019) also noted a distinct bias in solar radiation NLDAS-2 surrounding the Great Lakes, which presents as abnormally high solar radiation values that follow political boundaries in areas adjacent to the Great Lakes. In the case of our derived surfaces, we implemented a filtering algorithm to identify pixels in a bounding box around the Great Lakes (latitude > 40.96°N, longitude < 92.3°W) that exhibit the behavior of the additional bias. To identify pixels in the artifact, we standardized the bias corrected gridMET mean gride by z-scores and identified pixels with standardized growing season solar radiation means were greater than +0.20. Pixels identified as part of the bias were filled with values from their longitudinal nearest neighbor that fell under the +0.20 normalized threshold.

SALUS Experiments

To evaluate the effects of the correction on crop growth and water use we implemented a process based agricultural model, System for Agricultural Land Use Sustainability (SALUS) (Basso and Ritchie 2015) to conduct a series of simulations to examine the different combinations of climatic data inputs on simulated dry grain weight (GWAD), ET, Drought Stress (Drght_Fac). These experiments were performed at five locations within our study region that exhibit different energy balance regimes ranging from moisture-limited environments in the western areas of the region to more energy-limited environments in the east. Simulations were performed using several combinations of weather data as inputs for the simulations (Table 2.1). Each combination of weather data has standing in the literature and are common practice in agricultural modeling studies. Particularly for solar radiation, it is common practice, in lieu of in-situ solar observations, to substitute gridded solar radiation data and combine this with in-situ temperature and precipitation data (Baule et al. 2017; Gunn et al. 2018; Battisti et al. 2019).

Table 2.1 Combinations of weather inputs used in SALUS simulations.

Dataset Abbreviation	Temperature Dataset	Precipitation Dataset	Solar Dataset
CRN	CRN	CRN	CRN
GM	raw gridMET	raw gridMET	raw gridMET
GM_BC_DF	raw gridMET	raw gridMET	BC/DF gridMET
GM_CRN	CRN	CRN	BC/DF gridMET

The experiment setup simulated continuous corn at both locations from 2009-2015. Simulations were conducted with water-stress and no-nutrient stress, to better simulate a well-managed cropping system. Cultivar parameters were obtained from model developers (Basso 2019, personal communication). Parameters used are representative of county-level maize cultivars across the Midwestern United States. Representative county soils were obtained from gSSURGO, and the majority soil type associated with agricultural production in the county was

used. Planting date was determined by the model as a date during a specified planting window, where certain soil temperature and moisture thresholds are met. Spring tillage prior to planting was also simulated.

Results

Dataset Quality of Control of CRN Observations

Of the twenty-six stations shown in Figure 2.1, only 6 stations had a substantial number (>= 100 days from 2008-2015) of datapoints that exceeded the +5% of estimated clear sky radiation (Table 2.2). An example from the Sandstone, MN CRN site is shown in Figure 2.3a. In general, the problematic data points at each location occurred prior to 2012 and were concentrated in a specific block of time at each location. Data points that exceeded the estimated 5% threshold were excluded from the calculations for the bias correction procedure. Otherwise, the CRN data were not changed from their original values. As seen in the Figure 2.3, the point values generally show a larger daily range and annual cycle when compared to gridded solar radiation data. This difference is notable but is expected due to the difference between a 4km grid cell and a point in-situ observation. Despite the differences, the annual cycle and daily variability of solar radiation is well represented across the region by gridMET when compared to in-situ CRN observations. CRN and gridMET solar radiation time series post-filtering are shown in Figure 2.3b.

Table 2.2 Stations with more than 100 days where observed radiation was greater than 5 percent above estimated clear sky radiation and number of days at each location.

Site	Days > 5% above estimated clear sky radiation
Aberdeen, SD	193
Sandstone, MN	189
Necedah, WI	169
Champaign, IL	139
Bowling Green, KY	136
Goodridge, MN	121
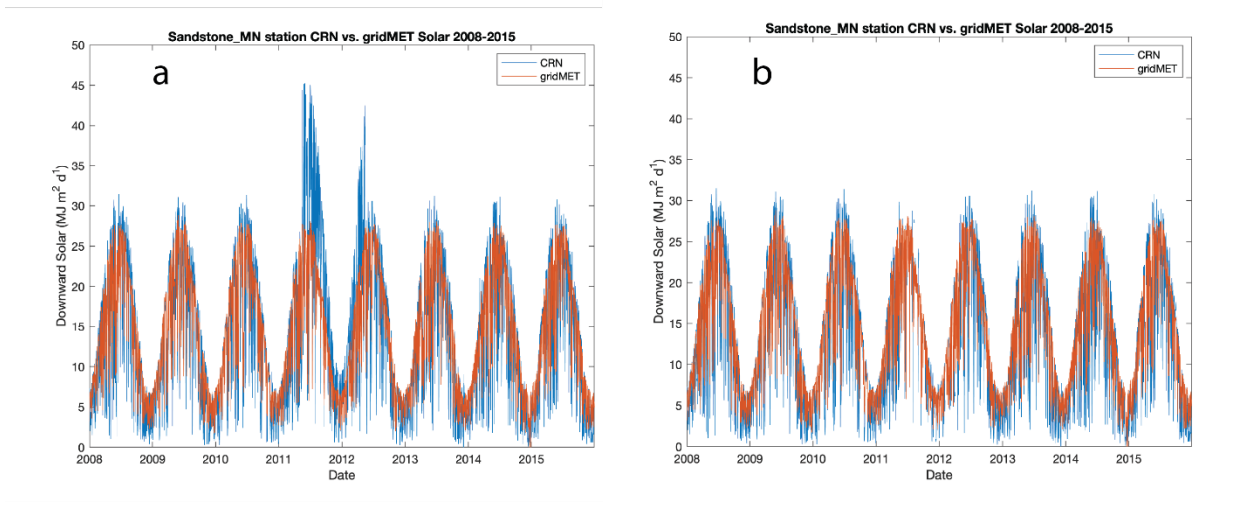
Figure 2.3. Time series from Sandstone, MN (92.99°W, 46.11°N) of solar radiation from unfiltered CRN station observations (blue) and closest gridMET data point (orange). The left side (a) plot shows all days, and the right side (b) shows all days following filtering.

Bias in the gridMET Long-Term Mean

Following the filtering of anomalously high solar radiation values at the CRN stations, the differences in between the means of the annual, growing season, and non-growing season solar radiation were calculated for both gridMET and CRN solar radiations values (Figure 2.4). Across the region, twenty five of the twenty six stations showed a positive bias in gridMET solar radiation values when compared to the CRN observations . The maximum positive gridMET bias

27

in annual solar radiation was at Gaylord, MI (+1.49 MJ m^{-2} d^{-1}). The only location in the study region that showed a negative (-0.26 MJ m^{-2} d^{-1}) annual gridMET bias (where CRN was higher than gridMET) was Whitman, NE in the west-central area of the study region. All stations showed a positive bias in the mean difference during the growing season, though the magnitude varied geographically, with Batesville, AR (+2.37 MJ m^{-2} d^{-1}) showing the largest growing season mean bias and Whitman, NE (+0.01 MJ m^{-2} d^{-1}) showing the smallest. Geographically, the differences are scattered station to station, however the three largest differences between the CRN observations and gridMET values are observed at stations in Arkansas and Missouri, while the stations with the smallest differences in the growing season means occurred in Minnesota, South Dakota, and Nebraska. Although the differences in the non-growing season means are not the primary focus of this study, there are substantial regional differences that should be commented on. The largest positive differences between the two datasets for non-growing season solar radiation were observed at Gaylord, MI (+1.66 MJ m^{-2} d^{-1}) and Goodridge, MN (+1.10 MJ m^{-2} d^{-1}). Negative biases, where the gridded data are biased low compared to the observations were located in the Great Plains states, with the largest negative biases being observed at Whitman, NE (-0.54 MJ m^{-2} d^{-1}) and Oakley, KS (-0.40 MJ m^{-2} d^{-1}). Other stations with negative biases during the non-growing season are located in Kansas, Nebraska, and South Dakota. Still, the majority (19/26) of stations during the non-growing season also exhibit positive biases. The only stations where positive non-growing season biases were larger in magnitude than growing season biases were Gaylord, MI and Goodridge, MN. The stations where negative non-growing season biases are larger than the magnitude of the bias during growing season were Harrison, NE, Oakley, KS, and Whitman, NE.
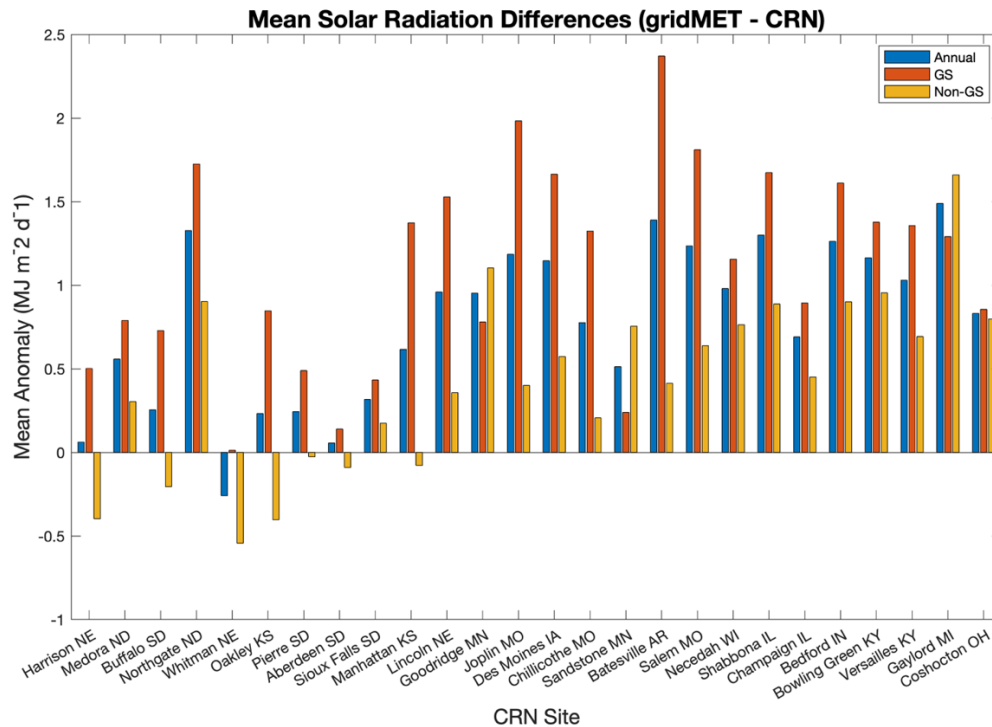
Figure 2.4 Mean differences (MJ m^{-2} d^{-1}) between filtered CRN solar radiation observations and gridMET solar radiation values for the period from 2008-2015. Blue bars indicate the Annual (January-December), orange bars indicate Growing Season (April-September), and yellow bars indicate Non-Growing Season (October-March) differences in solar radiation.

Bias Correction/Distribution Fit Correction

As shown here, the growing-season means obtained from gridMET were biased high at all 26 CRN locations. This bias varied substantially between stations (Figure 2.4) and our first step is to correct the mean bias following equations 1 and 2 for growing season days. A visual example of the nature of the correction is shown in Figures 2.5 and 2.6 for Coshocton, OH (81.78°W, 40.37°N).
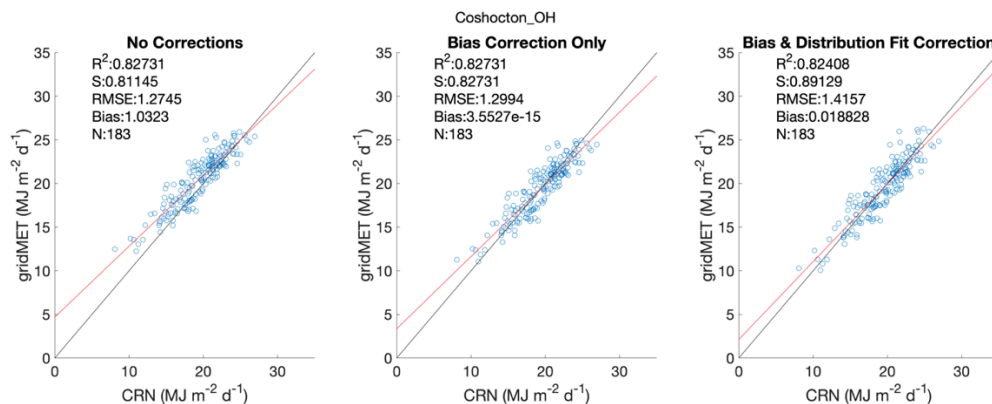
Figure 2.5 Scatterplots and regression statistics of gridMET (y-axis) and CRN (x-axis) solar radiation growing season DOY means for Coshocton, OH at each stage in the correction process. Black line is the 1:1 line and the red line is the least-squares regression fit for the data.
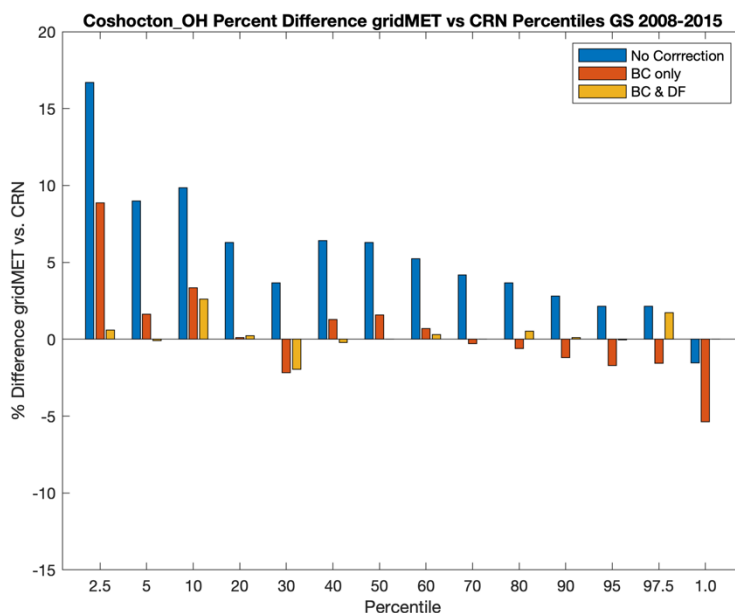
Figure 2.6 Percent differences between gridMET and CRN for Coshocton, OH growing season solar radiation across the selected percentile bins. Percentile values are shown for gridMET data with no correction (blue bars), only the bias correction is performed (red bars), and data after bias correction and distribution fit (yellow bars).

The sign of the growing season solar radiation bias across the region was similar among locations, though the magnitude of the biases varied from location to location. In general, during the growing season, gridMET solar radiation values are biased high not only at the mean, but the behavior of the biases also depends on the strength of the incoming solar radiation. Days with

30

lower mean solar radiation values tend to be overestimated more frequently at the lower tail of

the distribution than the mean. Higher solar radiation values are underestimated in gridMET

when compared to the CRN observations, although the magnitude is generally smaller than the

overestimation at the low end of the distribution. When a standard bias correction is applied to

gridMET data to match those of the mean of the CRN observations, the result is a dataset with a

matched mean (note: elimination of Bias between the left and middle plot of Figure 2.5).

However, the behavior of the tails is modified as well, generally forcing the smaller values closer

to the mean, but still overestimated. At the high end of the distribution, the application of the

mean correction further amplifies or introduces underestimation bias in values (Figure 2.5,

middle). In terms of fit statistics, the R^2 is essentially unchanged, S is improved, the RMSE is

degraded slightly, and the mean Bias (as previously mentioned) is essentially eliminated.
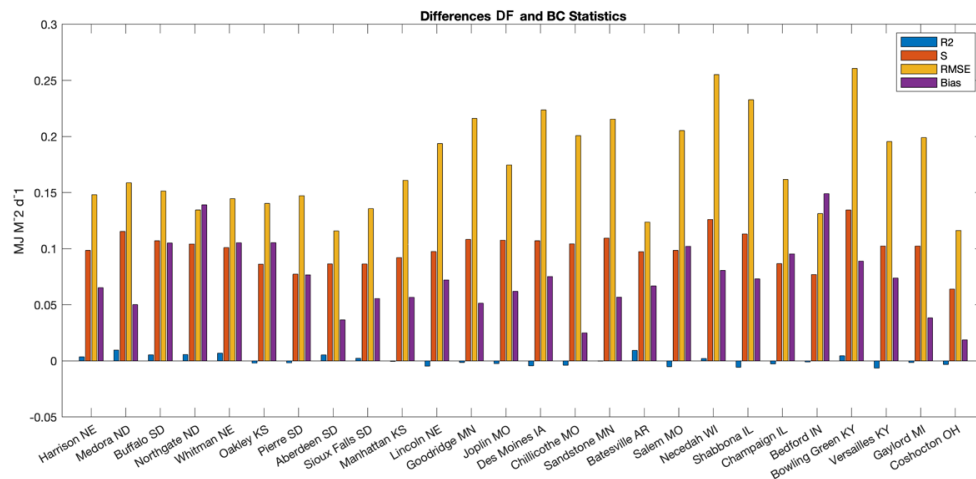
Figure 2.7 Differences in coefficient of determination (R^2), slope fit (S), root mean squared error
(RMSE), mean bias (Bias) between the mean distribution fit/bias corrected (DF) dataset and the
mean bias correction (BC) only dataset. Stations are roughly ordered from west to east.

To correct the behavior across the distribution of growing season solar radiation, the

distribution fitting technique outlined in Eqn. 4 was implemented to further adjust the gridMET

data to fit the nature of the CRN observations (Figure 2.6 & 2.7). In the case of the Coshocton,

OH site, distribution fitting of the chosen percentiles results in a dataset that is much closer to a 1:1 slope (S) with the CRN observations than the uncorrected or simply mean bias corrected dataset. The R^2 is degraded slightly at some locations by the distribution fitting, a slight mean bias is reintroduced at all locations and RMSE is slightly increased. However, the dataset is a much better fit (S) across the range of data following correction than the untreated data. Geographically the largest changes between the DF and BC data series are at locations where there were large biases evident at the tails of the solar radiation distribution. This behavior was more evident at locations in southern and eastern sections of the study region.

IDW Interpolation of Bias Correction and Distribution Fit Method

Eighteen surfaces (using a power weighting of 4) were generated across the study region (Ex. Figure 2.2). These surfaces are based on the regression coefficients and percentile corrections as determined in Eqns. 2-4 from the CRN and gridMET data for the individual station locations. Interpolated corrections were then applied on a daily basis to each grid cell within the study region to create daily time series of corrected solar radiation from 1979-2019. Surface plots of mean growing season radiation from both the uncorrected gridMET and the BC/DF gridMET data are shown for comparison (Figure 2.8).
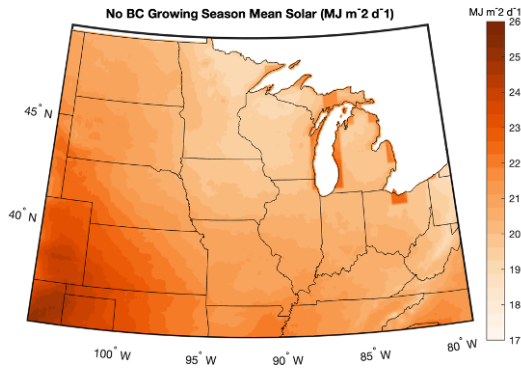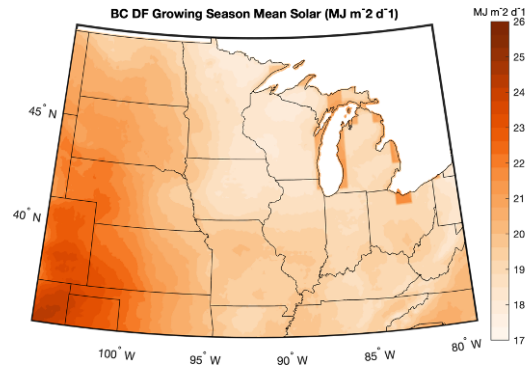
Figure 2.8 (a) Mean growing season radiation from 1979-2018 for the study region obtained from daily uncorrected gridMET values and (b) bias-corrected/distribution fit gridMET based on the corrections derived from 2008-2015 CRN observations.

Visually, between the two plots, the BC/DF gridMET solar radiation shows lower solar radiation values across the study region when compared to the uncorrected gridMET solar radiation. As expected with IDW interpolation, the largest differences in terms of $MJm^{-2}d^{-1}$ are in the southern and eastern portions of the study region. Percentage differences between the two surfaces are shown in Figure 2.9. As with any interpolation with a relatively sparse spatial network of observations, there appears to be some artifacts relative to station density (e.g. southern Ohio, Kentucky, Arkansas). Due to the sparseness of the CRN network in some areas of the study region, it is difficult to discern whether these are artifacts introduced by the interpolation or if these corrections are amplifying already existing features in the dataset (Figure 2.9).
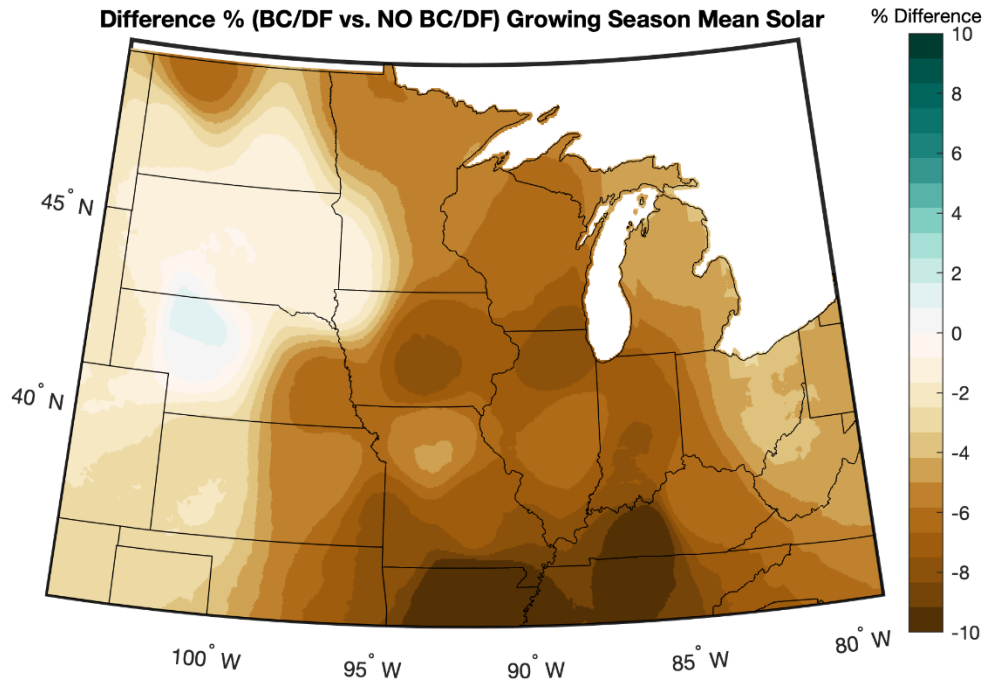
Figure 2.9 Percentage difference between BC/DF gridMET and uncorrected gridMET growing season mean solar radiation from 1979-2018.

In Figure 2.8, the artifact around the Great Lakes region that exhibits anomalously high solar radiation values is visually evident in both the uncorrected and BC/DF gridMET data. Following the identification and nearest neighbor interpolation outlined to reduce the impact of this artifact on data in the impacted cells, the growing season solar radiation grids show a substantially reduced artifact around the lakes (Figure 2.10) and the resulting grid is more in line with geographical expectations of highest solar radiation values observed in the southwest portion of the grid decreasing eastward and northward across the region.
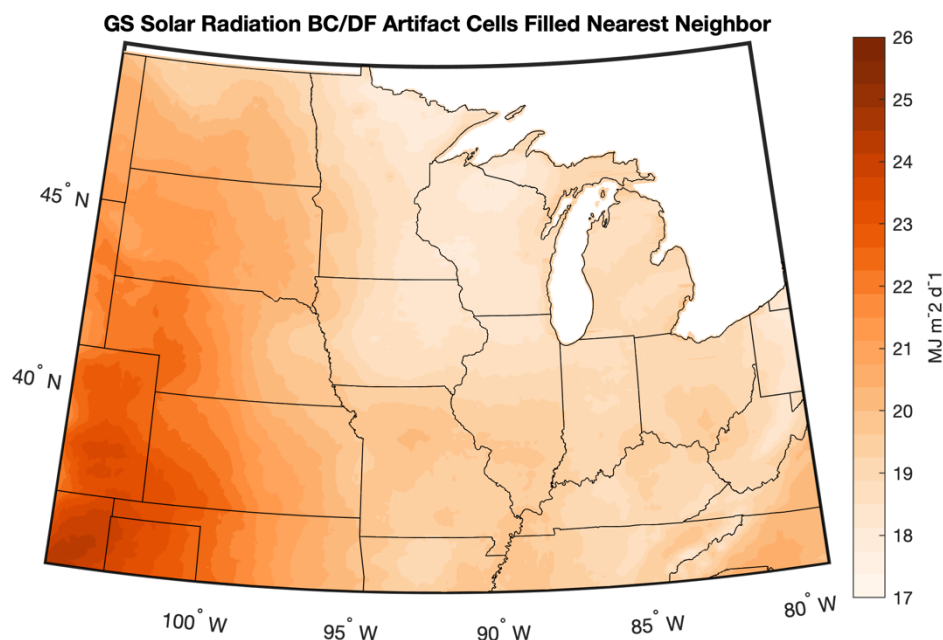
Figure 2.10 Mean growing season solar radiation from BC/DF gridMET with Great Lakes artifact filled.

SALUS Crop Model Experiments

The use of process-based crop simulation models allows us to evaluate the effects of different combinations of weather inputs on a range of potential crop production system impacts including yield and water use/stress. As a variable in the numerator of the equations that estimated PET, such as the Penman-Monteith or Priestly Taylor estimation methods, solar radiation plays a direct, linear role through the net radiation variable in determining potential evapotranspiration in most crop simulation models (Basso and Ritchie, 2018). Theoretically, each additional Megajoule per square meter of net radiation incident on the surface results in up to 0.4mm of additional potential evapotranspiration per square meter, which in turn impacts the amount of water available for a number of other crop growth and production cycles and processes. In Figure 2.1,1 simulated crop yields and growing season temperature differences between CRN and gridMET from 2009 through 2015 are shown. Since the goal of this research

does not directly include the bias correction of gridded temperature data, the variable behavior of

the temperature differences between point and grid are important to note as they may also

influence the simulated yields and water use. Still, the effect of the solar correction is evident at

both locations shown in Figure 2.11.
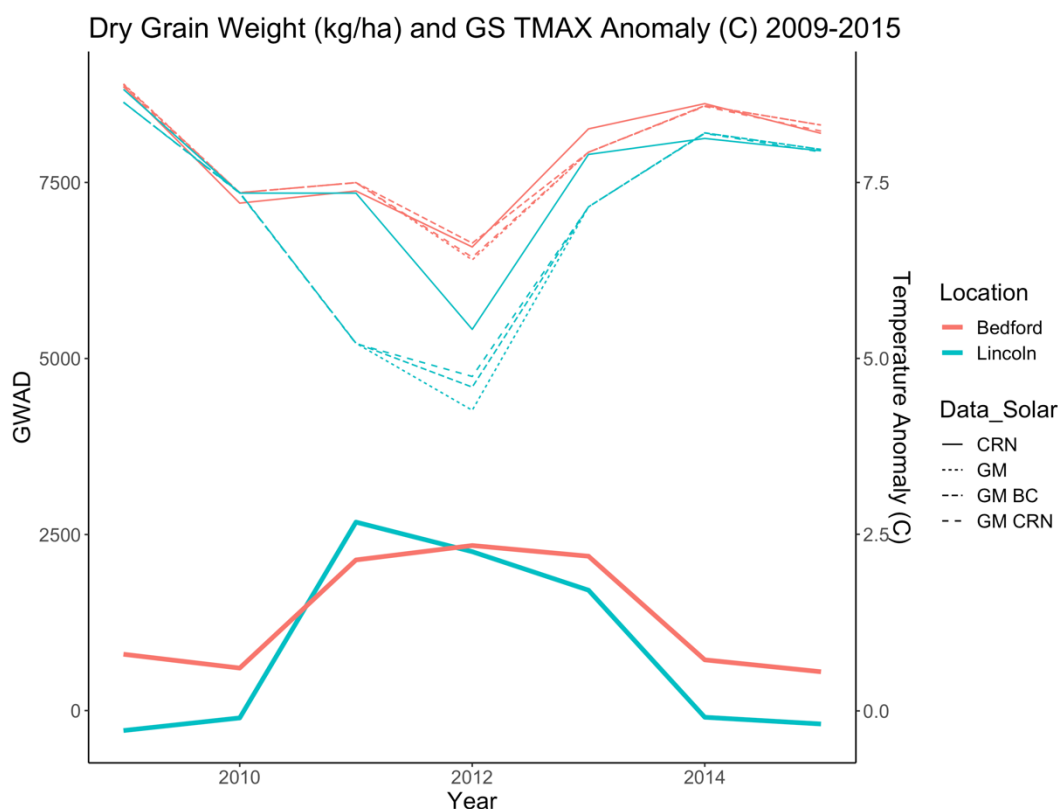
Figure 2.11 Simulated dry grain weight for maize simulations using different combinations of gridded and point-based temperature solar radiation and temperature observations and differences between gridMET values and CRN temperature observations. Data Key: CRN = CRN Temp, Precip, Solar; GM = raw gridmet Temp, Precip, & Solar; GM_BC = raw gridmet, Temp, Precip, corrected Solar, GM _CRN = gridmet Temp & Precip with CRN solar directly substituted for gridMET.

We see that yields can vary substantially with different combinations of weather inputs.

Using the four combinations of weather series, we can isolate the effect of solar radiation on crop

yield. There are differences between simulated crop yields for all four input series, the largest

differences appearing in 2012, which was a significant drought year at both locations. At

Lincoln, NE (a more moisture-limited location), the yields simulated in 2012 were lower than all previous and subsequent simulation years. With Lincoln's solar radiation being overestimated by gridMET and a substantial warm bias in the 2012 season, simulations run with raw gridMET output showed the lowest yields in 2012 (4265 kg/ha). Simulations run with the corrected gridMET solar radiation (4594 kg/ha) and CRN solar radiation with gridMET temperature and precipitation were similar (4745 kg/ha). Data run with CRN in-situ observations yielded highest in 2012 (5414 kg/ha). In the case of Bedford and Lincoln, the temperature biases led to larger yield differences than the differences in solar radiation. However, simulations forced with corrected gridMET solar radiation are in relatively good agreement with runs forced with gridMET temperature and precipitation and CRN solar radiation, suggesting that our correction method does improve results when corrected gridded solar radiation are used in place of raw gridMET data. As these results demonstrate, individual bias correction of other key variables needed in crop models is also important to ensuring accurate results and caution must be exercised by modelers incorporating gridded meteorological data in lieu of in-situ observations.

Drought Stress Days

Simulation results indicate that each location experienced substantial drought stress across the region during the study period of record. Results showing annual and cumulative totals over the 5 simulation growing seasons are shown in Figure 2.12. The geographical pattern of drought stress over the 5 years simulations window (Figure 2.12b) with the highest drought stress values at the two westernmost locations (Aberdeen and Lincoln) decreasing to the east across the region with lowest values at Necedah, Bedford, and Chillicothe. Though the means follow geographical expectations with a general decrease of mean annual precipitation from east to west across the region, there is substantial interannual variability at all locations (Figure

2.12a). Maximum annual values of drought stress occurred in 2012 at all locations. The overall

range of drought stress days among all locations and years was from 0 days to 50 days of

simulated drought stress in a given growing season.
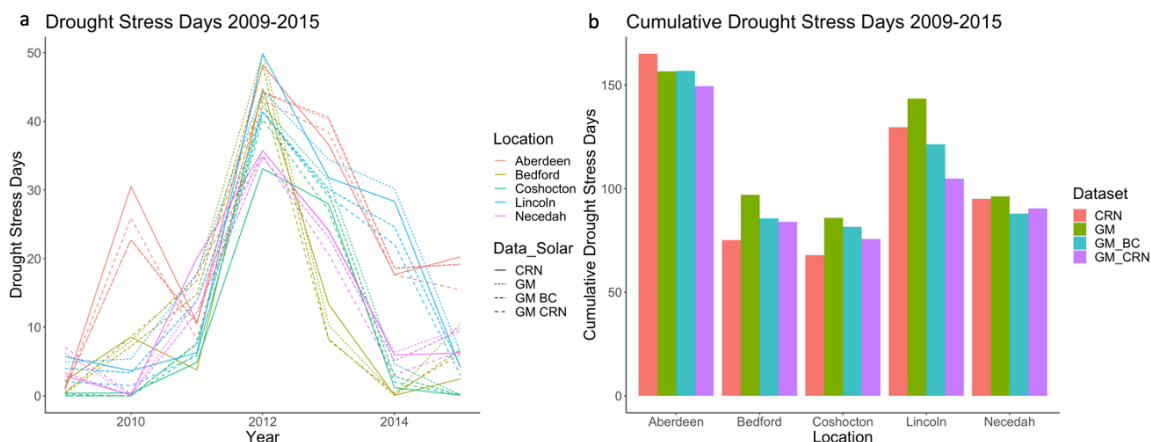
Figure 2.12 Annual (a) and Cumulative Totals (b) of Drought Stress Days as simulated by
SALUS with various combinations of meteorological data used as inputs.

The weather series used to force the SALUS simulations also had a substantial impact on

simulated drought stress. At four out of the five locations, simulated drought stress was lower

when CRN observations were used than the raw gridMET data. At all five locations the relative

difference between simulations forced CRN only and raw gridMET in simulated drought stress

was similar in magnitude to the difference in solar radiation between the two datasets. At

Aberdeen, where gridMET solar radiation was not biased high during growing season, drought

stress under CRN conditions is accordingly higher than under raw gridMET conditions. The

effect of the solar radiation correction on simulated drought stress is such that, without additional

bias correction for temperature and precipitation, the corrected gridMET solar radiation falls in

the middle between the two unaltered datasets.

Table 2.3 Difference in cumulative simulated Eta and Etp at Lincoln, NE using different combinations of gridded and point-based temperature and solar radiation data.

Station	Dataset Temp/Precip	Dataset Solar	Difference gridded vs raw CRN %	Différence gridMET T/P CRN Solar %	Difference gridded vs raw CRN %	Difference gridMET T/P with CRN Solar %
Lincoln, NE			Eta	Eta	Etp	Etp
	CRN	CRN	0.00	-1.48	0.00	-1.32
	gridMET raw	gridMET raw	5.36	3.81	6.42	5.02
	gridMET raw	gridMET BC/DF	3.29	1.77	2.06	0.72
	gridMET raw	CRN	1.50	0.00	1.33	0.00

Simulated potential and actual evapotranspiration for Lincoln, NE are shown in Table 2.3. With additional bias correction of temperature and precipitation simulations, the gridded data results could be brought closer in line with in-situ observations. Although the percentage improvements in terms of simulated actual evapotranspiration and potential evapotranspiration are modest in percentage terms, the differences are consistent and over long simulation windows (i.e. 10 + years) could compound and result in differences in simulated water balances over time. Ideally, additional bias correction of temperature and precipitation is not a trivial task and the method employed in additional bias correction would likely need to be multivariate and consider the interactions between the driving meteorological variables not simply addressing each bias on its own (Maraun et al. 2017).

Discussion/Conclusions

Solar radiation is one of the key variables in understanding the water and energy balance of a wide array of systems across a range of scales. As shown by previous research, gridded meteorological data often contains several biases when compared to observations that can alter

the conclusions drawn from impact models used by researchers and decision-makers. Each gridded dataset is unique and biases need to be assessed prior to incorporation in impact models. This study examines the bias in gridMET solar radiation data in comparison to point stations from the CRN network. The majority of the twenty-six stations analyzed for bias against gridMET solar radiation showed a positive bias in the gridMET solar radiation, with greater biases in the south and east of the study region.

A flexible bias correction/distribution fit correction method was incorporated on a site-by-site basis and applied to gridded growing season solar radiation values only. The bias at each CRN station location was first corrected through a mean-bias correction, which improved the statistics towards the mean/median of the dataset but introduced or exacerbated existing biases in the tails of the distribution. To correct the bias in the tails following mean-bias correction, a distribution fit procedure was developed to correct the behavior. Following the distribution fit procedure, the fit at each site was better across the whole distribution of solar radiation values (S), with the improvements in S ranging from 0.06 at Coshocton, Ohio to 0.13 at Necedah, Wisconsin. As these corrections are location specific, IDW was used to generate regional grids of the correction coefficients and applied corrections were applied on a grid cell by grid cell basis across the region. Following interpolation, cleaning procedures were applied to handle anomalously high radiation values in an artifact affecting the Great Lakes region. As NARR-derived datasets are well-represented in the body of literature regarding agriculture and climate in the United States (Perdinan et al. 2020; Kiefer et al. 2019; Slater 2016; Ho et al. 2018), gridMET seems an ideal illustrative example though conceptually this approach could be extended to any number of gridded datasets. Ideally, efforts to correct meteorological data such as solar radiation include a cross-validation (e.g. Wang et al. 2018) to evaluate the effectiveness

of bias correction methodology by subjecting data not used in the training to the correction model and evaluating the performance of the model on independent data. While the authors understand the importance of this in most cases, the authors determined that the period of record available for CRN with sufficient spatial density was too short for a true cross-validation approach to be used. When compared with other studies that deal with correction of gridded solar radiation data in our study region (namely Kiefer et al. 2019; Slater 2016), our corrected grids are similar in appearance and improvement over uncorrected grids. If a more spatially robust and homogenous network of solar radiation observations were to exist across the study region, our approach would likely produce a closer fit to observations than previous methods applied to correct growing season solar radiation in our region. However, current limitations regarding solar radiation observations and heterogeneities between existing observing networks make this an ongoing and difficult task.

Incorporating the corrected weather data into a process-based crop model (SALUS) illustrates the sensitivity of crop models to different combinations of weather data. The warm bias in many growing seasons from 2009-2015, particularly 2012 and 2013, and the high bias in the raw gridMET data typically led to lower yields and the highest water use over the seasons simulated. CRN was generally associated with the highest yields and lower water demands in the crop model simulations, while the corrected gridMET was generally in between simulations forced with in-situ CRN or raw gridMET data. Simulations forced with CRN solar and gridMET temperature and precipitation were closest to simulations forced with BC/DF gridMET solar. If simulations were to be run over longer time horizons (i.e. 10+ years), the difference in water use between simulations could amount to approximately one-year's worth of precipitation at many locations. Uncertainty in crop models, gridded meteorological data, and climate projections are

all common sources of frustration for researchers and end-users in a broad array of sectors (Asseng et al. 2015; Basso et al. 2015; Xiong et al. 2020; Tao et al. 2020). In this regard, many previous studies have addressed uncertainties associated with temperature input data and to a lesser extent precipitation. Much less research has concerned solar radiation. Our method provides an examination of the relative effects of solar radiation on crop yield and drought stress; temperature and precipitation biases also need to be considered before results from crop models are used in an operational capacity.

Looking towards the future there are several opportunities for additional research. Incorporating additional observational sites from non-Federal networks (i.e. state mesonets) that observe solar radiation would improve the spatial representativeness of the observational sites from which the correction coefficients are generated (e.g. Slater 2016). In addition, testing of this method on other gridded solar radiation datasets would provide more a comprehensive understanding of the biases present in gridded meteorological observations in our study region and their potential impacts on agricultural systems.

REFERENCES

REFERENCES

Abatzoglou, J. T., 2013: Development of gridded surface meteorological data for ecological applications and modelling. *Int. J. Climatol.*, **33**, 121–131, doi:10.1002/joc.3413. http://doi.wiley.com/10.1002/joc.3413 (Accessed September 25, 2019).

Allen, R. G., L. S. Pereira, D. Raes, and M. Smith, 1998: *Crop evapotranspiration -Guidelines for computing crop water requirements -FAO Irrigation and drainage paper 56*. 15 pp. https://s3.amazonaws.com/academia.edu.documents/40878584/Allen_FAO1998.pdf?AWS AccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1517432475&Signature=VIWgU uV3imzGaeS3pIJsh3PKkiA%3D&response-content-disposition=inline%3B filename%3DAllen_FAO1998.pdf (Accessed January 31, 2018).

Asseng, S., and Coauthors, 2015: Rising temperatures reduce global wheat production. *Nat. Clim. Chang.*, **5**, 143–147, doi:10.1038/nclimate2470.

Basso, B. and J. T. Ritchie, 2018: Evapotranspiration in high-yielding maize and under increased vapor pressure deficit in the US Midwest. *Agricultural & Environmental Letters*, **3**, 170039, doi: 10.2134/ael2017.11.0039

Basso, B., and J. Ritchie, 2015: Simulating crop growth and biogeochemical fluxes in response to land management using the SALUS model. *The Ecology of Landscapes: Long-Term Research on the Path to Sustainability*, S. Hamilton, J. Doll, and G.P. Roberston, Eds., Oxford University Press, New York, New York USA, 252–274.

Basso, B., D. W. Hyndman, A. D. Kendall, P. R. Grace, and G. P. Robertson, 2015: Can Impacts of Climate Change and Agricultural Adaptation Strategies Be Accurately Quantified if Crop Models Are Annually Re-Initialized? *PLoS One*, **10**, e0127333, doi:10.1371/journal.pone.0127333. http://www.ncbi.nlm.nih.gov/pubmed/26043188 (Accessed September 25, 2019).

Battisti, R., F. D. Bender, and P. C. Sentelhas, 2019: Assessment of different gridded weather data for soybean yield simulations in Brazil. *Theor. Appl. Climatol.*, **135**, 237–247, doi:10.1007/s00704-018-2383-y. https://doi.org/10.1007/s00704-018-2383-y (Accessed March 3, 2021).

Baule, W., B. Allred, J. Frankenberger, D. Gamble, J. Andresen, K. M. Gunn, and L. Brown, 2017: Northwest Ohio crop yield benefits of water capture and subirrigation based on future climate change projections. *Agric. Water Manag.*, **189**, doi:10.1016/j.agwat.2017.04.019.

Briley, L. J., W. S. Ashley, R. B. Rood, and A. Krmenec, 2017: The role of meteorological processes in the description of uncertainty for climate change decision-making. *Theor. Appl. Climatol.*, **127**, 643–654, doi:10.1007/s00704-015-1652-2. http://dx.doi.org/10.1007/s00704-015-1652-2.

Briley, L. J., R. B. Rood, and M. Notaro, 2021: Large lakes in climate models: A Great Lakes case study on the usability of CMIP5. *J. Great Lakes Res.*, **47**, 405–418, doi:10.1016/J.JGLR.2021.01.010.

Cosgrove, B. A., and Coauthors, 2003: Real-time and retrospective forcing in the North American Land Data Assimilation System (NLDAS) project. *J. Geophys. Res. Atmos.*, **108**, 2002JD003118, doi:10.1029/2002JD003118. https://onlinelibrary.wiley.com/doi/abs/10.1029/2002JD003118 (Accessed February 22, 2021).

Cucchi, M., G. P. Weedon, A. Amici, N. Bellouin, S. Lange, H. Müller Schmied, H. Hersbach, and C. Buontempo, 2020: WFDE5: Bias-adjusted ERA5 reanalysis data for impact studies. *Earth Syst. Sci. Data*, **12**, 2097–2120, doi:10.5194/ESSD-12-2097-2020.

Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.*, **137**, 553–597, doi:10.1002/qj.828. http://doi.wiley.com/10.1002/qj.828 (Accessed January 31, 2018).

Gunn, K. M., W. J. Baule, J. R. Frankenberger, D. L. Gamble, B. J. Allred, J. A. Andresen, and L. C. Brown, 2018: Modeled climate change impacts on subirrigated maize relative yield in northwest Ohio. *Agric. Water Manag.*, **206**, doi:10.1016/j.agwat.2018.04.034.

Ho, S.-T., J. E. Ifft, B. J. Rickard, and C. G. Turvey, 2018: Alternative Strategies to Manage Weather Risk in Perennial Fruit Crop Production. *Agric. Resour. Econ. Rev.*, 1–25, doi:10.1017/age.2017.29. https://www.cambridge.org/core/product/identifier/S1068280517000296/type/journal_article (Accessed October 21, 2018).

Hubbard, K. G., X. Lin, and C. B. Baker, 2005: On the USCRN temperature system. *J. Atmos. Ocean. Technol.*, **22**, 1095–1101, doi:10.1175/JTECH1715.1. https://journals.ametsoc.org/view/journals/atot/22/7/jtech1715_1.xml (Accessed February 22, 2021).

Hunt, E. D., H. E. Birge, C. Laingen, M. A. Licht, J. McMechan, W. Baule, and T. Connor, 2020: A perspective on changes across the U.S. Corn Belt. *Environ. Res. Lett.*, **15**, 071001, doi:10.1088/1748-9326/ab9333. https://doi.org/10.1088/1748-9326/ab9333 (Accessed March 8, 2021).

Jin, Z., S. V. Archontoulis, and D. B. Lobell, 2019: How much will precision nitrogen management pay off? An evaluation based on simulating thousands of corn fields over the US Corn-Belt. *F. Crop. Res.*, **240**, 12–22, doi:10.1016/J.FCR.2019.04.013. https://www.sciencedirect.com/science/article/pii/S0378429018314369 (Accessed September 25, 2019).

Kiefer, M. T., J. A. Andresen, D. Doubler, and A. Pollyea, 2019: Development of a gridded reference evapotranspiration dataset for the Great Lakes region. *J. Hydrol. Reg. Stud.*, **24**, 100606, doi:10.1016/j.ejrh.2019.100606.

Liu, D. L., F. Ji, B. Wang, C. Waters, P. Feng, and R. Darbyshire, 2020: The implication of spatial interpolated climate data on biophysical modelling in agricultural systems. *Int. J. Climatol.*, **40**, 2870–2890, doi:10.1002/joc.6371. https://onlinelibrary.wiley.com/doi/abs/10.1002/joc.6371 (Accessed April 17, 2020).

Maestrini, B., and B. Basso, 2018: Drivers of within-field spatial and temporal variability of crop yield across the US Midwest. *Sci. Rep.*, **8**, 14833, doi:10.1038/s41598-018-32779-3. http://www.nature.com/articles/s41598-018-32779-3 (Accessed September 25, 2019).

Maraun, D., and Coauthors, 2017: Towards process-informed bias correction of climate change simulations. *Nature Climate Change*, Vol. 7 of.

Mourtzinis, S., J. I. Rattalino Edreira, S. P. Conley, and P. Grassini, 2017: From grid to field: Assessing quality of gridded weather data for agricultural applications. *Eur. J. Agron.*, **82**, 163–172, doi:10.1016/J.EJA.2016.10.013. https://www.sciencedirect.com/science/article/pii/S1161030116302076 (Accessed April 19, 2018).

Ngo-Duc, T., J. Polcher, and K. Laval, 2005: A 53-year forcing data set for land surface models. *J. Geophys. Res. Atmos.*, **110**, n/a-n/a, doi:10.1029/2004JD005434. http://doi.wiley.com/10.1029/2004JD005434 (Accessed February 22, 2021).

Perdinan, J.A. Winkler, and J.A. Andresen, 2020: Evaluation of multiple approaches to estimate daily solar radiation for input to crop process models. *Atmosphere*, **12,** 8, doi: 10.3390/atmos12010008

Ramírez, J. A., M. T. Hobbins, and T. C. Brown, 2005: Observational evidence of the complementary relationship in regional evaporation lends strong support for Bouchet's hypothesis. *Geophys. Res. Lett.*, **32**, 15401, doi:10.1029/2005GL023549. https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2005GL023549 (Accessed July 15, 2021).

Slater, A. G., 2016: Surface solar radiation in North America: A comparison of observations, reanalyses, satellite, and derived products. *J. Hydrometeorol.*, **17**, 401–420, doi:10.1175/JHM-D-15-0087.1. http://dx.doi.org/10.1175/JHM-D-15- (Accessed February 22, 2021).

Tao, F., and Coauthors, 2020: Why do crop models diverge substantially in climate impact projections? A comprehensive analysis based on eight barley crop models. *Agric. For. Meteorol.*, **281**, 107851, doi:10.1016/j.agrformet.2019.107851.

Wang, Y., G. Sivandran, and J. M. Bielicki, 2018: The stationarity of two statistical downscaling methods for precipitation under different choices of cross-validation periods. *Int. J. Climatol.*, **38**, e330–e348, doi:10.1002/joc.5375. https://rmets.onlinelibrary.wiley.com/doi/full/10.1002/joc.5375 (Accessed June 22, 2021).

Weedon, G. P., and Coauthors, 2011: Creation of the WATCH Forcing Data and Its Use to Assess Global and Regional Reference Crop Evaporation over Land during the Twentieth Century. *J. Hydrometeorol.*, **12**, 823–848, doi:10.1175/2011JHM1369.1. http://journals.ametsoc.org/doi/abs/10.1175/2011JHM1369.1 (Accessed January 31, 2018).

Wei, Y., and Coauthors, 2014: The north american carbon program multi-scale synthesis and terrestrial model intercomparison project - Part 2: Environmental driver data. *Geosci. Model Dev.*, **7**, 2875–2893, doi:10.5194/gmd-7-2875-2014.

Xiong, W., and Coauthors, 2020: Different uncertainty distribution between high and low latitudes in modelling warming impacts on wheat. *Nat. Food*, **1**, 63–69, doi:10.1038/s43016-019-0004-2. https://www.nature.com/articles/s43016-019-0004-2 (Accessed June 22, 2021).

Zhang, W., and G. Villarini, 2019: On the weather types that shape the precipitation patterns across the U.S. Midwest. *Clim. Dyn.*, 1–16, doi:10.1007/s00382-019-04783-4. http://link.springer.com/10.1007/s00382-019-04783-4 (Accessed May 6, 2019).

CHAPTER 3. TRENDS IN QUALITY CONTROLLED PRECIPITATION INDICATORS IN THE UNITED STATES MIDWEST AND GREAT LAKES REGION

Abstract

Changes in precipitation can have broad and significant societal impacts. A number of previous studies that analyzed changes in precipitation across the Great Lakes and Midwest for a variety of time periods and using a range of quality-control standards and methods observed increased precipitation rates and totals, although there was considerable site-to-site variability, even for sites in close physical proximity. Biases and discontinuities in precipitation observations may contribute to this variability. This study identifies and examines changes in precipitation utilizing a unique approach to observation series screening over a region encompassing the Great Lakes and broader Midwestern region of the United States for the period 1951-2019. A multiple tier procedure was utilized to identify high quality input data series from the Global Historical Climatology Network-Daily dataset. Annual and seasonal time series of precipitation indicators were calculated and subjected to breakpoint analysis as further quality control. Trends were analyzed across a broad range of related indicators, from totals and frequencies of threshold events to event duration and potential linkages with total precipitable water. Results indicate that annual precipitation has generally increased across the region in terms of totals, although there is substantial variation across the study domain in the significance and magnitude of annual trends by indicator. Annual trends were spatially most consistent across eastern areas of the study domain while relatively greater station-to-station variability in trend significance and magnitude was observed across northern and western portions. Significant trends were generally fewer in number for seasonal precipitation indicators and less spatially coherent. The greatest number of significant trends occurred in fall with the fewest in spring. Correlation of indicator trends with trends of mean total precipitable water suggests weak correlations annually and moderate correlations at the seasonal scale. The trends of the precipitation indicators in our study exhibited

more coherent spatial patterns when compared with studies with different quality control criteria, illustrating the importance of quality control of observations in climatic studies and highlighting the complexity of the changing character of precipitation.

Introduction

Precipitation is the longest observed and most widely reported meteorological variable and is an essential component of the Earth's hydrologic cycle (Legates and Willmott, 1990). Precipitation is commonly defined as "the amount, usually expressed in millimeters or inches of liquid water depth, of the water substance that has fallen at a given point over a specified period of time" (Huschke, 1959, p.438). Although precipitation accumulation at daily, monthly, seasonal and annual scales has received the most attention in the climatological literature (e.g., Contractor et al. 2021), other precipitation characteristics such as frequency, intensity, and duration are as much, if not more, of a concern for many natural and human systems (Trenberth et al., 2003; Bartels et al., 2020). Moreover, changes in one or more precipitation characteristics can have substantial societal implications impacting many sectors, including, among others, agriculture (e.g., Pielke et al., 2000; Rosenzweig et al., 2002; Hunt et al. 2020; Kiefer et al., 2021), transportation (e.g., Attavanich et al., 2013; Talukder & Hipel, 2020), and tourism (e.g., Chin et al., 2018).

Changes in precipitation characteristics are a particular concern for the Midwest and Great Lakes region of the United States given the region's unique hydrology (Gronewold et al., 2021) and its agricultural importance and contribution to regional, national and global food security (Angel et al., 2018; Takle & Gutowski, 2020). Not surprisingly, a number of studies have investigated temporal trends in precipitation characteristics either specifically for the region (e.g., Zhang & Villarini, 2019) or as part of larger analyses of precipitation trends in the United

States (e.g., Kunkel et al., 2020a). For the most part, these analyses have focused on trends in annual and seasonal precipitation totals (e.g., Schoof et al., 2010), extreme precipitation (e.g., Pryor et al., 2009; Walsh et al., 2014), and/or the frequency of wet days (e.g., Roque-Malo & Kumar 2017; Bartels et al. 2018). In general, precipitation frequency and total accumulation appear to have increased across the region over the last several decades (Higgins et al., 2007; Dai et al., 2016; Contractor et al, 2021). In addition, the amount of precipitation falling during the heaviest events has increased at a greater rate in the Midwest and Great Lakes region compared to the national average (Angel et al. 2018). Extended dry periods have become less frequent, but their intensity (i.e. length) has increased slightly in recent decades (Groisman & Knight, 2008).

One constraint to comprehensive and accurate analysis of temporal trends in precipitation characteristics at the regional scale is the availability and quality of precipitation observations (Costa and Soares, 2009). Although numerous authors have examined the homogeneity of time series for various climatic variables including daily precipitation (Winkler, 2004; Daly et al., 2007; Wang et al., 2010), many studies employing in-situ climate observations fail to take data quality, other than data completeness, into account, even though the magnitude and sign of temporal trends can be biased by changes in technology, station siting, observing practices and other inhomogeneities that are not necessarily captured by station record completeness or recorded in standard metadata archives (Wang et al., 2010; Williams et al., 2012; Baule and Shulski, 2014). Recent progress in the development of spatial and temporal interpolation schemes and gridded datasets, the integration of radar and satellite derived precipitation estimates with in-situ observations, the development of atmospheric reanalysis products, and the availability of simulations from regional and global climate models have only partially alleviated concerns about data quality (Zhang et al., 2011). The limited periods of record for radar and

satellite precipitation estimates constrain their use for estimating temporal trends, and gridded

datasets can inherit the inhomogeneities of the underlying station observations, with developers

of these datasets often advising against their use for time series analysis (e.g., Daly et al. 2010).

Consequently, station-based climatologies, in spite of their limitations, remain the benchmark for

the assessment of long-term trends (Kiefer et al., 2021), although caution in their application is

critical to guard against misinterpreting temporal trends. Earlier studies of precipitation trends

for the Midwest and Great Lakes region frequently used station-level daily precipitation

observations from the Global Historical Climatology Network-Daily (GHCN-D) database

(Menne et al. 2012) for trend estimation (e.g., Villarini et al. 2011; Janssen et al. 2014; Guilbert

et al. 2015; Wu 2015; Hoerling et al. 2016; Roque-Malo and Kumar 2017; Huang et al. 2017,

2018; Kunkel et al. 2020a, b). For the most part, quality-control procedures have been limited to

those applied by the GHCN-D dataset developers to identify and/or correct for errors and

inhomogeneities in the precipitation data (Durre et al. 2008, 2010), supplemented by an

evaluation of data completeness (e.g., Kunkel et al. 2002b).

Other studies have investigated the synoptic-scale drivers of precipitation, particularly

those associated with extreme precipitation, finding that extreme precipitation events across the

Midwest and Great Lakes region are often associated with a westward expansion and

strengthening of subtropical high pressure across the western Atlantic Basin (Gutowski et al.

2008) as well as the advection of low-level moisture from the Gulf of Mexico ahead of slow

moving tropospheric waves (Winkler 1988; Zhang and Villarini 2019), with the latter being more

prevalent in the western portions of the region and the former in the eastern areas (Bell and

Janowiak 1995; Konrad 2001; Weaver and Nigam 2008). Consistent with these findings, Kunkel

et al. (2020a) showed that extreme daily precipitation events across the contiguous United States,

including the Midwest and Great Lakes region, are directly related to total precipitable water. Specifically, Kunkel et al. (2020b) examined the relationship between regional trends in total precipitable water and regional trends in extreme precipitation as calculated from GHCN-D station-level time series. These large-scale drivers of precipitation are often amplified or suppressed by regional and local climate drivers such as topography, water bodies, and land use/land cover (Myhre et al., 2016; Kunkel et al., 2020a), which can introduce considerable spatial variability in the temporal trends of precipitation characteristics, especially in the Midwest and Great Lakes region with its large water bodies and varied land use/land cover. For the most part, quality control of the precipitation observations employed in these studies of the synoptic, regional and local drivers of precipitation and their contribution to temporal trends in precipitation was confined to an assessment of data completeness of the precipitation time series.

This study provides a comprehensive assessment of the temporal trends in precipitation characteristics for the Midwest and Great Lakes region that focuses on the quality of available precipitation time series. We employ a three-step quality-control procedure that evaluates the GHCN-D precipitation time series for data completeness, possible observer bias, and potential breakpoints (i.e., discontinuities) with the goal of identifying those GHCN-D stations in which we have the greatest confidence for precipitation trend analysis, thereby increasing confidence in the sign, magnitude, significance, and spatial coherence of precipitation trends. We include a range of precipitation indicators that capture the frequency and persistence of high frequency, low magnitude and low frequency, high magnitude events. Furthermore, we examine the associations between temporal trends in the quality-controlled suite of precipitation indicators and trends in atmospheric moisture. The study findings provide the region's many stakeholders with needed information on long-term trends in precipitation characteristics of concern to them,

greater certainty in incorporating these data in planning processes and a high-quality baseline for

assessing future trends.

Methods

For this study, the Midwest and Great Lakes region was defined as the states of

Pennsylvania, Ohio, Indiana, Michigan, Illinois, Wisconsin, Minnesota, Iowa, Missouri, Kansas,

Nebraska, South Dakota, New York, and North Dakota (Figure 3.1).
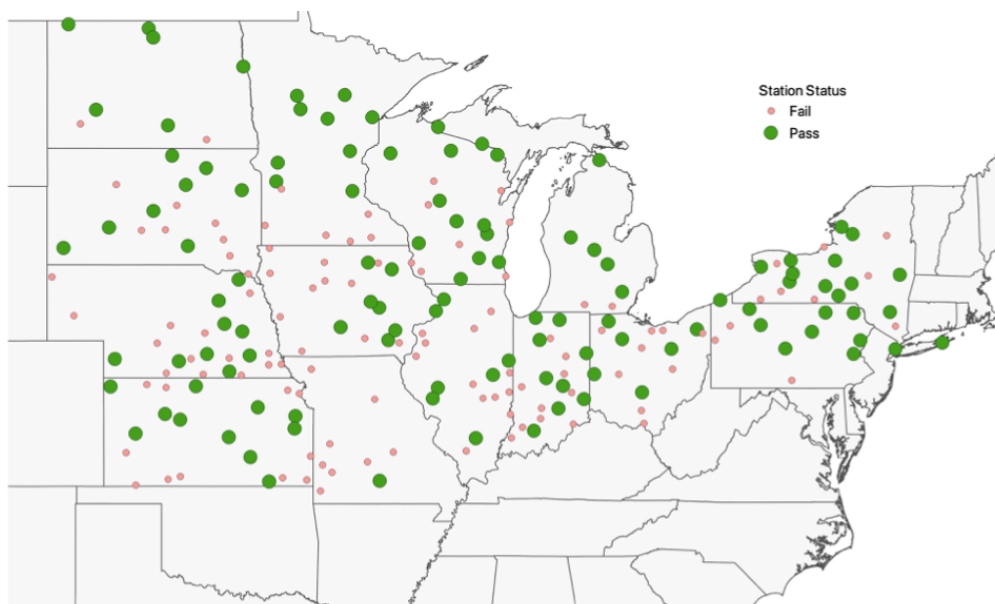
Figure 3.1 Study region and the United States Historical Climate Network (USHCN) stations
(green circles) within the study region that passed the quality control checks for data
completeness and lack of observer bias as outlined in the methods. Stations that passed the data
completeness check but failed at least one of the tests for lack of observer bias are shown as pink
circles. The number of stations that passed the third quality control test (no breakpoints) is given
in Table 3.2.

We analyzed a subset of individual site climate series from the National Centers for

Environmental Information's (NCEI) GHCN-D collection (Menne et al. 2012). As a first step in

selecting stations for the analysis, we examined the GHCN-D database for station series included

in the United State Historical Climatology Network (USHCN; Easterling 2002) which had at

least 90% data completeness for daily precipitation during 1951-2019. Only USHCN sites were considered, as these stations were preselected by NCEI based on record length, data completeness, and historical stability (Menne & Williams, 2012). Data flagged by the GHCN-D quality control procedures as suspicious were marked as missing (Menne et al. 2012). The length of the study period allowed for trends in the second half of the 20th century and the early 21st century to be assessed while maintaining a relatively large pool of potential stations and reasonable spatial coverage. This first data quality control step led to an initial subset of 317 stations over the study region.

The next quality control step involved using tests proposed by Daly et al. (2007) to check for observer bias in precipitation time series, specifically the underreporting of light (1.26 mm) precipitation amounts and the overreporting of precipitation amounts evenly divisible by 5 and/or 10 when expressed as inches. As these tests were designed for data originally measured in inches, they are described here using inches (in.) in place of millimeters. The under-reporting check consisted of calculating the ratio of counts between 0.06 -0.10 in. (C_{6-10}) and 0.01-0.05 in. (C_{1-5}) as follows:

$$R_L = \frac{C_{6-10}}{C_{1-5}},$$

where C_{6-10} is the total observation count in the 0.06-0.10-in. range, and C_{1-5} is the total observation count in the 0.01-0.05-in. range. If the ratio, R_L, between C_{6-10} and C_{1-5} exceeded 0.60, the station failed the check (Daly et al., 2007).

The tests for errant reporting of values divisible by 5 or 10 were conducted by binning precipitation into 0.01 in. bins, fitting a gamma distribution to the data between 0.03 in. and 1.00 in., and comparing the predicted (P) and observed (O) frequency of the binned observations with the residual (R) calculated as:

$$R = 100 * (P - O)$$

The test for biases in amounts divisible by 5 and amounts only divisible by 10 were carried out separately. For the divisible by 5 test, the first residual mean was calculated by averaging the residuals over all amounts except those divisible by 5 (R_1) and the second residual (R_5) consisted of the mean of residuals for only amounts divisible by 5 as follows:

$$\bar{R}_1 = \frac{\sum_{i=1}^{n_1} R_{1_i}}{n_1}; \ \bar{R}_5 = \frac{\sum_{i=1}^{n_5} R_5}{n_5}$$

where n_1 and n_5 are the number of ones and fives bins and R_1 and R_5 are residuals calculated from equation 2. The means for the divisible by 10 bias were calculated similarly, instead using values only divisible by 10. The means were compared using a two-tailed t-test with an alpha level of 0.01.

Examples of output from the second quality control procedures are shown in Figure 3.2 for two locations, Manhattan, KS HCN which passed all the bias tests at $p \leq 0.01$ or $R_L \leq 0.6$ despite showing a small divisible by 10 bias, and Lamar 7N, MO HCN which failed the bias tests showing a strong under reporting bias, a strong divisible by 5 bias, and a strong divisible by 10 bias. Stations that failed any of the tests were removed from the analysis, leaving a subset of 114 long-term climate series across the Midwest and Great Lakes region for the period from 1951-2019 for precipitation indicators.
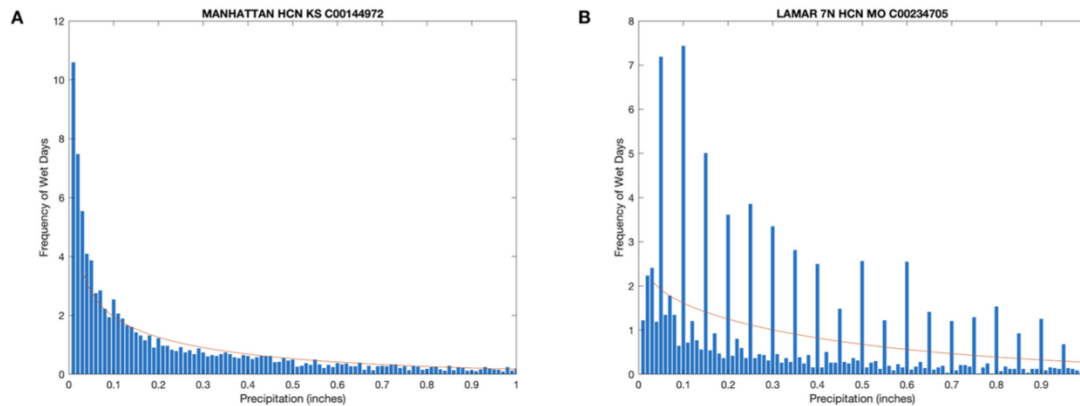
Figure 3.2 Histograms from two example stations in the study region showing precipitation frequency (blue bars) over the period from 1951-2019 binned in 0.01" increments and a gamma distribution (red line) fit to the data following Daly et al. (2007). Manhattan, KS HCN (left) passes (p ≤ .01) the tests for underreporting of daily precipitation amounts less than 0.05 in. (1.26 mm) and for overreporting of daily precipitation amounts (in inches) evenly divisible by 5 or 10 despite showing a small divisible by 10 bias. Lamar 7N, MO HCN (right) fails all three tests (p ≤0 .01), showing a strong under reporting bias, a strong divisible by 5 bias, and a strong divisible by 10 bias.

The third quality control step involved checking the time series of the precipitation indicators for breakpoints. Possible sources of discontinuities in the time series include, among others, instrument changes, station moves, and changes in observation protocols including time of observation (Winkler 2004). Following Mallakpour and Villarini (2016), the Pettitt test (Pettitt 1979) was applied to identify years when a breakpoint is likely, indicating a non-homogenous time series. A breakpoint was considered significant at p ≤ 0.01, and the time series for that station was excluded from further analysis. This resulted in a variable number of stations per indicator, with the number of excluded stations ranging from none to a maximum of 17. The description of the Pettitt test, following Jaiswal et al. (2015) is: a method that detects a significant change in the mean of a time series, where the exact time of the change (i.e. breakpoint) is unknown, According to the Pettit test, if $x_1, x_2, x_3, ...x_n$ is a series of observed data which has a break point at t so that $x_1, x_2, x_3, ...x_t$ has a distribution ($F_1(x)$) which is different from

the distribution ($F_2(x)$) of the second part of the series $x_{t+1}, x_{t+2}, x_{t+3}, ... x_n$. The non-parametric test statistic is described as follows:

$$U_t = \sum_{i=1}^{t} \sum_{j=t+1}^{n} sgn(x_i - x_j)$$

$$sgn(x_i - x_j) = \begin{cases} 1, if(x_i - x_j) > 0 \\ 0, if(x_i - x_j) = 0 \\ -1, if(x_i - x_j) < 0 \end{cases}$$

The test statistic K and the associated confidence level (ρ) for the sample length (n) is described as:

$$K = Max|U_t|$$

$$\rho = e^{(\frac{-K}{n^2+n^3})}$$

When ρ is smaller than the specified confidence level (p), a breakpoint is considered significant.

Precipitation Indicators

This study included a range of precipitation indicators. Several indicators were used to characterize the frequency of non-extreme precipitation, including the number of days with measurable precipitation (e.g. Pryor et al. 2009) and the probabilities of wet-wet day and dry-dry day sequences (e.g., Ines et al., 2011). A wet day was defined as a precipitation total $\geq$ 1.26mm (0.05 in.) (Groisman et al. 1999). Extreme precipitation was represented in the analysis by indices developed by the Expert Team on Climate Change Detection and Indices (ETCCDI) (Donat et al. 2013) and annual values were calculated using the software packages provided by the ETCCDI Working Group (available at http://www.climdex.org). The extreme precipitation indicators include 10 wet indices and 1 dry index that can be further grouped into percentile-based indices (2), threshold indices (3), absolute value indices (2), duration indices (2), annual accumulation, and "simple" intensity (annual total precipitation divided by the number of wet

days). For the percentile-based indices, the base period for defining the percentile value was the 30-year climate normal period of 1981-2010. Descriptions of each of the non-extreme and extreme precipitation indicators are provided in Table 3.1.

All precipitation indicators were also defined for the climatological seasons of spring (MAM), summer (JJA), fall (SON), and winter (DJF). This is in contrast to most previous studies where precipitation indicators were calculated for annual time steps, with less attention paid to the seasonal variations in the precipitation indicators beyond the frequency of high intensity daily precipitation events (e.g. Mallakpour and Villarini, 2017). Given the importance of precipitation timing and sequencing for numerous regional applications, such as soil moisture and nitrogen movement in agricultural systems (Bowles et al., 2018; Riha et al., 1996) and plant disease risk (Komoto et al. 2021), this study extended the precipitation indicators to the seasonal time step. Additionally, the seasonal analyses provide greater context to more clearly interpret annual trends.

Table 3.1 Precipitation indicators included in the analysis. All indicators were calculated annually and seasonally except those marked with an asterisk (*) were only calculated annually.

Index Name	ID	Definition	Units
Accumulation/Simple Intensity			
Annual total wet day precipitation	PRCPTOT	Total precipitation on wet days (PRCP ≥ 1.26mm)	mm
Simple daily intensity index	SDII	Total precipitation divided by the number of wet days	mm day^{-1}
Duration			
Consecutive wet days	CWD	Maximum number of consecutive days with PRCP ≥1.26 mm	days
Consecutive dry days	CDD	Maximum number of consecutive days with PRCP < 1.26 mm	days
Percentile (Percentile values were calculated for the period 1951-1980)			
Precipitation on very wet days	R95pTOT	Total precipitation on days when PRCP ≥ 95th percentile	mm
Precipitation on extremely wet days	R99pTOT	Total precipitation on days when PRCP ≥ 99th percentile	mm
Threshold			
Number of days with measurable precipitation	R1.26mm	Number of days with PRCP ≥ 1.26mm	days
Number of heavy precipitation days	R10mm	Number of days with precipitation ≥ 10mm	days
Number of very heavy precipitation days	R20mm	Number of days with precipitation ≥ 20mm	days
Number of days with consecutive days with measurable precipitation	WW	Annual count of days when PRCP ≥ 1.26 mm on consecutive days	days
Number of days with consecutive days without measurable precipitation	DD	Annual count of days when PRCP < 1.26 mm on consecutive days	days
Absolute			
Maximum 1-day precipitation*	Rx1day	Maximum 1-day precipitation	mm
Maximum 5-day precipitation*	Rx5day	Maximum consecutive 5-day precipitation	mm

Following the lead of numerous recent studies (e.g., Alexander et al., 2006; Pryor et al., 2009; Shulski et al., 2015; Dai et al., 2016; Roque-Malo and Kumar, 2017), non-parametric

statistical methods were employed to estimate the significance and magnitude of temporal trends

in the precipitation indicators at the study locations. While a number of potential non-parametric

methods were available (e.g. Sneyers, 1990; Şen, 2015; Onyutha,2021), we chose the two-tailed

Mann-Kendall trend test (Mann, 1945; Kendall, 1955;1975) to test for the significance of

potential temporal trends due to its prevalence in previously mentioned studies across the region

to allow for intercomparison of results. A strength of the Mann-Kendall method is its ability to

assess the significance of trends that are monotonic but not necessarily linear in character. For

those locations with significant trends as identified by the Mann-Kendall test, the nonparametric

Kendall's tau-based slope estimator (Sen, 1968) was used to obtain a numerical estimate of the

temporal trend. All analyses were conducted using three significance levels ($p \leq 0.05$, $p \leq 0.10$,

and $p \leq 0.20$) to examine how significance level affects the number of significant trends and their

spatial representation. The equations describing the Mann-Kendall test are as follows:

$$S = \sum_{i=1}^{n} \sum_{j=1}^{i-1} sgn(x_i - x_j)$$

where n is the total length of data, x_i and x_j are two generic sequential data values, and the

following function assumes the values:

$$sgn(x_i - x_j) = \begin{cases} 1, if \ (x_i - x_j) > 0 \\ 0, if \ (x_i - x_j) = 0 \\ -1, if \ (x_i - x_j) < 0 \end{cases}$$

The mean of S is E[S] = 0 and the variance σ^2 is

$$\sigma^2 = \frac{1}{n}\left[n(n-1)(2n+5) - \sum_t t(t-1)(2t+5))\right]$$

where *n* is the length of the time series and t is the extent of any given ties and Σt denotes the

summation over all tied values. The statistic S is approximately normally distributed provided

that the following Z-transformation is employed:

$$Z = \begin{cases} \dfrac{S-1}{\sigma} & if \ S < 0 \\ 0 & if \ S = 0 \\ \dfrac{S+1}{\sigma} & if \ S > 0 \end{cases}$$

The Sen's (1968) slope was calculated as follows: first, a set of linear slopes is calculated

$$d_k = \frac{X_j - X_i}{j - i}$$

for $(1 \leq i < j \leq n)$, where *d* is the slope, *X* denotes the variable, *n* is the sample length, and *i*, *j*,

and *k* are indices. Sen's slope is then calculated as the median from all slopes (d_k).

Total Precipitable Water

Daily values of total precipitable water (TOTPRCPWAT) were obtained from the

NCEP/NCAR I Reanalysis (Kalnay et al., 1996) at a 2.5°x2.5° spatial resolution. The daily

values were used to calculate mean daily annual and seasonal total precipitable water for each

year during the 1951-2019 study period for a bounding box ranging from 106°W - 69°W

longitude and 34°N - 54°N latitude. Only grid cells that contained observing sites used in our

analyses were subjected to analysis. Pearson correlation coefficients (r) and non-parametric

Kendall rank correlation coefficients (τ) were calculated between the trend value of four

representative precipitation indicators (WW, PRCPTOT, R1.26mm, R95pTOT) at each station

considered previously and the trend of total precipitable water of the nearest reanalysis grid cell.

Pettitt tests were conducted on the NCEP NCAR time series of precipitable water for each grid

cell to examine for potential heterogeneities prior to the satellite era (e.g. Kunkel et al. 2020b.).

Significant breakpoints (p < 0.01) were evident at some grid cells, however they were not clustered in time. Given the noted strength of the NCEP-NCAR reanalysis in areas where radiosonde observations are available (Trenberth et al. 2005), as in our study region, we deemed the data appropriate for our analyses.

Results

Trends in Precipitation Indicators

<u>Annual Indicators</u>

For the annually-derived precipitation indictors, the number of stations with statistically-significant ($p \leq 0.10$) trends varied substantially among the different indicators, ranging from 67% of the station sites for annual total precipitation (PRCPTOT) to only 20% of the stations for the maximum number of consecutive wet days per year (CWD) (Table 3.2). With the exception of the maximum number of consecutive dry days (CDD) and the number of dry-dry day sequences (DD), more than 90% of the statistically significant trends over time when summed across the indicator variables are positive, indicating a generally wetter climate. The negative trends observed for CDD and DD are also indicative of a wetter climate. In addition to PRCPTOT, the majority of the locations display significant upward trends for the simple intensity index (PRCPTOT divided by the number of days with precipitation ≥ 1 mm; SDII), the number of days per year with precipitation ≥ 10mm (R10mm), the number of days per year with precipitation ≥ 20 mm (R20mm), and the total precipitation on days with daily precipitation $\geq$ 95th percentile (R95pTOT). A majority (54%) of stations also have statistically significant negative trends for DD, whereas only 39% of the locations have statistically significant positive trends in the annual number of wet-wet day sequences (WW). A considerably smaller number of stations displayed significant trends for several of the other indicators, with the fraction of

stations with significant trends falling below 30% for CDD, CWD, total precipitation on days with daily precipitation $\geq 99^{th}$ percentile (R99pTOT), maximum one-day precipitation amount (Rx1day) and maximum consecutive 5-day precipitation (Rx5day). With the exception of CDD and DD, statistically significant negative trends were infrequent for the various indicators, ranging from no significant negative trends for CWD, SDII, R95pTOT, R99pTOT, SDII, Rx1day, and Rx5day to 6% of the locations for WW. The number of stations exhibiting significant breakpoints was greatest for PRCPTOT (14) and R.126MM (17). Rx1day, Rx5day, CDD, and CWD showed no significant breakpoints.

Table 3.2 Number of stations exhibiting statistically significant trends (Mann Kendall, p ≤ 0.05 two-tailed, p ≤ 0.10 two-tailed, p ≤ 0.20, two-tailed) from 1951-2019 in the annual precipitation indicators. Indicators where more than 50% of stations analyzed showed a significant trend are shown in bold. See Table 3.1 for definition of the abbreviations for the precipitation indicators.

Precipitation Indicator	Total number of stations after breakpoint analysis	Number of stations with significant positive trends (p≤0.05)	Number of stations with significant positive trends (p≤0.10)	Number of stations with significant positive trends (p≤0.20)	Number of stations with significant negative trends (p≤0.05)	Number of stations with significant negative trends (p≤0.10)	Number of stations with significant negative trends (p≤0.20)
PRCPTOT	100	47	**75**	**79**	1	1	1
R1.26mm	97	42	**53**	**61**	1	3	5
SDII	112	42	52	**62**	0	0	3
CWD	114	17	23	36	0	0	0
CDD	113	0	2	2	16	28	44
WW	105	27	44	51	5	7	9
DD	107	2	3	3	41	**61**	**65**
R10mm	104	42	**60**	**72**	0	1	1
R20mm	101	35	**58**	**59**	1	1	1
R95pTOT	104	38	**62**	**63**	0	0	0
R99pTOT	108	11	33	39	0	0	0
Rx1day	114	13	29	43	0	0	0
Rx5day	114	23	33	47	0	0	2

A subset of indicators that encompass the range of precipitation characteristics included in the analysis, namely PRCPTOT, WW, the number of wet days with precipitation ≥ 1.26 mm (R1.26mm), and R95pTOT, is used to illustrate the spatial variability across the study region in the temporal trends for the annual indicators (Figure 3.3). For all four indicators, statistically significant positive trends are distributed across the study region, although the magnitude of these trends is generally larger in the eastern two-thirds of the study area, including in the vicinity of the Great Lakes. In the western third of the study region, although the trends in the selected indicators are generally positive, the magnitude of the trends is smaller with relatively

fewer stations meeting the threshold for statistical significance. Regardless of precipitation indicator, negative trends are evident for only a few stations and are insignificant. Significance of the same indicators but with a weaker significance threshold ($p \leq 0.20$) is shown in Figure 3.4. When the significance level is lowered, the number of significant positive trends increases substantially with no or very little increase in the number of significant negative trends and the stations with significant positive trends are more spatially coherent. When a stricter ($p \leq 0.05$) threshold is used, the number of stations exhibiting significant trends decreases when compared to the moderate ($p \leq 0.10$) and weak ($p \leq 0.20$) thresholds. Spatially, when the strict threshold is used, the largest groupings of sites with significant trends are in the central and eastern portions of the study region (Figure 3.5). The number and spatial coherence of significant positive trends in the western areas of the study region are reduced under the strict criterion.
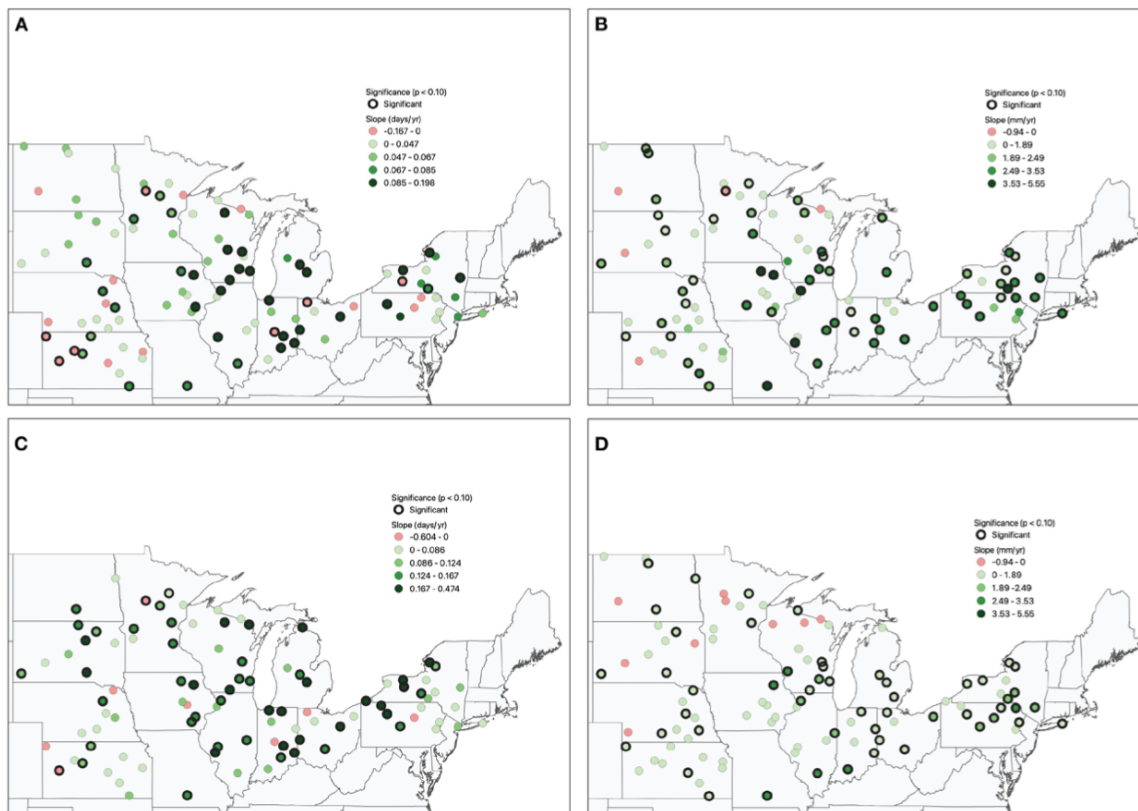
Figure 3.3 Trends for 1951-2019 in representative annual indicators of precipitation characteristics at locations in the Midwest and Great Lakes region that passed the quality control checks described in the methods section: a) annual counts of wet-wet day sequences (ANN WW; days; count year^{-1}; upper left), b) annual total precipitation on wet days (PRCPTOT; mm year^{-1}; upper right), c) number of days with precipitation $\geq$ 1.26 mm (R1.26mm; days year^{-1}; lower left), and d) total precipitation on days when precipitation is $\geq$ 95th percentile (R95pTOT; mm year^{-1}; lower right).
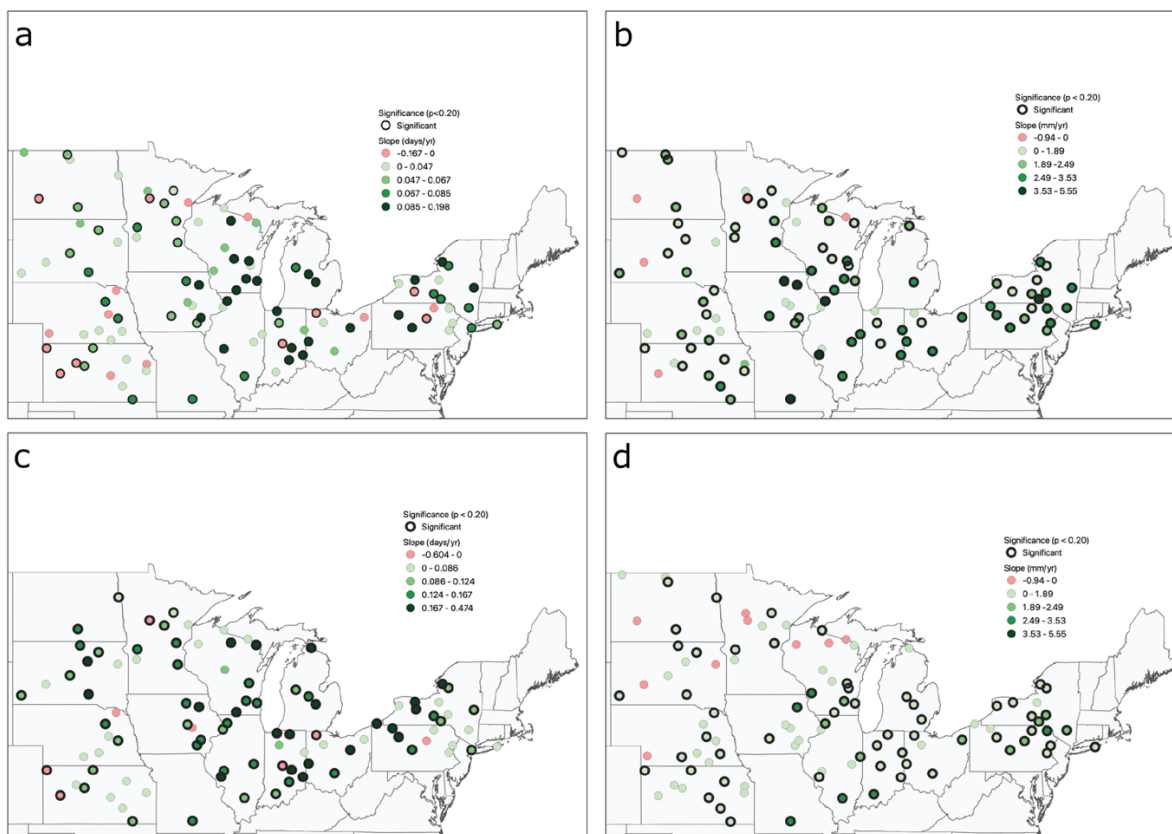
Figure 3.4 Trends (units year^{-1}) for 1951-2019 in representative annual indicators of precipitation characteristics at locations in the Midwest and Great Lakes region that passed the quality control checks described in the methods section: a) annual counts of wet-wet day sequences (ANN WW; days; upper left), b) annual total precipitation on wet days (PRCPTOT; mm; upper right), c) number of days with precipitation ≥ 1.26 mm (R1.26mm; days; lower left), and d) total precipitation on days when precipitation is ≥ 95th percentile (R95pTOT; mm; lower right). Significance threshold was reduced to p≤ 0.20.
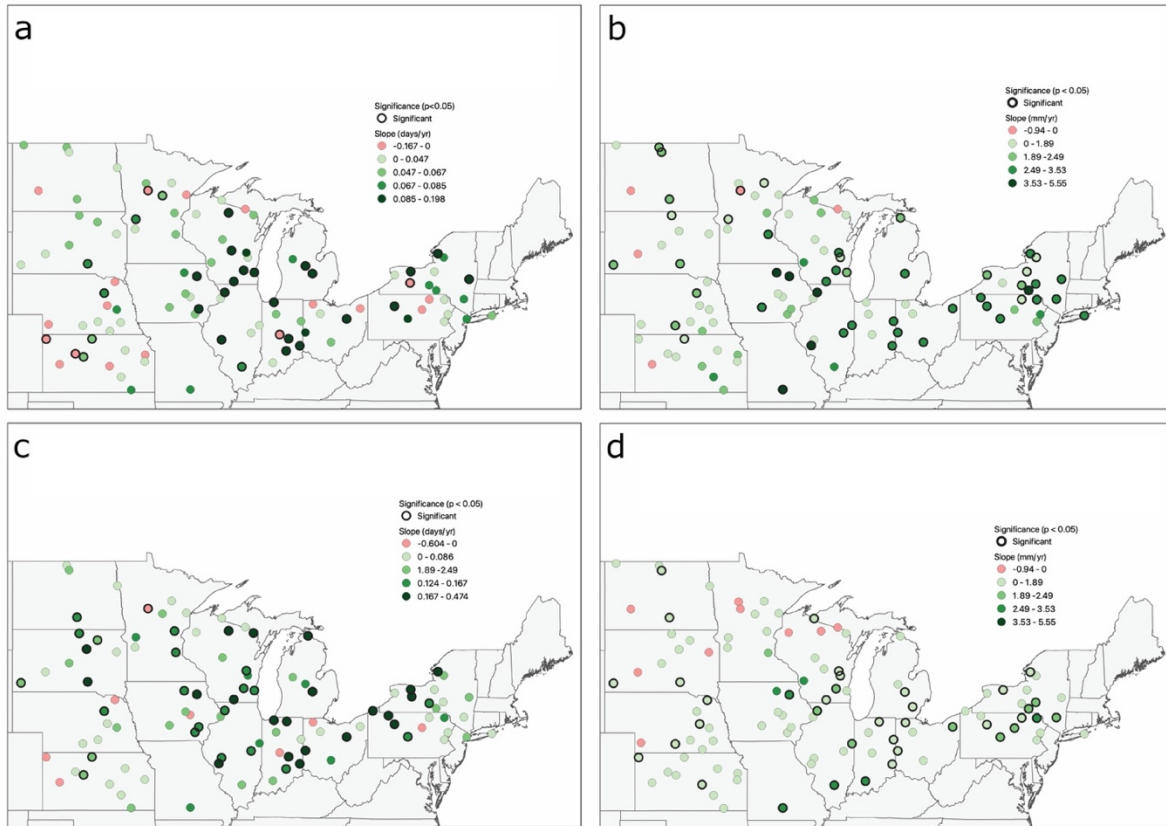
Figure 3.5 Trends (units year^{-1}) for 1951-2019 in representative annual indicators of precipitation characteristics at locations in the Midwest and Great Lakes region that passed the quality control checks described in the methods section: a) annual counts of wet-wet day sequences (ANN WW; days; upper left), b) annual total precipitation on wet days (PRCPTOT; mm; upper right), c) number of days with precipitation ≥ 1.26 mm (R1.26mm; days; lower left), and d) total precipitation on days when precipitation is ≥ 95th percentile (R95pTOT; mm; lower right). Significance threshold was increased to p≤ 0.05.

We also evaluated the ratio of the trend estimates for the annual indicators of R95pTOT and PRCPTOT for stations with a significant positive trend in PRCPTOT, as an indicator of the relative contribution of precipitation on very wet days to trends in total precipitation (Figure 3.6). In general, precipitation on very wet days has contributed the most (ratios >0.60 and at some locations >1.0) to annual total precipitation in eastern New York/Pennsylvania, Indiana, southern Wisconsin/eastern Iowa, and eastern Nebraska/Kansas, compared to elsewhere in the study region. The modest (<0.60) ratios at many locations elsewhere suggest that the overall increase in total precipitation is not exclusively, or even primarily, tied to increases in the frequency of

higher intensity events. Rather, changes in the frequency of lighter accumulations are also

contributing to the trends in total precipitation.
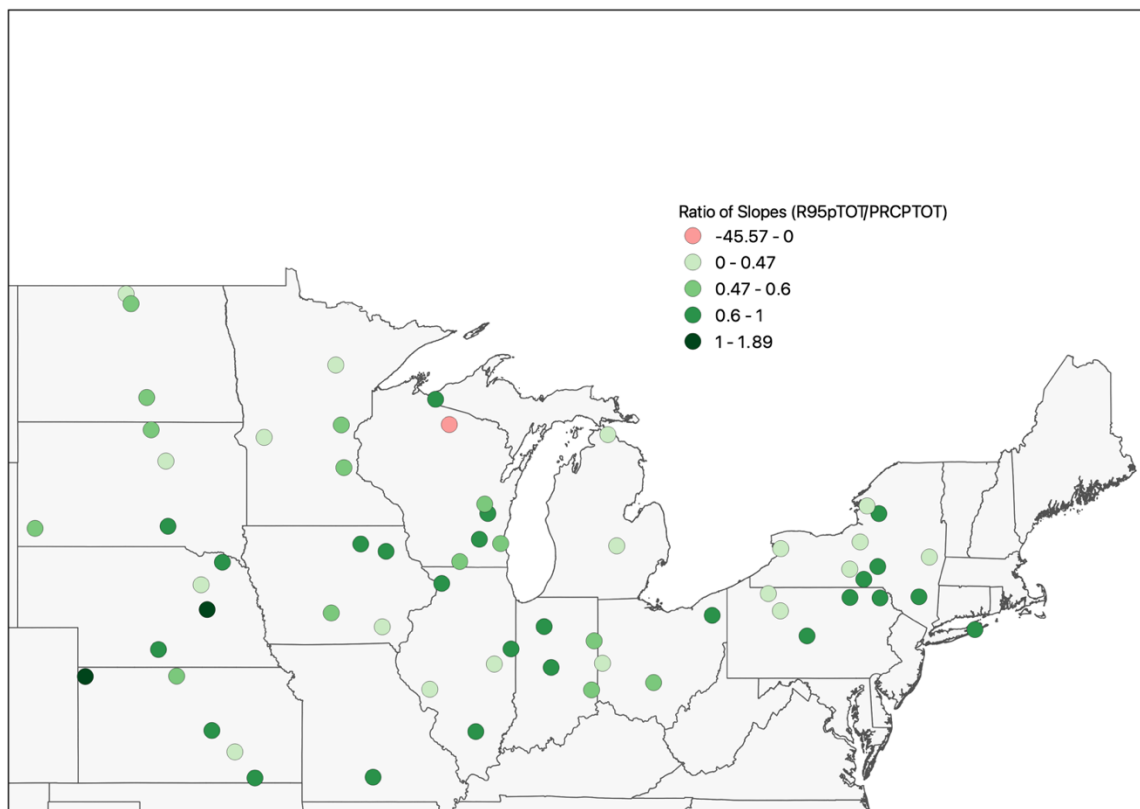
Figure 3.6 Ratio of the trend in annual total precipitation on days when precipitation is ≥ 95th percentile (R95pTOT) to the trend in annual total precipitation on all wet days (PRCPTOT) for those stations with statistically significant positive trends in PRCPTOT.

To better understand the consistency at individual locations of the trends across

precipitation indicators, a four-sided Venn diagram was used to plot the number of significant

positive trends and the percentage of significant positive trends for all possible combinations of

the four representative precipitation indicators, PRCPTOT, WW, R1.26mm, and R95pTOT,

recognizing that the number of available stations varies among indicators due to differing

frequency of breakpoints identified in the time series (Figure 3.7). The number of locations with

significant positive trends for multiple indicators is substantial, with 16.3% of the stations with

positive trends for all the indicators considered and 19.8% stations with positive trends for the two accumulation indicators PRCPTOT and R95pTOT. The number of stations with significant trends for the three variable combinations and the other two variable combinations is relatively small.
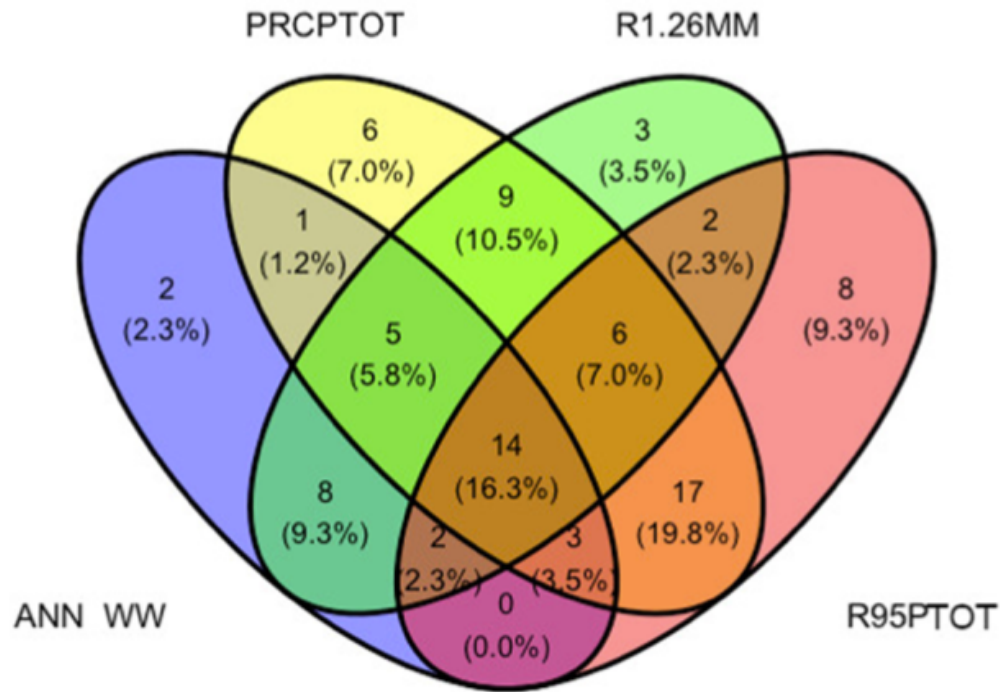
Figure 3.7 Venn Diagram of the number of stations with significant positive trends for all possible combinations of four representative annual precipitation indicators: the probability of wet-wet days (WW), total annual precipitation (PRCPTOT), the number of wet days (R1.26mm), total precipitation on days with precipitation $\geq 95^{th}$ percentile (R95pTOT). Percentages are relative to largest number of significant positive trends which was 86. Percentage of significant ($p \leq 0.10$) positive trends falling in each category is shown in parentheses.

For brevity, seasonal results are shown only for PRCPTOT, R95pTOT, WW, and R1.26mm, which capture the breadth of the different precipitation indicators. As for the annual precipitation characteristics, we observe that in all seasons the number of significant positive trends at all significance thresholds considered substantially exceeds the number of significant negative trends for the selected indicators (Table 3.3). However, the proportion of stations with significant trends varies by season. The seasonal trends for PRCPTOT indicate that no single season is solely responsible for the annual increase in precipitation observed at the majority of the stations. Significant ($p \leq 0.10$) positive trends are observed at over 35% of the stations in fall and winter, 30% of the stations in summer, and 20% of stations in spring. With the exception of one location in winter, no significant negative trends in seasonal PRCPTOT are evident. For R95pTOT, significant positive trends are evident at 33% of the stations in fall but at only approximately 20% of stations in the other seasons. For WW, over 24% percent of the station locations have significant positive trends in summer, fall, and winter, whereas only 15% of all stations observed in spring have positive trends. When the threshold for significance is weaker ($p < 0.20$), the number of significant trends increases substantially for most indicators in most seasons. Almost all of the additional significant trends that emerge by lowering the threshold are positive in sign. The number of significant trends in any one season is typically less than the number of significant trends for the corresponding annual indicator. Under the weak ($p \leq 0.20$) threshold, no individual seasonal indicator presents significantly positive trends at more than 50% of stations, with the exception of R1.26mm in fall Similar to the annual indicators, when the threshold for significance is stricter ($p \leq 0.05$), fewer stations exhibit statistically positive trends, while the number of statistically significant negative trends remains largely unchanged.

Table 3.3 Number of stations exhibiting statistically significant trends (Mann Kendall, p ≤ 0.05 two-tailed, p ≤ 0.10 two-tailed, p ≤ 0.20, two-tailed) from 1951-2019 in four representative seasonal indicators: total seasonal precipitation (PRCPTOT), the number of wet days (R1.26mm), the count of wet-wet days (WW), and the total precipitation on days with precipitation ≥ 95th percentile (R95pTOT). Indicators where more than 50% of stations analyzed showed a significant trend are shown in bold.

Precipitation Indicator	Season	Total number of stations after breakpoint analysis	Number of stations with significant positive trends (p≤0.05)	Number of stations with significant positive trends (p≤0.10)	Number of stations with significant positive trends (p≤0.20)	Number of stations with significant negative trends (p≤0.05)	Number of stations with significant negative trends (p≤0.10)	Number of stations with significant negative trends (p≤0.20)
PRCPTOT								
	Annual	*100*	*47*	**75**	**79**	*1*	*1*	*1*
	Spring	112	12	21	31	0	0	5
	Summer	111	20	31	49	0	0	2
	Fall	114	34	42	50	0	0	0
	Winter	104	22	32	44	0	0	1
R1.26mm								
	Annual	*97*	*42*	**53**	**61**	*1*	*3*	*5*
	Spring	114	10	15	21	4	6	12
	Summer	113	21	31	37	0	1	3
	Fall	111	25	33	**66**	0	0	1
	Winter	110	23	27	35	1	2	5
WW								
	Annual	*105*	*27*	*44*	*51*	*5*	*7*	*9*
	Spring	113	8	17	26	4	5	9
	Summer	111	16	24	35	4	6	9
	Fall	111	18	27	41	0	1	1
	Winter	106	18	25	39	0	0	1
R95pTOT								
	Annual	*104*	*47*	**62**	**63**	*0*	*0*	*0*
	Spring	111	12	22	37	0	0	0
	Summer	114	15	22	37	0	0	3
	Fall	114	28	38	56	0	0	0
	Winter	111	15	22	37	0	0	0

Seasonal variations in the spatial patterns in trends over time are shown for PRCPTOT, R95pTOT, and WW. The spatial distributions for R1.26mm are not shown as they are similar to those for WW. Large between season differences in the spatial variability of the temporal trends

are evident. For instance, locations with significant ($p \leq 0.10$) positive trends in seasonal

PRCPTOT are distributed across the study area in fall but are largely confined to the vicinity of

the Great Lakes (Wisconsin, the Lower Peninsula of Michigan, northeastern Ohio, western

Pennsylvania, western New York) in winter (Figure 3.8). In spring, most of the significant

positive trends are found in the western two thirds of the study region with few significant trends

in New York, Pennsylvania, and Ohio, whereas in summer the greatest density of significant

trends along with the largest trend magnitudes are found in the eastern and central portions of the

study region. For most stations, significant positive trends are observed in only one or two

seasons. A significant negative trend across all seasons is observed at only one location.
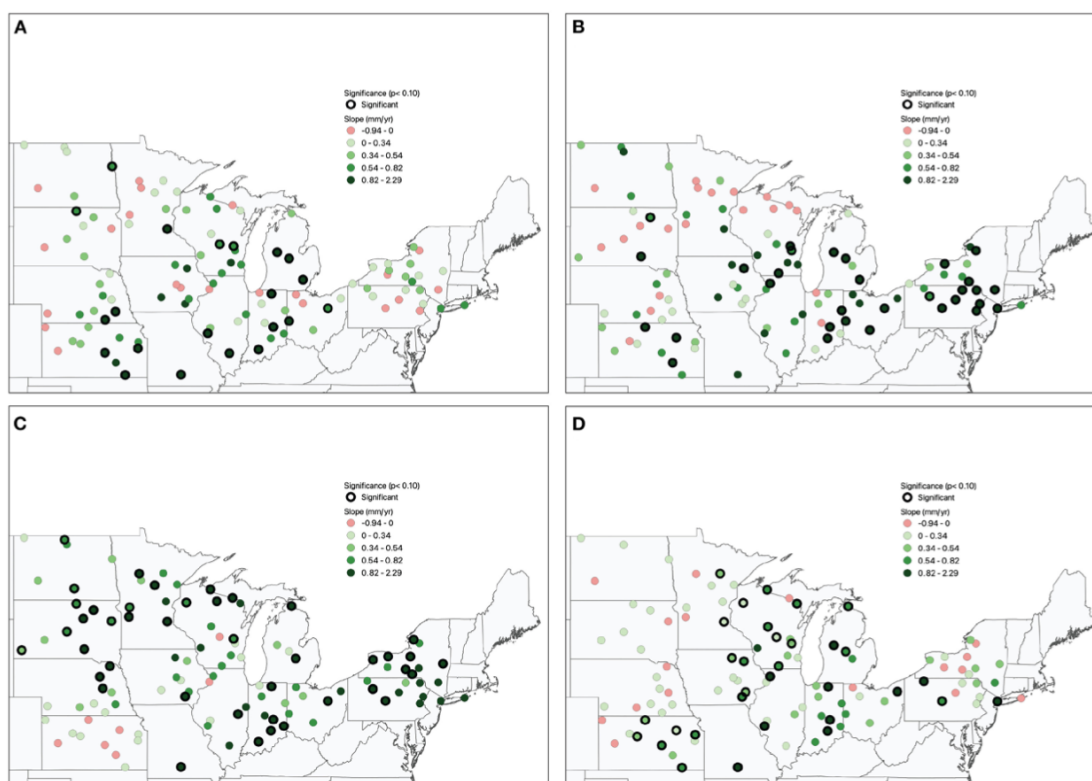
Figure 3.8 Trends (mm year[-1]) in seasonal total precipitation (PRCPTOT) for: a) spring, b) summer, c) fall, and d) winter.

The seasonal trends of R95pTOT display less spatial coherence when compared to seasonal PRCPTOT and to annual R95pTOT, with locations with significant positive trends often surrounded by locations with insignificant positive, and sometimes insignificant negative, trends (Figure 3.9). The number of locations in winter with significant positive trends is relatively small and these locations are mostly found in the vicinity of the western Great Lakes. The spatial extent of significant positive trends expands in spring to include most of the southern and eastern portions of the study area, with few significant trends evident in the northwestern portion of the study area. In summer, locations with significant positive trends are clustered in New York/Pennsylvania, Ohio/Indiana, and southern Wisconsin. The largest magnitude trends in R95pTOT are generally observed during the summer months. Significant positive trends are evident in fall across much of the area except for the extreme western portion of the study region and in the Lower Peninsula of Michigan. Although negative trends are evident for a number of locations in all seasons for R95pTOT, these trends are significant at only one location in spring and two locations in winter.
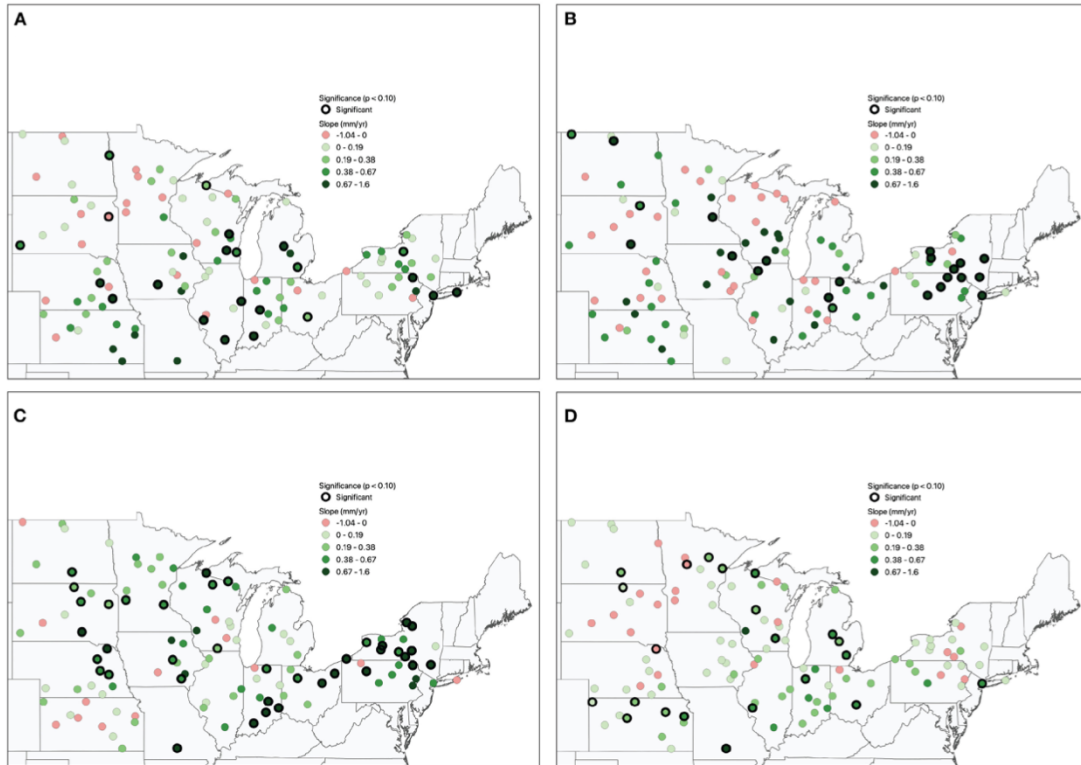
Figure 3.9 Trends (mm year^{-1}) in the seasonal amount of total precipitation falling on days with precipitation $\geq$ 95th percentile (R95pTOT) for a) spring, b) summer, c) fall, and d) winter.

Significant trends in seasonal WW are less spatially coherent than the annual WW indicator (Figure 3.10). As with seasonal R95pTOT, significant positive trends are often surrounded by insignificant trends or in a few cases significant negative trends. In general, stations with significant positive trends are more clustered for WW than for seasonal R95pTOT but less so than for seasonal PRCPTOT. The number of stations with significant positive trends is small in spring, and there are several (5) significant negative trends. The stations with significant positive trends are relatively dispersed, although some clustering is evident near the center of the study region. A more distinct spatial pattern is present in summer. Stations with significant positive trends are concentrated in Iowa, Indiana, Wisconsin, Ohio, Pennsylvania, and New York. In contrast, mostly insignificant trends are evident throughout the Plains states and east across Minnesota and northern Wisconsin, and the few significant trends in this area are

76

negative. In fall, significant positive trends in WW are found across the northern half of the study

region, whereas mostly insignificant trends of mixed sign are observed for the southern half of

the region with the exception of Illinois. Little spatial coherence is evident in the wintertime

trends of WW, other than some clustering of significant positive trends in the central and

extreme northeast sections of the study region.

Figure 3.10 Trends (days year[-1]) in the seasonal count of wet-wet-day sequences (WW) for a) spring, b) summer, c) fall, and d) winter.

When compared to annual indicators, the Venn diagrams of seasonal indicators show that

the groupings of indicators are more dispersed among the possible combinations of the four

representative indicators. (Figure 3.11). In fall and winter the three-indicator combination of

PRCPTOT, WW, and R1.26MM and the two-indicator combination of PRCPTOT and R95pTOT

are more frequent, while in summer the most common combination is PRCPTOT and R1.26MM.

In spring, locations are less likely compared to the other seasons to experience significant positive trends for two or more of the representative precipitation indicators, in part a reflection the smaller number of significant trends. For all seasons except spring a substantial number of stations display a significant trend only for R95pTOT. Venn diagrams can also be used to assess whether individual locations are likely to experience significant trends in a particular indicator during more than one season. Our results indicate that, regardless of the indicator type, significant positive trends are most likely to be observed during only one season (Figure 3.12).
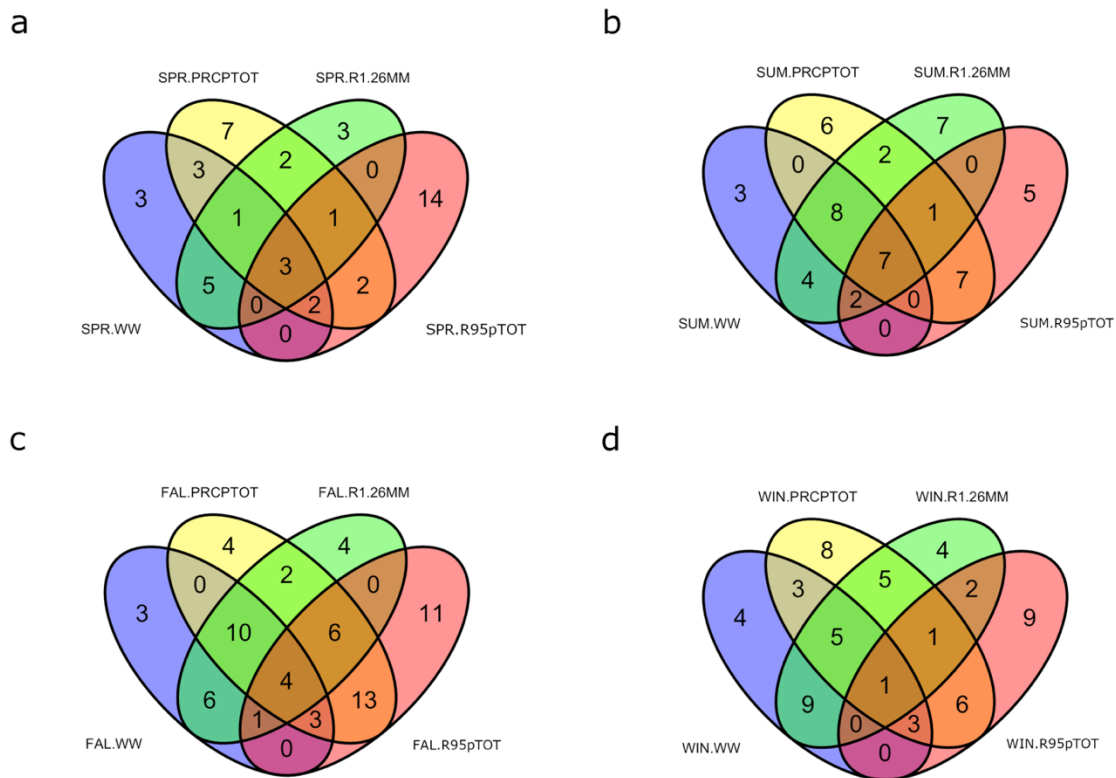
Figure 3.11 Venn diagram of the number of stations with significant ($p \leq 0.10$) positive trends for all possible combinations of four representative seasonal precipitation indicators (the probability of wet-wet days (WW), total precipitation (PRCPTOT), the number of wet days (R1.26mm), and total precipitation on days with precipitation $\geq 95^{th}$ percentile (R95pTOT) for a) spring, (b) summer, c) fall, and d) winter.
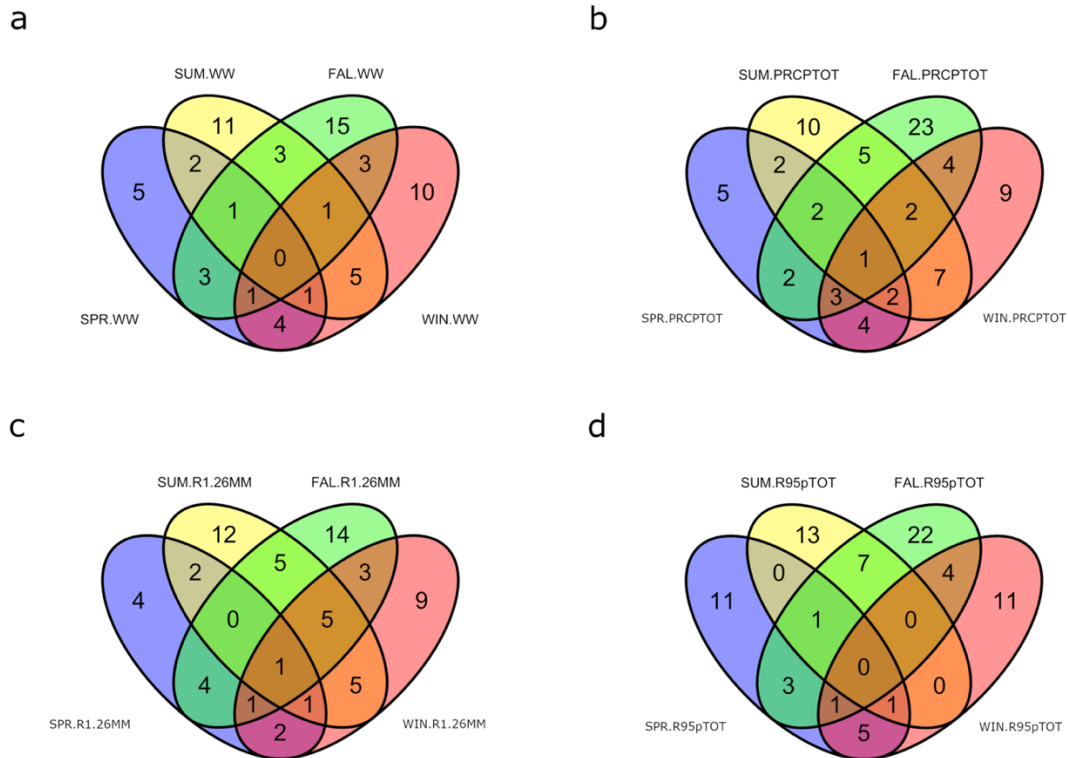
Figure 3.12 Venn diagram of the number of stations with significant (p ≤ 0.10) positive trends during one or more seasons for four representative seasonal precipitation indicators: a) the probability of wet-wet days (WW), b) total precipitation (PRCPTOT), c) the number of wet days (R1.26mm), and the total precipitation on days with precipitation≥ 95th percentile (R95pTOT).

Total Precipitable Water

Trends in annual daily mean TOTPRCPWAT during the 1951-2019 study period are positive in sign and significant over the southern two-thirds of the study region (Figure 3.13). The largest trends are in the south-central portion of the study region with the smallest trends located over the central and western Great Lakes. A significant increase in TOTPRCPWAT is also evident over the Great Plains, with the magnitude of the trend decreasing from south to north. Correlations between the trends in annual TOTPRCPWAT and the trends in the annual values of the four representative precipitation indicators are weak to moderate (Table 3.4; Figure 3.14), as indicated by the parametric Pearson's r and non-parametric Kendall's τ correlation

79

coefficients. Correlations for the annual trends are insignificant ($p \leq 0.05$) and negative for WW and R1.26mm, and insignificant but positive for PRCPTOT. Only the correlation between the annual trends in TOTPRCPWAT and those in R95pTOT is significant, with the sign of the correlation indicating a positive association between the annual trends of these two variables.
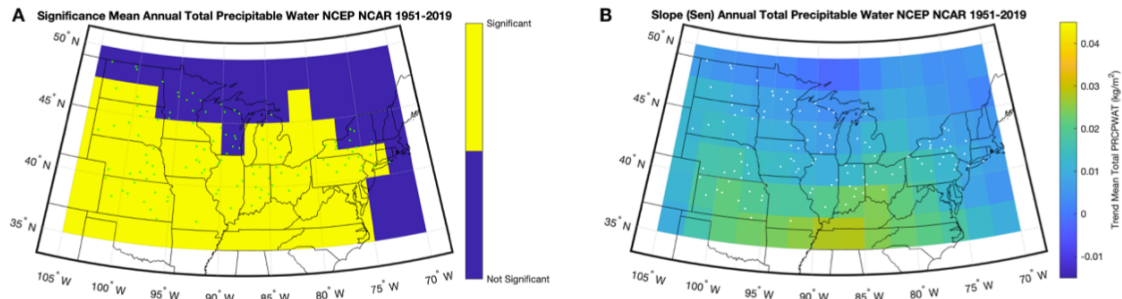
Figure 3.13 a) Significance ($p \leq 0.05$) and b) Sen's slope (kg m^{-2} yr^{-1}) of the trend in annual mean daily total precipitable water for the study region from 1951-2019. The stations with quality-controlled precipitation time series are shown on both maps as dots.
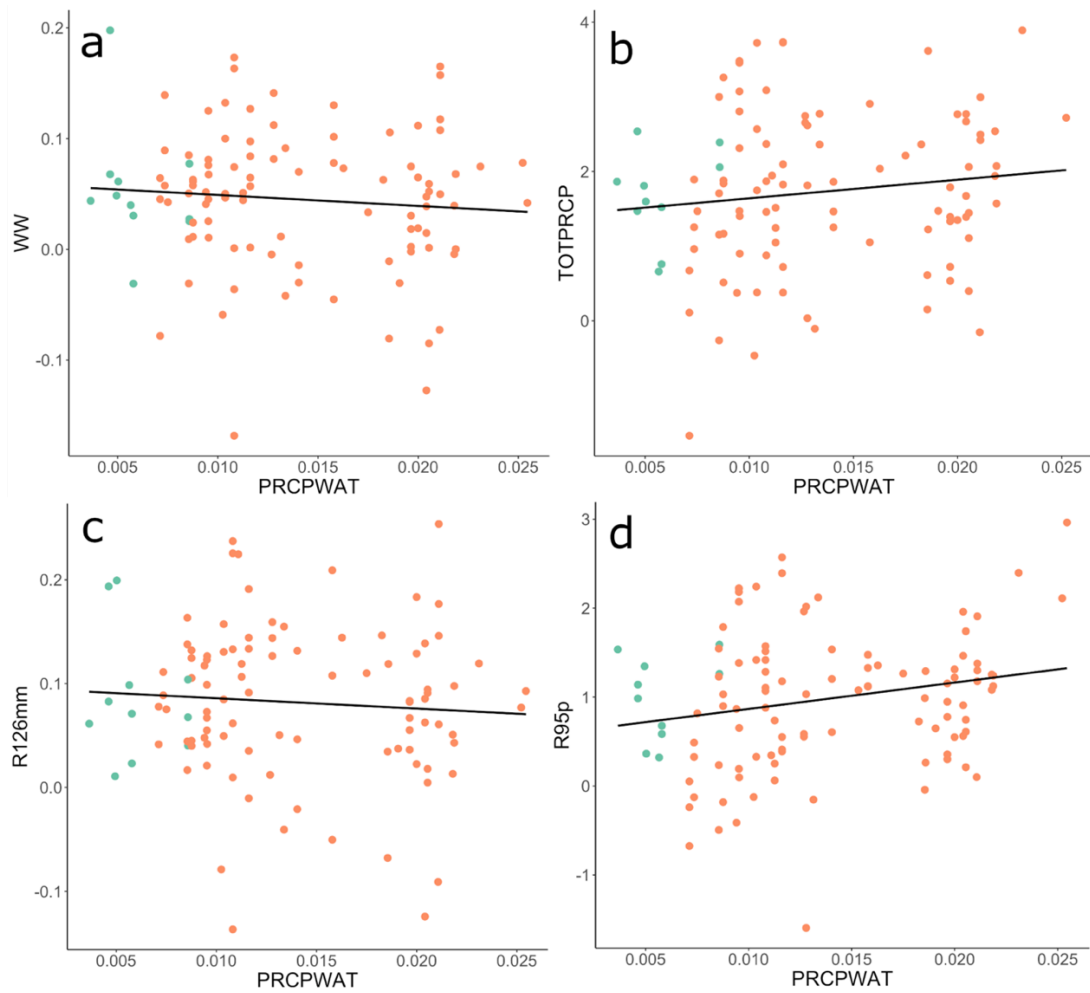
Figure 3.14 Scatter plots of annual trends (1951-2019) in mean total precipitable water and annual trends in WW (a), PRCPTOT (b), R1.26MM (c), R95pTOT (d). Significant (p ≤ 0.05) trends in mean total precipitable water are shown in orange, insignificant trends are shown in green. Estimated least squares regression line shown in black.

Table 3.4. Pearson correlation coefficients (r) and Kendall rank correlation coefficients (τ) between annual and seasonal trends from 1951-2019 in precipitation indicators and total precipitable water. Significant correlations ($p \leq 0.05$) are noted in bold. P-values for Pearson's r are noted in p_r and Kendall's τ are noted under pτ.

Indicator	Season	r	p_r	τ	pτ
WW	Spring	0.11	0.26	0.062	0.33
	Summer	**0.37**	**< 0.01**	**0.23**	**<0.01**
	Fall	-0.18	0.06	-0.10	0.13
	Winter	0.036	0.71	0.03	0.68
	Annual	*-0.089*	*0.37*	*-0.023*	*0.73*
PRCPTOT	Spring	**0.33**	**< 0.01**	**0.21**	**<0.01**
	Summer	**0.4**	**< 0.01**	**0.26**	**<0.01**
	Fall	-0.14	0.15	**-0.13**	**0.05**
	Winter	**0.25**	**0.01**	**0.18**	**0.01**
	Annual	*0.13*	*0.20*	*0.097*	*0.15*
R1.26MM	Spring	**0.2**	**0.04**	**0.15**	**0.02**
	Summer	**0.42**	**< 0.01**	**0.27**	**<0.01**
	Fall	-0.086	0.36	-0.04	0.55
	Winter	0.058	0.55	0.01	0.85
	Annual	*-0.076*	*0.45*	*-0.015*	*0.82*
R95pTOT	Spring	**0.19**	**0.05**	0.06	0.33
	Summer	**0.26**	**0.01**	**0.17**	**0.01**
	Fall	-0.13	0.16	**-0.16**	**0.014**
	Winter	**0.32**	**< 0.01**	**0.22**	**<0.01**
	Annual	***0.21***	***0.03***	***0.15***	***0.03***

Assessment of the possible contribution of seasonal trends in TOTPRCPWAT to seasonal trends in the representative precipitation indicators is complicated by seasonal variations in the significance of the TOTPRCPWAT trends, although in general correlation coefficients at the seasonal time scale are larger than those at the annual scale. Significant ($p \leq 0.05$) positive trends in TOTPRCPWAT (Figure 3.15) are evident during spring, summer, and fall for portions of the study area, although insignificant trends are observed for a substantial number of the reanalysis grid cells, with the location of the insignificant trends varying by season. In contrast, trends in TOTPRCPWAT are insignificant for all grid cells in winter. Most of the significant ($p = 0.05$)

correlations between the seasonal trends in TOTPRCPWAT and the season trends in the precipitation indicators are positive, although the significance of the trends varies by season and indicator. Correlations between the seasonal trends are significant in spring (PRCPTOT, R1.26mm, R95pTOT), summer (WW, PRCPTOT, R1.26mm, R95pTOT), and winter (PRCPTOT and R95pTOT), although the significant winter trends should be treated cautiously given the weak trends in TOTPRCPWAT at this time of year. No significant correlations were observed in the fall under Pearson's r. When Kendall's τ is used, correlations for fall are significant and negative for PRCPTOT and R95pTOT.
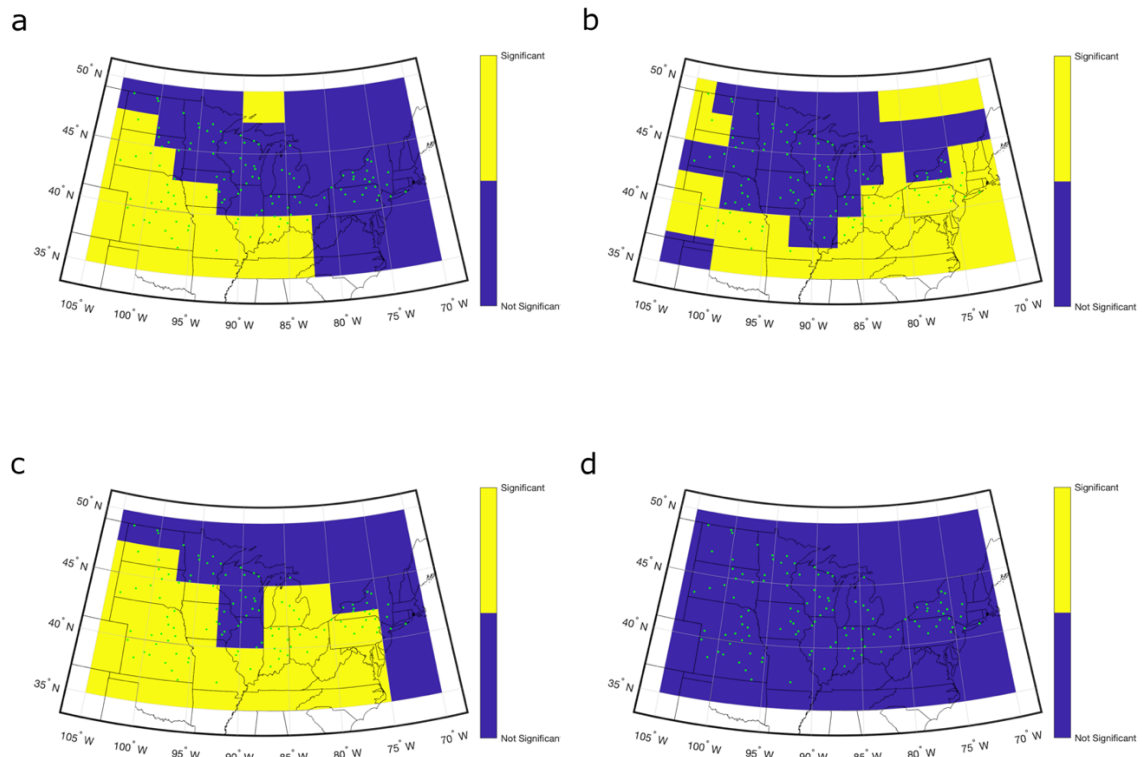
Figure 3.15 Significance (p<0.05) of seasonal trend in mean daily total precipitable water (kg m^{-2} yr^{-1}) for 1951-2019: a) spring, b) summer, c) fall, and d) winter. The stations with quality-controlled precipitation time series are shown on both maps as dots.

Discussion/Conclusion

The impact of the additional quality control measures on the number of stations available for precipitation trend analysis is striking. Of the 317 stations in the Midwest and Great Lakes region that met the initial criterion of 90% completeness, 203 stations were removed at the second step because they failed the tests for observer bias (underreporting of precipitation ≤ 1.26 mm and overreporting of precipitation amounts divisible by 5 or 10 when precipitation is recorded in inches). In contrast, the breakpoint analyses, which were conducted separately for each precipitation indicator in recognition that discontinuities can impact the indicators differently, removed only a small portion of the remaining stations (17 or fewer, depending on the indicator). This is a somewhat surprising result given the well documented discontinuities in

observations from the United States Cooperative Observer Network (Karl & Williams 1987; Winkler 2004; Menne et al., 2010), which is the largest source of precipitation data for the United States in the GHCN-D database (Menne et al. 2012). One interpretation is that many of the precipitation time series were affected by both observer bias and discontinuities and were removed following the tests for observation bias. The number of stations with breakpoints was largest for the "accumulation" and "threshold" precipitation indicators, suggesting the tests for observation bias did not remove all afflicted time series for these indicators. The final suite of quality-controlled time series has a much coarser station density than the datasets used in previous studies, and, while not suitable for investigating local-scale variations in precipitation trends, provides high confidence in the estimation of regional-scale variations. The quality-control routines implemented here also allow for more confidence in trends across the range of indicators from high frequency light events to low frequency extreme events, as observer bias affects various indicators differently and may not be captured in studies relying solely on data completeness and documented changes for data screening.

One finding from the use of the carefully quality-controlled time series is that the estimated trends for 1951-2019 in the Midwest and Great Lakes region are predominantly positive for all the "wet" precipitation indicators and negative for the "dry" precipitation indicators. In fact, there is a near absence of significant negative trends across the region for all indicators, with the exception of DD and CDD, and for all seasons and at all three significance levels included in the analysis. On the other hand, the proportion of stations with significant positive trends varies by precipitation indicator, season, and significance level. In general, significant trends at the moderate ($p \leq 0.10$) significance level are most likely for the indicators involving precipitation accumulation and counts of days with precipitation above specified

thresholds, and less likely for indicators of maximum reported precipitation and the indictors defined in terms of the sequencing of precipitation. Thus, users need to be cautious of inferring from significant trends in common precipitation characteristics, such as total precipitation, that significant trends are also occurring in other precipitation characteristics at a particular location. The larger number of significant positive trends for the "wet" indicators under the weak ($p \leq$ 0.20) significant level obviously need to be interpreted cautiously because of the greater probability of a Type I error (rejecting the null hypothesis of no trend when it is true). However, the greater spatial coherence of the locations with significant trends for the weak significance level compared to the moderate and stringent levels is consistent with a regional-scale trend toward a wetter climate that is emerging from interannual variability.

Our results also confirm that precipitation indicators that are defined annually often mask strong seasonal variations in the temporal trends of both high frequency, low magnitude events and low frequency, high magnitude events. For almost all locations, one cannot assume based on the trends in an annual precipitation indicator that a location is experiencing similar trends seasonally. Instead, a significant trend in a particular precipitation indicator typically is observed during only one season.

While the low spatial density of the stations that met all three of the quality control criteria somewhat constrains inferences regarding subregional variations in precipitation trends, our results, especially those using the weaker and moderate significance levels, suggest that the character of precipitation is not changing uniformly across the Midwest and Great Lakes region. In terms of the annual values of four representative precipitation indicators (PRCPTOT, R1.26mm, WW, R95pTOT), significant positive trends are observed across the central and eastern portions of the study region for all four indicators, whereas in the west there is a notable

absence of significant positive trends for R1.26mm events. Seasonal differences in the spatial distribution of significant trends are also evident, particularly for winter when significant trends for the four representative indicators are largely confined to western Great Lakes portion of the study region. The smaller number of significant trends present under the strict criteria, highlights the strength and relative cohesiveness of trends in precipitation in the central and eastern portions of the region, where most of the significant ($p \leq 0.05$) trends are located.

The quality-controlled time series are also useful for evaluating relationships between trends in the precipitation characteristics and physical processes potentially contributing to these trends. Expanding on the intriguing findings of Kunkel et al. 2020b who found a significant positive correlation between regionally-averaged trends in extreme precipitation and trends in precipitable water for the contiguous United States, we correlated, at annual and seasonal temporal scales, the trends in PRCPTOT, R1.26mm, WW, and R95pTOT for the quality-controlled station time series with trends in average daily precipitable water at a 2.5° latitude x 2.5° longitude resolution from the NCEP/NCAR reanalysis (Kalnay et al. 1996). The correlations for R95pTOT support for the Midwest and Great Lakes region the coarser-scale findings from Kunkel et al 2020a that the trend in extreme precipitation increases with an increasing trend in precipitable water, but also point to a more complex interpretation of the relationship between in trends in precipitation characteristics and trends in precipitable water for the study region. In particular, significant ($p \leq 0.05$) correlations are evident during spring and summer for PRCPTOT and R1.26mm and in summer for WW, suggesting that increases in precipitable water may also contribute to positive trends in high frequency precipitation events and even to the sequencing of wet days. Also, the correlation between the trend in R95pTOT and that for precipitable water is insignificant in fall for the parametric correlation coefficient and significant

but negative for the non-parametric correlation coefficient, suggesting that changes in atmospheric lifting mechanisms (e.g. fronts, extratropical cyclones) rather than increased atmospheric humidity may be more important for explaining the positive trend in R95pTOT in the Midwest and Great Lakes region in fall. Our findings of insignificant trends in precipitable water for large portions of the study area, especially in winter when the precipitable water trends are insignificant for the entire NCEP/NCAR grid over the study area, point to the need for cautious interpretation of the relationship between trends in precipitable water and trends in precipitation characteristics.

We have demonstrated the usefulness of quality-controlled precipitation time series for evaluating trends in precipitation characteristics and for investigating their relationship with processes. However, the limitations of the quality-controlled dataset should also be considered in interpreting the findings presented here and when applying the time series in future work. A key limitation is the coarse spatial resolution of the quality-controlled time series, limiting their usefulness in investigating potential contributions of local-scale features such as lake surfaces or topography on trends in precipitation characteristics. Another concern is that identified breakpoints in the time series that are attributed to changes in instrumentation, station moves or observation protocols may instead be caused by changes in circulation regimes. Also, some types of precipitation indicators may be less sensitive to observer bias than others, and a less stringent protocol for removing time series for consideration would be appropriate. Moreover, for any quality control routine that is not manual, there are almost always time series with data issues relevant to a particular analysis that pass through filters and checks and those without data issues that are incorrectly removed.

In sum, our analysis focused on quality control of station time series to improve the quality of data prior to analysis. As a result of this effort, the trends in our study tended to exhibit a more cohesive spatial and temporal similarities when compared with studies with different quality control criteria, illustrating the importance of quality control of observations in climatic studies. Also, our results indicate, at least for the Midwest and Great Lakes region, that not only is extreme precipitation increasing but the entire distribution of precipitation has been shifting upward over time.

REFERENCES

REFERENCES

Alexander, L. V., Zhang, X., Peterson, T. C., Caesar, J., Gleason, B., Klein Tank, A. M. G., et al. (2006). Global observed changes in daily climate extremes of temperature and precipitation. *J. Geophys. Res.* **111**, D05109. doi:10.1029/2005JD006290.

Angel, J., C. Swanston, B.M. Boustead, K.C. Conlon, K.R. Hall, J.L. Jorns, K.E. Kunkel, M.C. Lemos, B. Lofgren, T.A. Ontl, J. Posey, K. Stone, G. Takle, and D. Todey. (2018). Midwest. In Impacts, Risks, and Adaptation in the United States: Fourth National Climate Assessment, Volume II [Reidmiller, D.R., C.W. Avery, D.R. Easterling, K.E. Kunkel, K.L.M. Lewis, T.K. Maycock, and B.C. Stewart (eds.)]. U.S. Global Change Research Program, Washington, DC, USA, pp. 872–940. doi: 10.7930/NCA4.2018.CH21

Attavanich, W., McCarl, B. A., Ahmedov, Z., Fuller, S. W., and Vedenov, D. V. (2013). Effects of climate change on US grain transport. *Nat. Clim. Chang.* **2013** *37* 3, 638–643. doi:10.1038/nclimate1892.

Bartels, R. J., Black, A. W., and Keim, B. D. (2020). Trends in precipitation days in the United States. *Int. J. Climatol.* **40**, 1038–1048. doi:10.1002/joc.6254.

Baule, W. J., and Shulski, M. D. (2014). Climatology and trends of wind speed in the Beaufort/Chukchi Sea coastal region from 1979 to 2009. *Int. J. Climatol.* **34**. doi:10.1002/joc.3881.

Bell, G. D., and Janowiak, J. E. (1995). Atmospheric Circulation Associated with the Midwest Floods of 1993. *Bull. Am. Meteorol. Soc.* **76**, 681–695. doi:10.1175/1520-0477(1995)076<0681:ACAWTM>2.0.CO;2.

Bowles, T. M., Atallah, S. S., Campbell, E. E., Gaudin, A. C. M., Wieder, W. R., and Grandy, A. S. (2018). Addressing agricultural nitrogen losses in a changing climate. *Nat. Sustain.* **1**, 399–408. doi:10.1038/s41893-018-0106-0.

Chin, N., Byun, K., Hamlet, A. F., and Cherkauer, K. A. (2018). Assessing potential winter weather response to climate change and implications for tourism in the U.S. Great Lakes and Midwest. *J. Hydrol. Reg. Stud.* **19**, 42–56. doi:10.1016/J.EJRH.2018.06.005.

Contractor, S., Donat, M. G., and Alexander, L. V. (2021). Changes in Observed Daily Precipitation over Global Land Areas since 1950. *J. Clim.* **34**, 3–19. doi:10.1175/JCLI-D-19-0965.1.

Costa, A. C., and Soares, A. (2009). Homogenization of Climate Data: Review and New Perspectives Using Geostatistics. *Math. Geosci.* **41**, 291–305. doi:10.1007/s11004-008-9203-3.

Dai, S., Shulski, M. D., Hubbard, K. G., and Takle, E. S. (2016). A spatiotemporal analysis of Midwest US temperature and precipitation trends during the growing season from 1980 to 2013. *Int. J. Climatol.* **36**, 517–525. doi:10.1002/joc.4354.

Daly, C., Gibson, W. P., Taylor, G. H., Doggett, M. K., and Smith, J. I. (2007). Observer bias in daily precipitation measurements at United States Cooperative Network Stations. *Bull. Am. Meteorol. Soc.* 88, 899–912. doi:10.1175/BAMS-88-6-899.

Donat, M. G., Alexander, L. V., Yang, H., Durre, I., Vose, R., Caesar, J., et al. (2013). Global Land-Based Datasets for Monitoring Climatic Extremes. *Bull. Am. Meteorol. Soc.* **94**, 997–1006. doi:10.1175/BAMS-D-12-00109.1.

Durre, I., Menne, M. J., Gleason, B. E., Houston, T. G., and Vose, R. S. (2010). Comprehensive Automated Quality Assurance of Daily Surface Observations. *J. Appl. Meteorol. Climatol.* **49**, 1615–1633. doi:10.1175/2010JAMC2375.1.

Durre, I., Menne, M. J., and Vose, R. S. (2008). Strategies for Evaluating Quality Assurance Procedures. *J. Appl. Meteorol. Climatol.* 47, 1785–1791. doi:10.1175/2007JAMC1706.1.

Easterling, D. R. (2002). United States Historical Climatology Network Daily Temperature and Precipitation Data (1871-1997). Oak Ridge, TN (United States): Oak Ridge National Laboratory doi:10.2172/814188..

Groisman, P. Y., Knight, R. W., Groisman, P. Y., and Knight, R. W. (2008). Prolonged Dry Episodes over the Conterminous United States: New Tendencies Emerging during the Last 40 Years. J. Clim. **21**, 1850–1862. doi:10.1175/2007JCLI2013.1.

Groisman, P. Y., Karl, T. R., Easterling, D. R., Knight, R. W., Jamason, P. F., Hennessy, K. J., et al. (1999). Changes in the Probability of Heavy Precipitation: Important Indicators of Climatic Change. *Clim. Change* **42**, 243–283. doi:10.1023/A:1005432803188.

Gronewold, A. D., Do, H. X., Mei, Y., and Stow, C. A. (2021). A Tug-of-War Within the Hydrologic Cycle of a Continental Freshwater Basin. *Geophys. Res. Lett.* **48**, e2020GL090374. doi:10.1029/2020GL090374.

Guilbert, J., Betts, A. K., Rizzo, D. M., Beckage, B., and Bomblies, A. (2015). Characterization of increased persistence and intensity of precipitation in the northeastern United States. *Geophys. Res. Lett.* **42**, 1888–1893. doi:10.1002/2015GL063124.

Gutowski, W. J., Willis, S. S., Patton, J. C., Schwedler, B. R. J., Arritt, R. W., and Takle, E. S. (2008). Changes in extreme, cold-season synoptic precipitation events under global warming. *Geophys. Res. Lett.* **35**, L20710. doi:10.1029/2008GL035516.

Higgins, R. W., Silva, V. B. S., Shi, W., and Larson, J. (2007). Relationships between Climate Variability and Fluctuations in Daily Precipitation over the United States. *J. Clim.* **20**, 3561–3579. doi:10.1175/JCLI4196.1.

Hoerling, M., Eischeid, J., Perlwitz, J., Quan, X. W., Wolter, K., and Cheng, L. (2016). Characterizing Recent Trends in U.S. Heavy Precipitation. *J. Clim.* **29**, 2313–2332. doi:10.1175/JCLI-D-15-0441.1.

Huang, H., Winter, J. M., and Osterberg, E. C. (2018). Mechanisms of Abrupt Extreme Precipitation Change Over the Northeastern United States. *J. Geophys. Res. Atmos.* **123**, 7179–7192. doi:10.1029/2017JD028136.

Huang, H., Winter, J. M., Osterberg, E. C., Horton, R. M., and Beckage, B. (2017). Total and Extreme Precipitation Changes over the Northeastern United States. *J. Hydrometeorol.* **18**, 1783–1798. doi:10.1175/JHM-D-16-0195.1.

Hunt, E. D., Birge, H. E., Laingen, C., Licht, M. A., McMechan, J., Baule, W., et al. (2020). A perspective on changes across the U.S. Corn Belt. *Environ. Res. Lett.* 15, 071001. doi:10.1088/1748-9326/ab9333.

Huschke, R. E. (1959). *Glossary of Meteorology*. Boston: American Meteorological Society.

Ines, A. V. M., Hansen, J. W., and Robertson, A. W. (2011). Enhancing the utility of daily GCM rainfall for crop yield prediction. Int. J. Climatol. **31**, 2168–2182. doi:10.1002/joc.2223.

Jaiswal, R.L, Lohaniu, A.K., and Tiwari, H.L. (2015). Statistical Analysis for Change Detection and Trend Assessment in Climatological Parameters. *Environmental Processes*. 2, 729-749. doi: 10.1007/s40710-015-0105-3

Janssen, E., Wuebbles, D. J., Kunkel, K. E., Olsen, S. C., and Goodman, A. (2014). Observational- and model-based trends and projections of extreme precipitation over the contiguous United States. *Earth's Futur.* **2**, 99–113. doi:10.1002/2013EF000185.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., et al. (1996). The NCEP/NCAR 40-year reanalysis project. *Bull. Am. Meteorol. Soc.* **77**, 437–471. doi:10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2.

Karl, T. R., Williams, C. N., Karl, T. R., and Jr., C. N. W. (1987). An Approach to Adjusting Climatological Time Series for Discontinuous Inhomogeneities. *J. Clim. Appl. Meteorol.* **26**, 1744–1763. doi:10.1175/1520-0450(1987)026<1744:AATACT>2.0.CO;2.

Kendall, M. G. (1955). *Rank correlation methods*. 2nd ed. Oxford: Hafner Publishing Co. Available at: http://psycnet.apa.org/record/1956-03749-000.

Kiefer, M. T., Andresen, J. A., McCullough, D. G., Baule, W. J., and Notaro, M. (2021). Extreme minimum temperatures in the Great Lakes region of the United States: A climatology with implications for insect mortality. *Int. J. Climatol.* doi:10.1002/JOC.7434.

Komoto, K., Soldo, L., Tang, Y., Chilvers, M. I., Dahlin, K., Guentchev, G., et al. (2021). Climatology of persistent high relative humidity: An example for the Lower Peninsula of Michigan, USA. *Int. J. Climatol.* **41**, E2517–E2536. doi:10.1002/JOC.6861.

Konrad, C. E. (2001). The Most Extreme Precipitation Events over the Eastern United States from 1950 to 1996: Considerations of Scale. *J. Hydrometeorol.* 2, 309–325. doi:10.1175/1525-7541(2001)002<0309:TMEPEO>2.0.CO;2.

Kunkel, K. E., Stevens, S. E., Stevens, L. E., and Karl, T. R. (2020a). Observed Climatological Relationships of Extreme Daily Precipitation Events With Precipitable Water and Vertical Velocity in the Contiguous United States. *Geophys. Res. Lett.* **47**, e2019GL086721. doi:10.1029/2019GL086721.

Kunkel, K. E., Karl, T. R., Squires, M. F., Yin, X., Stegall, S. T., and Easterling, D. R. (2020b). Precipitation Extremes: Trends and Relationships with Average Precipitation and Precipitable Water in the Contiguous United States. *J. Appl. Meteorol. Climatol.* **59**, 125–142. doi:10.1175/JAMC-D-19-0185.1.

Legates, D. R., and Willmott, C. J. (1990). Mean seasonal and spatial variability in gauge-corrected, global precipitation. *Int. J. Climatol.* **10**, 111–127. doi:10.1002/joc.3370100202.

Mallakpour, I., and Villarini, G. (2016). A simulation study to examine the sensitivity of the Pettitt test to detect abrupt changes in mean. *Hydrol. Sci. J.* **61**, 245–254. doi:10.1080/02626667.2015.1008482/SUPPL_FILE/THSJ_A_1008482_SM3561.DOC.

Mallakpour, I., and Villarini, G. (2017). Analysis of changes in the magnitude, frequency, and seasonality of heavy precipitation over the contiguous USA. *Theor. Appl. Climatol.* **130**, 345–363. doi:10.1007/s00704-016-1881-z.

Mann, H. B. (1945). Nonparametric Tests Against Trend. *Econometrica* **13**, 245. doi:10.2307/1907187.

Menne, M. J., Williams, C. N., and Palecki, M. A. (2010). On the reliability of the U.S. surface temperature record. *J. Geophys. Res. Atmos.* **115**, 11108. doi:10.1029/2009JD013094.

Menne, M. J., Durre, I., Vose, R. S., Gleason, B. E., and Houston, T. G. (2012). An Overview of the Global Historical Climatology Network-Daily Database. *J. Atmos. Ocean. Technol.* **29**, 897–910. doi:10.1175/JTECH-D-11-00103.1.

Myhre, G., Forster, P. M., Samset, B. H., Hodnebrog, Ø., Sillmann, J., Aalbergsjø, S. G., et al. (2016). PDRMIP: A Precipitation Driver and Response Model Intercomparison Project, Protocol and preliminary results. *Bull. Am. Meteorol. Soc.* doi:10.1175/BAMS-D-16-0019.1.

Onyutha, C. (2021). Graphical-statistical method to explore variability of hydrological time series. Hydrol. Res. 52, 266–283. doi:10.2166/NH.2020.111.

Pettitt, A.N., (1979). A non-parametric approach to the change-point problem. Appl. Stat., **28**(2), 126–135

Pielke, R. A., Jr., and Downton, M. W. (2000). Precipitation and damaging floods: Trends in the United States, 1932–1997. *J. Clim*, 3625–3637..

Pryor, S. C., Howe, J. A., and Kunkel, K. E. (2009). How spatially coherent and statistically robust are temporal changes in extreme precipitation in the contiguous USA? *Int. J. Climatol.* **29**, 31–45. doi:10.1002/joc.1696.

Riha, S. J., Wilks, D. S., and Simoens, P. (1996). Impact of temperature and precipitation variability on crop model predictions. *Clim. Change* 32, 293–311. doi:10.1007/BF00142466.

Roque-Malo, S., and Kumar, P. (2017). Patterns of change in high frequency precipitation variability over North America. *Sci. Rep.* **7**, 10853. doi:10.1038/s41598-017-10827-8.

Rosenzweig, C., Tubiello, F. N., Goldberg, R., Mills, E., and Bloomfield, J. (2002). Increased crop damage in the US from excess precipitation under climate change. *Glob. Environ. Chang.* **12**, 197–202. doi:10.1016/S0959-3780(02)00008-0.

Schoof, J. T., Pryor, S. C., and Surprenant, J. (2010). Development of daily precipitation projections for the United States based on probabilistic downscaling. *J. Geophys. Res. Atmos.* **115**, D13106. doi:10.1029/2009JD013030.

Sen, P. K. (1968). Estimates of the Regression Coefficient Based on Kendall's Tau. *J. Am. Stat. Assoc.* **63**, 1379. doi:10.2307/2285891.

Şen, Z. (2015). Innovative trend significance test and applications. Theor. Appl. Climatol. **2015** 1273 127, 939–947. doi:10.1007/S00704-015-1681-X.

Shulski, M. D., Baule, W., Stiles, C., and Umphlett, N. (2015). A historical perspective on Nebraska's variable and changing climate. *Gt. Plains Res.* **25**. doi:10.1353/gpr.2015.0023.

Sneyers, R. (1990) On the Statistical Analysis of Series of Observations; Technical Note no. 143, WMO no. 415; Secretariat of the World Meteorological Organization: Geneva, Switzerland, 1990.

Takle, E. S., and Gutowski, W. J. (2020). Iowa's agriculture is losing its Goldilocks climate. *Phys. Today* **73**, 26. doi:10.1063/PT.3.4407.

Talukder, B., and Hipel, K. W. (2020). Diagnosis of sustainability of trans-boundary water governance in the Great Lakes basin. *World Dev.* **129**, 104855. doi:10.1016/J.WORLDDEV.2019.104855.

Teale, N., and Robinson, D. A. (2020). Patterns of water vapor transport in the Eastern United States. *J. Hydrometeorol.* **21**, 2123–2138. doi:10.1175/JHM-D-19-0267.1.

Trenberth, K. E., Dai, A., Rasmussen, R. M., and Parsons, D. B. (2003). The Changing Character of Precipitation. *Bull. Am. Meteorol. Soc.* **84**, 1205–1217. doi:10.1175/BAMS-84-9-1205.

Walsh, J., Wuebbles, D., Hayhoe, K., Kossin, J., Kunkel, K., Stephens, G., et al. (2014). Ch. 2: Our Changing Climate. Climate Change Impacts in the United States: The Third National Climate Assessment. doi:10.7930/J0KW5CXT.

Wang, X. L., Chen, H., Wu, Y., Feng, Y., Pu, Q., Wang, X. L., et al. (2010). New Techniques for the Detection and Adjustment of Shifts in Daily Precipitation Data Series. *J. Appl. Meteorol. Climatol.* **49**, 2416–2436. doi:10.1175/2010JAMC2376.1.

Weaver, S. J., and Nigam, S. (2008). Variability of the Great Plains Low-Level Jet: Large-Scale Circulation Context and Hydroclimate Impacts. *J. Clim.* **21**, 1532–1551. doi:10.1175/2007JCLI1586.1.

Williams, C. N., Menne, M. J., and Thorne, P. W. (2012). Benchmarking the performance of pairwise homogenization of surface temperatures in the United States. *J. Geophys. Res. Atmos.* **117**, n/a-n/a. doi:10.1029/2011JD016761.

Winkler, J. A. (1988). Climatological Characteristics of Summertime Extreme Rainstorms in Minnesota. *Ann. Assoc. Am. Geogr.* **78**, 57–73. Available at: http://www.jstor.org/stable/2563440.

Winkler, J. A. (2004). "The Impact of Technology Upon In Situ Atmospheric Observations and Climate Science," in *Geography and Technology* (Springer Netherlands), 461–490. doi:10.1007/978-1-4020-2353-8_20.

Wu, S. Y. (2015). Changing characteristics of precipitation for the contiguous United States. *Clim. Change* **132**, 677–692. doi:10.1007/S10584-015-1453-8/TABLES/2.

Zhang, W., and Villarini, G. (2019). On the weather types that shape the precipitation patterns across the U.S. Midwest. *Clim. Dyn.*, 1–16. doi:10.1007/s00382-019-04783-4.

Zhang, X., Alexander, L., Hegerl, G. C., Jones, P., Tank, A. K., Peterson, T. C., et al. (2011). Indices for monitoring changes in extremes based on daily temperature and precipitation data. *Wiley Interdisc. Rev. Clim. Chang.* **2**, 851–870. doi:10.1002/wcc.147.

CHAPTER 4. WITHIN-FIELD SIMULATION OF NITRATE LEACHING FOR MILLIONS OF FIELDS: AN INTERCONNECTED STUDY OF YIELD, MANAGEMENT, AND CLIMATE IN THE MIDWEST

Abstract

Agricultural fields traditionally have been treated as a uniform unit in terms of management. Recent studies have highlighted the sub-field variability of yield over time and potential benefits of variable management. Here we take a novel process-based simulation approach to simulate maize-soy production, while incorporating sub-field variability over millions of fields in the Midwestern United States. Simulations suggest that nitrate leaching has multiple drivers across the region and that these vary over space. The largest sources of nitrate leaching appear to be low stable and unstable yield zones within the states of Iowa, Illinois, and Indiana. Relative behavior of the zones at the sub-field spatial scale reveals that unstable zones are more sensitive to climate and weather than stable zones and suggest in-season management may be more effective in these areas. Analysis of climate trends in conjunction with model output suggests that precipitation regime and frequency of water stress are climatic variables most correlated with nitrate loss in all zones, with the strongest correlations present in unstable zones. While climatic trends and forcing have a substantial impact on nitrate leaching result, with increased precipitation and increased water stress resulting in greater leaching, these numbers are greatly affected by management, physical/biogeochemical properties of the soil, and position in the topography of the region. This novel approach to sub-field simulation of yield and nitrate leaching provides a framework to advance regional knowledge of potential sources of nitrate loss and the development of management strategies aimed at sustaining yields and reducing nitrogen loss.

Introduction

Since the late twentieth century through present, there has been an observed increase in annual precipitation throughout the Midwest region (Walsh et al. 2014) defined here as the 14-

98

state region encompassing: North Dakota, South Dakota, Nebraska, Kansas, Arkansas, Missouri, Iowa, Minnesota, Wisconsin, Illinois, Indiana, Michigan, Ohio, and Pennsylvania. At the same time there has been a regional increase in the amount of corn (*Zea mays* L.) and soybeans (*Glycine max*) produced and in the area planted, particularly in the northern and western portions of the region and in areas that historically have been considered marginally-productive for these crops (Mladenoff et al. 2016; Hatfield 2012). Typically, soils on these marginal lands have relatively lower organic matter levels and textures less suitable for agricultural production. If the soils have a higher sand content and coarser texture, they are inherently more susceptible to nitrate leaching from inorganic fertilizers (Basso et al. 2016; Zotarelli et al. 2007). Given economic and environmental concerns regarding nitrate loss and pollution (Kalkhoff et al. 2016), it is important to understand how climate and its inherent variability with respect to precipitation impacts nitrate loss on soils that are particularly vulnerable to nitrate leaching. In addition to the expansion on marginal lands, the corn belt as a whole has geographically shifted and expanded to include areas in primary production areas previously dominated by small grain crops and grasslands in the northern Great Plains (Lin and Henry 2016; Wimberly et al. 2017). This is in addition to a general increase in the area within the corn belt region being planted in either corn or soybeans in recent decades (Johnston 2014; Lark et al. 2015).

Given the recently observed trends in precipitation from multiple studies previously mentioned, our major objective is an exploration of the role of weather and climate in determining the spatial and temporal characteristics of nitrate loss and crop yields across the Midwest under historical climatic conditions. Our approach includes the application of a deterministic, process-based crop model that can help allow for a better understanding of the

complex relationships between weather, climate, agricultural management, production, and nutrient dynamics (Jones et al. 2016; Wallach et al. 2019).

Traditionally, studies on nitrogen loss in the Midwest and adjacent regions have relied on field experiments and field-level based crop simulation studies focused on one or a few locations (with similar climatic characteristics) over a relatively short periods of time such as a few growing seasons (Basso and Ritchie 2005; Basso et al. 2007; Congreves et al. 2016; Gerakis et al. 2006; Iqbal et al. 2018; Smith et al. 2019; Syswerda et al. 2012) and focus on the impacts of different fertilizer treatments (i.e. rate, timing, form of application), differences in rotation, or differences between coarse or fine-textured soils. These studies have provided a substantial amount of evidence for the mechanistic drivers of nitrogen loss at the field and sub-field scales.

Focusing less on nitrogen loss, Li et al. (2015) and Liu et al. (2011) carried out long-term validations of the DSSAT-CENTURY (Gijsman et al. 2002), a process-based crop model, on two long-term agricultural experiments with corn and wheat at a humid, semi-arid site in Canada and found that it satisfactorily simulated carbon, nitrogen, and water dynamics in areas with contrasting climatic regimes over long cropping sequences given adequate fertilization of the cropping systems considered.

Fewer studies have directly assessed the impacts of weather and climate variability on nitrogen loss/leaching in agricultural systems using a process based modeling approach, which allows researchers to simulate sequences of a cropping system with weather and climatic conditions that may not have occurred at field locations during a period of time when observations were taken or to explore the potential impacts of climate change on a system. In most cases these studies have also been limited to one or two locations with similar climatic regimes and the use of a variety of approaches to evaluate linkages between weather, climate,

and nitrogen loss in a variety of field cropping agricultural systems. Two recent examples

employing the DeNitrification and DeComposition (DNCD) biogeochemical model (Li 2000)

examine the linkage between climate variability on corn production for grain (Congreves et al.

2016) and silage (Smith et al. 2019) under conventional and best management practices at

locations in Southeast Ontario and Iowa, respectively. In each study the authors noted that

nitrogen dynamics varied substantially with different management practices, with respect to

climate variability the seasons with the highest leaching events across different soils and

management were associated with relatively higher daily temperatures, greater total

precipitation, and more frequent and intense precipitation events. Similarly, Iqbal et al. (2017)

used artificial sequences of extreme weather years in two-year sequences with the DAYCENT

model in Boone County, IA to examine the effects of extreme interannual sequences of weather

on nitrogen loss under different crop rotations and fertilizer rates. They found that regardless of

crop rotation or cover crop, extreme sequences of weather transitioning from wet to dry, or dry to

wet did not have identical effects on nitrogen loss (either through denitrification or nitrate

leaching), even though the total amount of precipitation was identical in each scenario which

suggested that the sequencing of precipitation is a key factor impacting nitrogen loss. The

authors also found the largest total simulated loss events were associated with wet-wet scenarios

at that location.

Additionally, there have been multiple studies in recent years that examine the

relationship between regional precipitation variability at interannual (Loecke et al. 2017; Sinha

and Michalak 2016) and interdecadal (Ballard et al. 2019) time scales and the correlation with

riverine nitrogen levels at the watershed scale. These studies noted a strong relationship between

interannual variability of precipitation (particularly wet springs following dry growing seasons)

and higher nitrogen concentrations in regional stream flow. Additionally, Ballard et al. (2019) noted an increase in stream nitrogen concentrations in and downstream from regions with general increases in the total amount and intensity of springtime precipitation. One weather-related index that has been significantly correlated with riverine nitrogen levels is the Weather Whiplash Index (Loecke et al., 2017), which is defined as the total of January through June precipitation minus the total of July through December of the previous year divided by the total amount of precipitation over the 12-month period. Additionally, riverine nitrogen levels have been found to have significant relationships with annual precipitation, total March to May precipitation above the historical 95th percentile, net anthropogenic nitrogen inputs, percentage of water shed covered in wetlands, and percentage in forest, shrubland, or herbaceous cover (Ballard et al., 2019; Sinha and Michalak, 2016), and temperature in March, April, and May (Ballard et al., 2019).

Regional approaches to process based crop simulations and statistical modeling exercises are not a new development in the field, however in many previous studies, fields were treated as individual homogenous units or aggregated to larger land units (e.g. Rosenzweig et al. 2018). In recent years, there have been several studies that have attempted to model/quantify sub-field variability within the context of large multi-state regions (Deines et al. 2021; Jin et al. 2019; Basso et al. 2019). Each study takes a unique approach to the quantification/representation of sub-field variability; however they share in common the idea that yield and process variability at sub-field scales is relevant at broad regional scales. Within these frameworks it is possible to begin asking question on scales that have policy-relevant implications within a region.

The studies mentioned in the previous paragraphs provide evidence for mechanisms of nitrogen loss in agricultural systems and their reflection in riverine nitrogen levels using a

combination of field observations (Deines et al. 2021; Jin et al. 2019; Basso et al. 2019), process-based agricultural systems models (Jin et al. 2019), and regression modeling (Deines et al. 2021). Field level studies provide mechanistic details and an understanding of nitrogen loss at the field scale. While the studies focused on riverine nitrogen loading provide a regional-scale perspective on the associations between weather and climatic variability and nitrogen loading, the climatology of nitrogen loss across the Midwest at the sub watershed scale and across space remain less clear. Using a combination of gridded climatological data and process based agricultural models run spatially over the Midwest, we can further elucidate the risks of nitrogen loss across the region and with varying climates and soil types, identifying areas where nitrogen losses are more or less likely to occur under observed climatological conditions (e.g. Bowles et al., 2018).

Methods

Study Region/CLU Delineation

For this study, the Midwest and Great Lakes region was defined as the states of Pennsylvania, Ohio, Indiana, Michigan, Illinois, Wisconsin, Minnesota, Iowa, Missouri, Arkansas, Kansas, Nebraska, South Dakota, New York, and North Dakota. These states represent a relatively contiguous zone of intensive agricultural activity, particularly in the cultivation of field crops. Following the methodology developed by Basso et al. (2019) which used the Cropland Data Layer (CDL), Google Earth Engine Landsat 5,7, and 8 Imagery, and Common Land Unit (CLU) datasets compiled by the United States Department of Agriculture (USDA), fields within the region where corn or soy were grown for at least 3 years during the period from 2010-2017 were identified, including fields with continuous corn, corn-soy, and less frequently-used complex crop rotations. Utilizing the Normalized Difference Vegetation Index (NDVI) and

30 meter by 30 meter resolution data from the CDL, each CLU was subset into five Yield

Stability Zones (YSZs) based on the year-to-year variability in NDVI. NDVI is used as a proxy

for yield. The NDVI-based proxy estimates were then compared and validated (at 30-meter

resolution) for corn and soy with observed crop yields at 2-meter resolution gathered from

combine harvester yield monitors for 508 fields across the region. The majority of these fields

were in a corn-soy rotation. NDVI-based yield stability maps were then generated for the entire

region for the 2010-2017 study period. Following Basso et al. (2019), there are five yield

stability zones identified and utilized in this study. Two zones representing areas of a field with

relatively stable, less variable NDVI and yields: High Stable zones (HS) that exhibit consistently

higher NDVI compared to the mean NDVI of the field and Low Stable zones (LS) where NDVI

was consistently lower than the mean NDVI of the field. Three unstable zones where some years

NDVI was above the field mean and in other years below the field mean were also identified and

further delineated based on relative position in the landscape: Unstable zones within depressions

are noted as Unstable-Depression (USD), on hilltops Unstable-Hill (USH), and sloping areas are

noted as Unstable-Other (USO).

SALUS Simulation Experiments

Using the individual fields spatial dataset defined by CLU and the stability zone

classifications from Basso et al.(Basso et al. 2019) as a framework, this study implements the

process-based Systems Approach to Land Use Sustainability (Basso and Ritchie 2015) (SALUS)

crop simulation model to examine the potential mechanisms responsible for nitrate leaching and

the potential impacts of using traditional uniform field management on areas of a field that

produce differing yields over time. Process-based crop models require daily weather data in the

form of daily values for maximum temperature, minimum temperature, downward solar

radiation, and precipitation. In studies examining single sites or a collection of sites, weather data is often acquired from in-situ observations at or nearby the field being simulated. For daily crop simulations across space and time, the use of gridded meteorological data is often necessary. The gridMET(Abatzoglou 2013) dataset was used as the meteorological dataset to drive the model. This dataset provides temporally and spatially complete daily data over the study period from 1989-2019 at 4-km spatial resolution. While the principal variables of temperature and precipitation have been validated by the producers and subsequent users of this dataset, solar radiation has received less attention and researchers have noted spatially variable biases (Kiefer et al. 2019). Solar radiation data were corrected for systemic biases on a grid-cell by grid-cell basis using the methodology outlined in Chapter 2. In addition to weather data, gridded soils data were also used. Baseline soils were acquired at 10-meter horizontal resolution for the 14-state region from the gridded version of the Soil Survey Geographic database gSSURGO (USDA NRCS, 2014). As Yield Stability Zone classifications used are unique to each field as designated by CLU, modifications to the soil depth and Soil Hospitality Factor (SHF) are relative to the majority soil type identified in each CLU. For each stability zone, modifications to soil depth, saturated hydraulic conductivity, and soil hospitality factor were necessary (Table 4.1).

Table 4.1 Table showing the general experimental setup and modifications required for SALUS to represent the Yield Stability Zones (YSZs).

YSZ	SHF Mods	Soil Depth Mods	SWCN Mods	Plant Pop. Mods	PRCP Mods
HS	+ at depths below 30 cm	Extend to 100 cm if needed	None	Corn: 10 plants/m^2 Soy: 37.7 plants/m^2	None
LS	- at depths below 10 cm	Restricted to 41 cm	Reduced at surface to simulate compaction	Corn: 4 plants/m^2 Soy: 37.7 plants/m^2	None
USD	+ at depths below 30cm	Extend to 100 cm if needed	None	Corn: 8 plants/m^2 Soy: 37.7 plants/m^2	Days > 10mm of precipitation, increase daily 20%. Wet years reduce yield by 30%
USH	- at depths below 10 cm	Restricted to 41 cm	Reduced at surface to simulate compaction	Corn: 4 plants/m^2 Soy: 37.7 plants/m^2	Days > 10mm of precipitation, decrease daily 20%.
USO	- at depths below 10 cm	Restricted to 41 cm	Reduced at surface to simulate compaction	Corn: 4 plants/m^2 Soy: 37.7 plants/m^2	Days > 10mm of precipitation, decrease daily 20%.

The agronomic assumptions and management set-up for the SALUS simulations were designed to represent historical management practices with respect to year on a county level assuming uniform management for YSZs. For computational efficiency, the entire region was simulated as all maize or all soy in alternation years. Maize and soy were chosen as the crops to be simulated as they represent the two largest crops in terms of acres planted in recent decades (Hunt et al. 2020), often in rotation with one another. Cultivar parameters for maize and

soybeans reflect historical cultivars used in the region and were validated against yield maps from Basso et al. 2019 over the years from 2008-2017. Planting, harvest dates, and fertilizer rates were determined from observed county level data from the National Agricultural Statistics Service (NASS). Crop planting dates were based on historical observations of the 50th percentile of reported dates of that county's planted acreage for that particular crop in each year. Planting dates were weighted to the density of actual production in each year in each state rather than the geographic center of each county. For each year, the centroids for corn production were calculated for each state and using county centroids weighted by total land area planted under corn in each county. Interpolation between state-year centroids was achieved by fitting a generalized additive model (GAM) to date of planting using the following model:

$$DP = s(lat, lon) + ti(lat * year) + ti(Lon * year)$$

Where DP is the date of planting at a given level (20, 50, and 80%), s(Lat,Lon), is the two-dimensional smoothing function with latitude and longitude as covariates; ti(Lat*year) and ti(Lon*year) are tensor functions of the random effect of year, latitude, and longitude. For soy years, the process for determining planting date was simpler (as the soybean module in SALUS is less complex than corn) and consisted of the 50% plant date for soybeans in each county respectively. Applied fertilizer rates were prescribed based on state rate values from Basso et al. (2019) (Table 4.2). Corn crops were fertilized in a split application, with 1/3 applied at planting and 2/3 as a side-dress application 45 days after planting. Fertilizer was applied as Urea-Ammonium Nitrate, broadcast and incorporated (100%) to a depth of 10.2 cm (4 in.) at planting and 5.1 cm (2 in.) at side-dress. The simulations were all run under rain-fed conditions across the region except for Nebraska and Kansas, where irrigation was simulated on all fields with sprinkler irrigation to maintain a range of 60-90% of maximum plant available water within a

managed soil layer in the top 30 cm of the profile. Harvest dates were prescribed based on the 50% of harvest county reported dates from NASS for soy and corn crops. Three separate tillage scenarios were simulated: 1) a no-tillage system, 2) a minimum tillage of 10 cm with a tandem disk 2 days prior to planting, and 3) a conventional tillage of 20 cm with a chisel plow 7 days after harvest in the fall. To stabilize nutrient and soil moisture dynamics, a spin-up simulation of 30 years from 1989-2018 was conducted and final conditions from the spin-up simulations were used to initialize the SALUS simulations at each pixel.
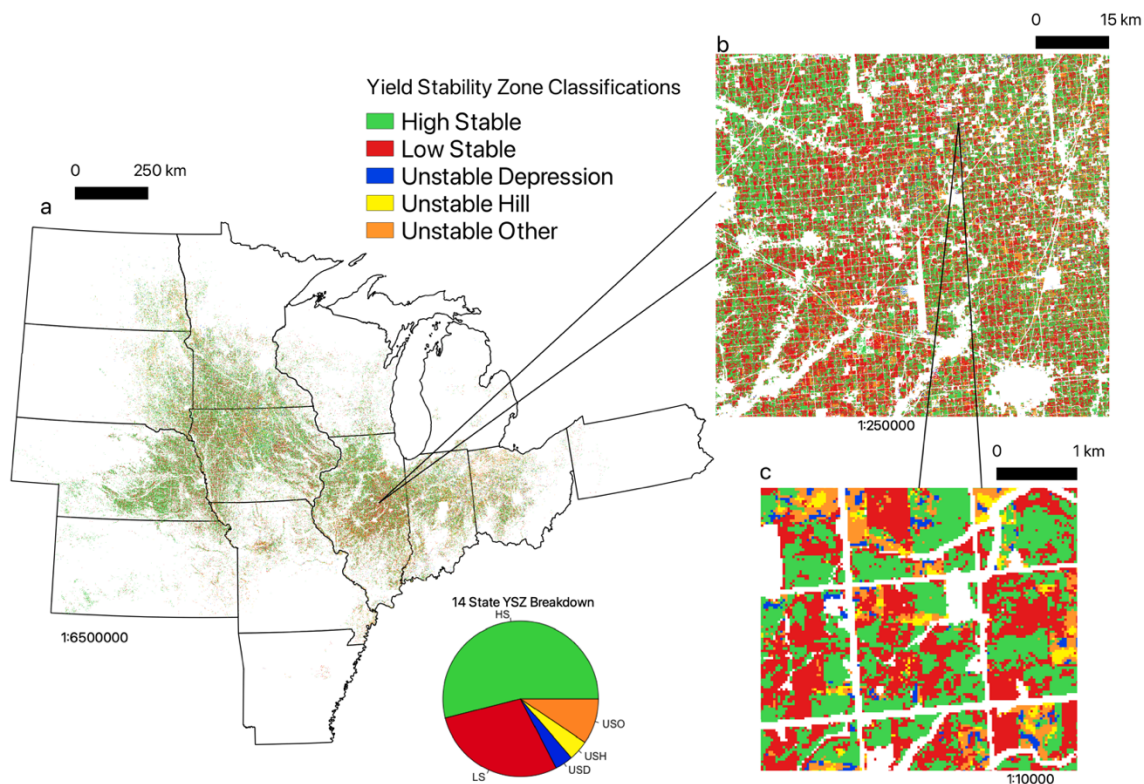
Figure 4.1 Maps showing the locations of simulations indicated by yield stability zone (YSZ) at three disparate scales; a) 1:6500000, b) 1:250000, c) 1:10000. The pie chart shows the percentage breakdown of the total number of 30-meter simulations across the region broken down by YSZ. HS zones are shown as green, LS as red, USD as blue, USH as yellow, and USO as orange.

Table 4.2 Reported range of fertilizer recommendations from Basso et al. (2019) and amount applied in SALUS simulations to corn crops. * The maximum rate was used in Wisconsin, as manure is more frequently used and manure was not captured in fertilizer rate recommendations. ** indicates Missouri rate used, *** indicates Iowa rate used, **** indicates Ohio rate used.

State	Range of Fertilizer N Rates (kg N/ha)	N Amount Applied in Simulations (kg N/ha)
IA	152-208	180
IL	179-229	204
IN	168-251	209.5
MI	143-198	170.5
MN	154-212	183
MO	189-260	224.5
ND	137-190	163.5
OH	155-234	194.5
SD	134-182	158
WI*	111-155	155
PA	155-234****	194.5
AR	189-260**	224.5
NE	152-208***	180

Model Output Analysis Methodology

SALUS and most process-based crop simulation models provide comprehensive output for a large number of variables on both a seasonal and daily basis. The total number of 30 meter simulations over the 16-state region broken down by YSZ are: HS (175,788,295), LS (93,222,151), USD (12,309,047), USH (13,689,920), and USO (30,735,605) for a total of 325,745,018 simulations conducted at a daily timestep from 1989-2019. Due to the very large number of unique simulations (Figure 4.1), most of the regional spatial analyses of this study were restricted to seasonal analyses (i.e. cumulative seasonal values at harvest). In this case, seasonal refers to the period from the day after the harvest of one crop until the harvest of the following crop. County-level subsets were chosen for 5 representative counties across the study domain to examine the daily characteristics of the output for county-level aggregate analysis and single site analysis (Figure 4.2). Model outputs available for analysis from this framework focused on the primary drivers of the nitrogen balance and are shown in Table 4.3. As an indirect

validation, aside from reported values in the literature and the closure of the nitrogen balance within the modeling framework, output of Total NLC from harvest 2012-harvest 2013 were visually and statistically compared with stream measurements of nitrate and total nitrogen collected during the spring/summer of 2013 Van Metre et al., 2016). One of the driving variables tested was the Weather Whiplash Index (WWI) (Loecke et al. 2017). This index looks at the variability in precipitation in 6-month periods centered on July. The WWI is defined as follows

$$WWI = \frac{\sum_{7_{i-1}}^{12_{i-1}} PRCP - \sum_{1_i}^{6_i} PRCP}{\sum(\sum_{7_{i-1}}^{12_{i-1}} PRCP + \sum_{1_i}^{6_i} PRCP)}$$

Where PRCP refers to monthly precipitation accumulation, i refers to the current calendar year, the subscripts on summation operator indicate starting month of period of accumulation, and the superscripts on the summation operators indicate the end month of the period of accumulation.
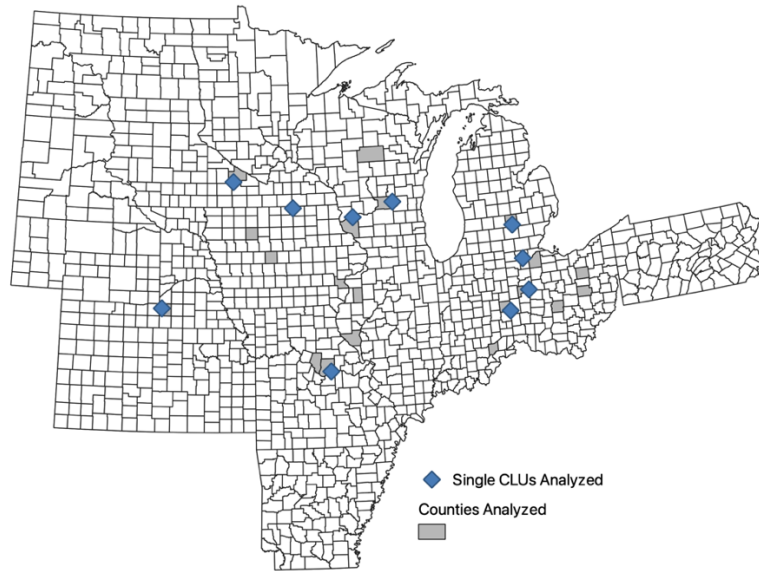
Figure 4.2 Counties within the study region where daily SALUS outputs were available and single sites where additional seasonal/daily analyses were conducted. *The single site located in Worth County, IA only had seasonal outputs available.

Table 4.3 SALUS output variables available for analysis.

Variable Name [units]	Description
GWAD [kg/ha]	Dry Grain Weight
NLC [kg/ha]	Nitrate Leached
Nitrate_Bl [kg/ha]	Nitrate below ground
N_Plant [kg/ha]	Plant nitrogen uptake
ROF (mm)	Cumulative runoff
DRN (mm)	Cumulative drainage
EOA (mm)	Evapotranspiration

Examining subsets of the data at differing spatial aggregations is helpful to elucidate the mechanics of nitrogen loss in this system. County-level analysis highlights the differing seasonal cycles present at different locations within our study region and identifies periods when nitrogen loss occurs relative to management. However, heterogeneities in soil type and intensity of production (i.e. number of simulations present in each county) obscure the individual contribution of each field and potential drivers of nitrogen loss are obscured. Analysis of the sub-

111

field variability (e.g. YSZs) at the individual field level allows us to examine the potential

mechanistic drivers responsible for the principal model outputs such as yield (GWAD) and

nitrate leaching (NLC). Individual field sites (CLU) were selected based on largest mean corn

yields in their respective counties. Once a field was selected, seasonal outputs were subjected to

correlation analysis (Pearson's r) to determine the general correlations between the various input

and output variables and links with yield and nitrogen loss (Table 4.4). Correlations were

performed both with the current season driving variables and the current seasons SALUS output

and with the previous seasons driving variable and the current season SALUS output (i.e. lag-1).

Correlation analyses were conducted for each stability zone available in the CLU considered.

Correlations were conducted both for each crop individually and in aggregate.

Table 4.4 Driving variables subject to correlation analysis with SALUS outputs.

Driving Variable	Description
PRCP_GS	Growing Season (May-Sept.) Precipitation Accumulation (mm)
WWI (Loecke et al. 2017)	Weather Whiplash Index (see Eqn. 1)
ET/PET	Ratio of October-September totals of SALUS simulated Evapotranspiration to Penman-Monteith grass reference Potential Evapotranspiration
MAXT GS	Annual Growing Season Mean High Temperature
PET WAT YR	Water year (October-September) Penman-Monteith grass reference Potential Evapotranspiration

Results

Background Hydroclimatic Environment/Trends

In recent years, a number of studies regarding precipitation trends and variability in the

Midwestern United States have generally concluded that precipitation totals and the amount of

water available on landscapes has increased (Baule et al. 2022; Kunkel et al. 2020a; Hayhoe et

al. 2018; Pryor et al. 2009) although there is substantial variability in these trends across the region. Trends in the meteorological data implemented in this study are in agreement with these findings. However examination of potential evapotranspiration (PETref) series in the region suggests that PETref has been increasing at a more rapid rate than precipitation, both annually (Figure 4.3 a & c) and within the growing season. Trends in PETref are statistically significant and positive in regions east of the Mississippi River, and for much of Missouri and western Kansas (Figure 4.3 d.). A mix of insignificant trends and significant positive trends is observed in Iowa, southern Nebraska, and Minnesota, and production areas of North Dakota and South Dakota. The larger trends in PETref compared to precipitation suggests that the increases in precipitation from 1989-2019 have been at least partially offset by increases in evaporative demand from 1989-2019 over portions of the region. Exceptions to this include swaths of Iowa, Indiana, Nebraska, and much of the Dakotas. Over much of the region positive trends in PET meet criteria for statistical significance, while negative slopes in PET are not statistically significant. Annual trends in precipitation were largely insignificant (Figure 4.3 b) from 1989-2019, save for portions of Nebraska, Wisconsin, Illinois, Indiana, and Michigan, where they exceeded +6 mm/yr in isolated areas (Figure 4.3 a).
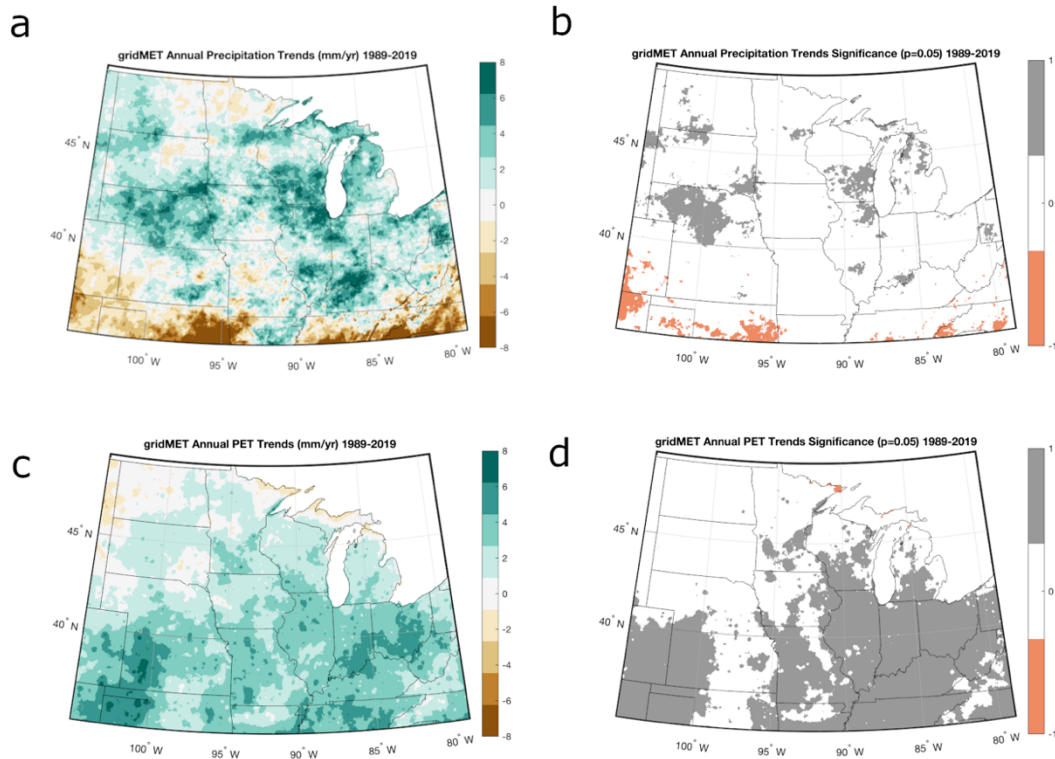
Figure 4.3 a) Sen's slope in annual precipitation accumulation (mm/yr) from 1989-2019, b) significance of trend (p<0.05,Mann-Kendall,two-tailed) in annual precipitation from 1989-2019 where a gray indicates significant positive trend, white is an insignificant (no-trend), and orange indicates a significant negative trend, c) Sen's slope in annual potential evapotranspiration (mm/yr) from 1989-2019, and d) significance of trend (p<0.05,Mann-Kendall,two-tailed) in annual PETref from 1989-2019 where gray indicates a significant positive trend, white is an insignificant trend (no-trend), and orange indicates a significant negative trend.

The increase in potential evapotranspiration has been driven primarily by increasing growing season temperatures from 1989-2019 (Figure 4.4 a & c). During the period from 1989-2019, maximum and minimum temperatures during the growing season increased at approximately the same rate over much of the region. Greater increases over time were evident in the eastern portions of the study region and were often associated with statistically significant positive trends (Figure 4.4 b and d). Exceptions to this pattern are largely restricted to the Dakotas, where minimum temperatures have increased, but more slowly than locations south and east, and maximum temperatures have been relatively flat or trending negative.
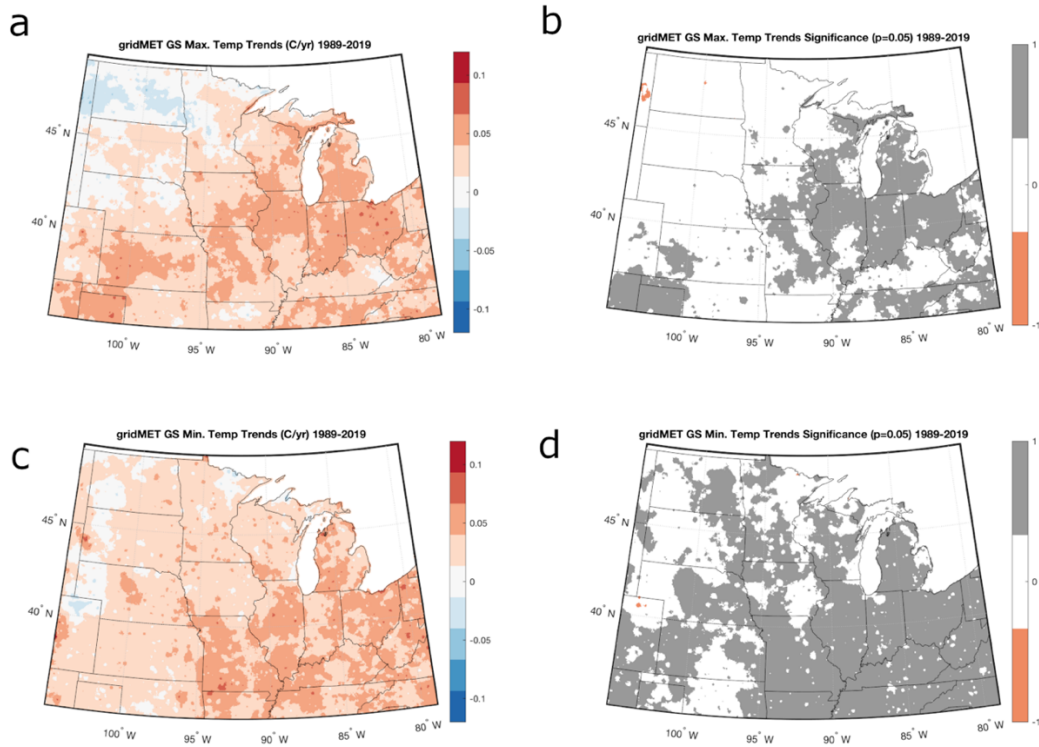
Figure 4.4 a) Trend in growing season maximum daily temperature (°C/yr) from 1989-2019, b) significance of trend (p<0.05,Mann-Kendall,two-tailed) in growing season maximum daily temperature from 1989-2019 where gray indicates a significant positive trend, white is an insignificant trend (no-trend), and orange indicates a significant negative trend, c) Trend in growing season minimum daily temperature (°C/yr) from 1989-2019, and d) significance of trend (p<0.05,Mann-Kendall,two-tailed) in growing season minimum daily temperature from 1989-2019 where gray indicates a significant positive trend, white is an insignificant trend (no-trend), and orange indicates a significant negative trend.

Figure 4.5 a) Map of simulated mean corn GWAD from 1989-2019 for all YSZs, b) histograms of mean corn GWAD from 1989-2019 by YSZ, and c) histograms of mean soy GWAD from 1989-2019 by YSZ.

For clarity of discussion, the results presented in this section will refer to simulations run with the no-tillage option. The impacts of tillage will be discussed in a later section as it largely results in a modulation/amplification type effect in regard to soil nitrogen. With respect to YSZ across the region and based on total grain produced (kg/ha) over the 1989-2019 period, HS zones were responsible for 76.9% of total production, LS were responsible for 16.6%, while USD, USH, and USO were responsible for 3.89%, 2.27%, and 0.32% respectively (Figure 4.5). In terms of mean yield, irrespective of intensity, HS zones exhibited the highest mean yields (10,184 kg/ha corn; 3,150 kg/ha soy), followed by USD (7,263 kg/ha corn; 2,898 kg/ha soy), and

LS (4,099 kg/ha corn; 2,408 kg/ha soy), USH (4,063 kg/ha corn; 2,180 kg/ha soy), and USO (4,057 kg/ha corn; 2,136 kg/ha) were substantially lower than HS and USD. In terms of yields by zone across the region (Figure 4.6), the highest corn yields by state were simulated in Nebraska (irrigated), followed by Iowa and Indiana. For soy, the highest yields were simulated in Minnesota, Pennsylvania, and Nebraska (irrigated). The three lowest yielding states in terms of corn were Missouri, North Dakota, and Arkansas. For soy, the three lowest yielding states were Wisconsin, Illinois, and Arkansas. Relatively greater interannual yield variability for both corn and soybean were observed across northern and western regions of the study area, particularly portions of Iowa, Minnesota, South Dakota, and North Dakota (not shown). In terms of yield trend (not shown) over the study period, there was a mix of negative and positive trends, with a concentration of negative yield trends in the southwestern portions of the study region and positive yield trends across the northern areas. However, the majority of yield trends were not statistically significant (Mann-Kendall, two-tailed, $p < 0.05$) and suggest a good match of crop parameters to the background climate. The three highest simulated yielding states in terms of combined corn and soy grain production were (in descending order) Iowa, Nebraska (irrigated), and Illinois while the three lowest total production states were North Dakota*, Pennsylvania, and Arkansas.
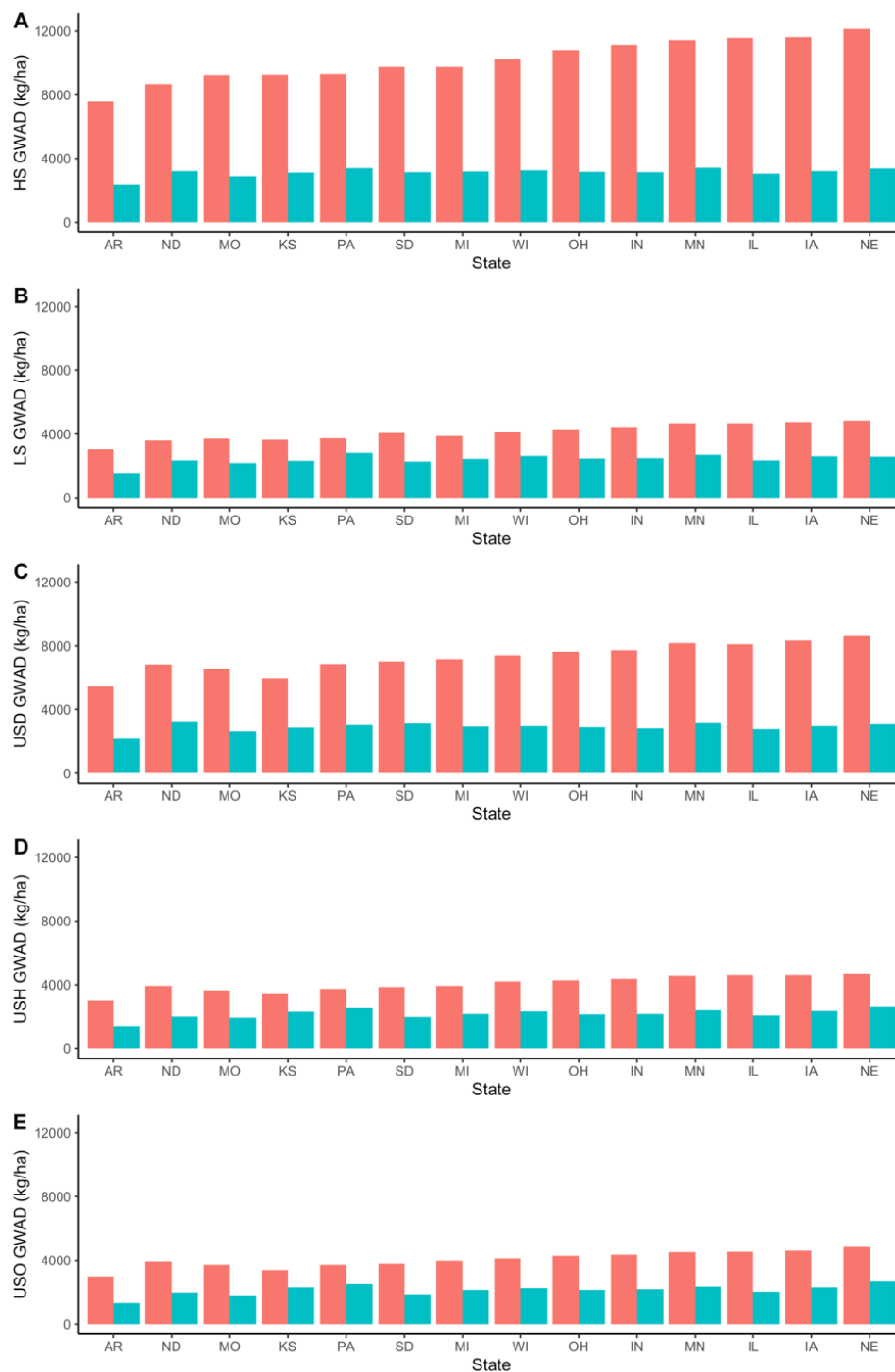
Figure 4.6 Bar plots of state level mean GWAD for corn (red) and soy (blue) from 1989-2019.
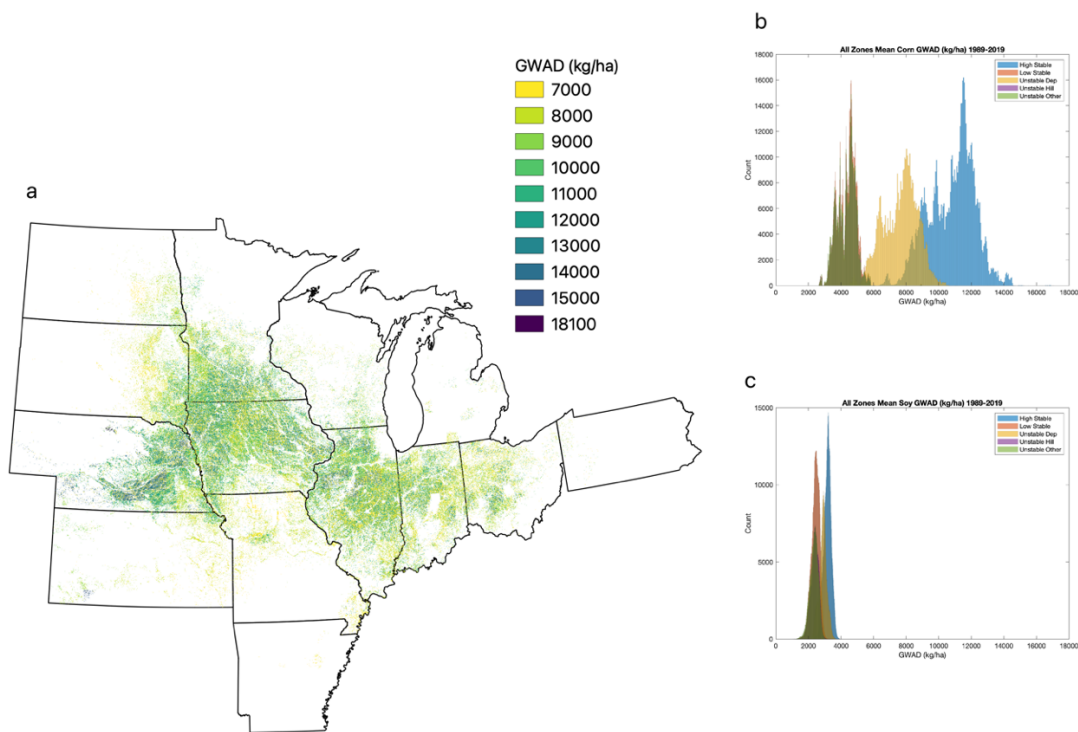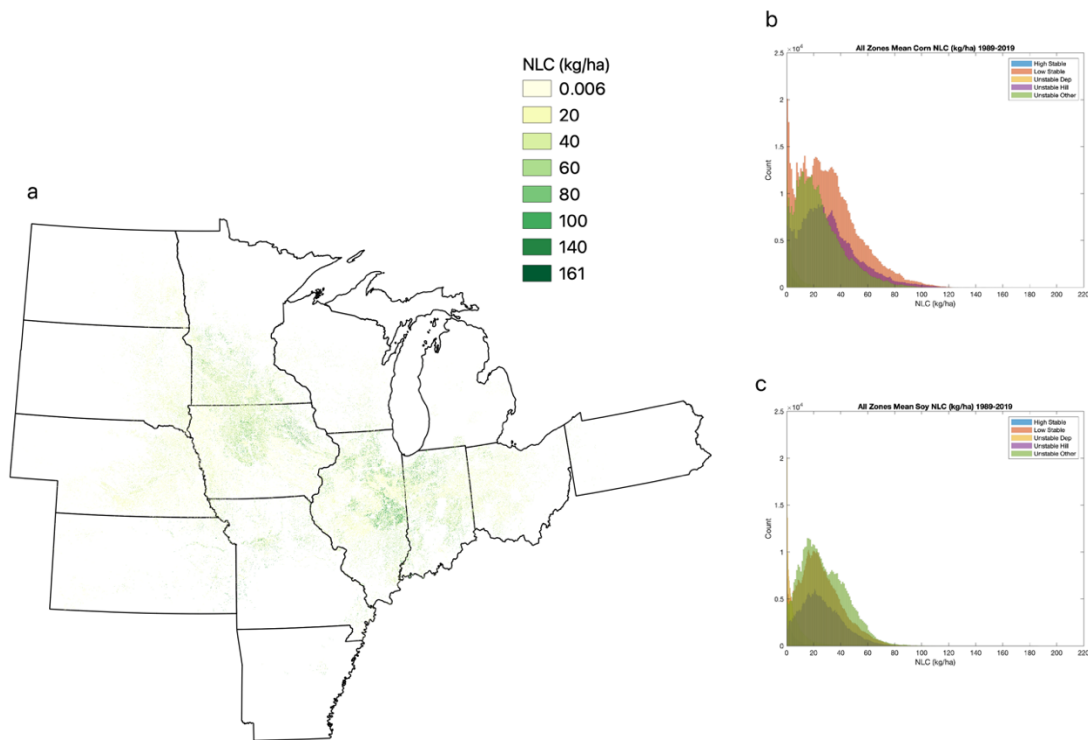
Figure 4.7 a) Map of simulated mean corn NLC from 1989-2019 for all YSZs, b) histograms of mean corn NLC from 1989-2019 by YSZ, and c) histograms of mean soy NLC from 1989-2019 by YSZ.

For Nitrate Leaching (NLC) with respect to YSZ, LS (26.96 kg/ha corn; 25.15 kg/ha soy), USH (26.96 kg/ha corn; 26.76 kg/ha soy), and USO (26.96 kg/ha corn; 30.18 kg/ha) were responsible for highest mean NLC values. HS (1.94 kg/ha corn; 8.31 kg/ha soy) and USD (4.58 kg/ha corn; 5.74 kg/ha soy) zones were responsible for less NLC on average (Figure 4.7). NLC exhibited more relative pixel to pixel spatial variability across the region than yield (Figure 4.7). The states with the highest simulated NLC mean values during corn production years were, in descending order Arkansas, Indiana, and Kansas (irrigated) while the states with the lowest mean annual NLC were Nebraska (irrigated), South Dakota, and North Dakota. Despite the drastic differences in yields between soy and corn crops in our simulations, leaching was more evenly

distributed between soy and corn years. NLC in corn years was somewhat higher in most states, likely due to the timing of fertilization during the growing season (Figure 4.8). This references to the dominant sources of NLC in the simulations coming from the LS, USH, and USO zones of the field. In contrast, for the HS and USD zones, leaching was generally higher during soy years than corn years (Figure 4.8 a & c). Since these zones are considered well-managed and in deeper soil profiles within our framework, this suggests the NLC differences between HS/USD and LS/USH/USO are associated with the uniform management scheme. Over the study period, NLC increased over time in northwestern and eastern sections of the region but tended to decrease in central sections (Figure 4.9 a). In the areas where trends were statistically significant, the vast majority were positive and concentrated in areas of Iowa, Minnesota, South Dakota, and North Dakota. Statistically significant negative trends in nitrate leaching are evident in east central Illinois. As with GWAD, the majority of the trends in NLC across the region were insignificant. The top three states in terms of total nitrate leached over the simulation from 1989-2019 were Illinois, Iowa, and Indiana. The lowest three states in terms of total nitrate leached were, in descending order, North Dakota*, Arkansas, and Pennsylvania. The order of the states for total amount of nitrate lost are largely reflective of the intensity of production and not necessarily the most vulnerable locations for nitrogen loss on a field-by-field basis (Figure 4.10). The relatively high contribution of the LS zones to the state total highlights these stability zones as the primary contributors of leached nitrate. LS zones on average are responsible for 65% of the total nitrate leached on a state-by-state basis.
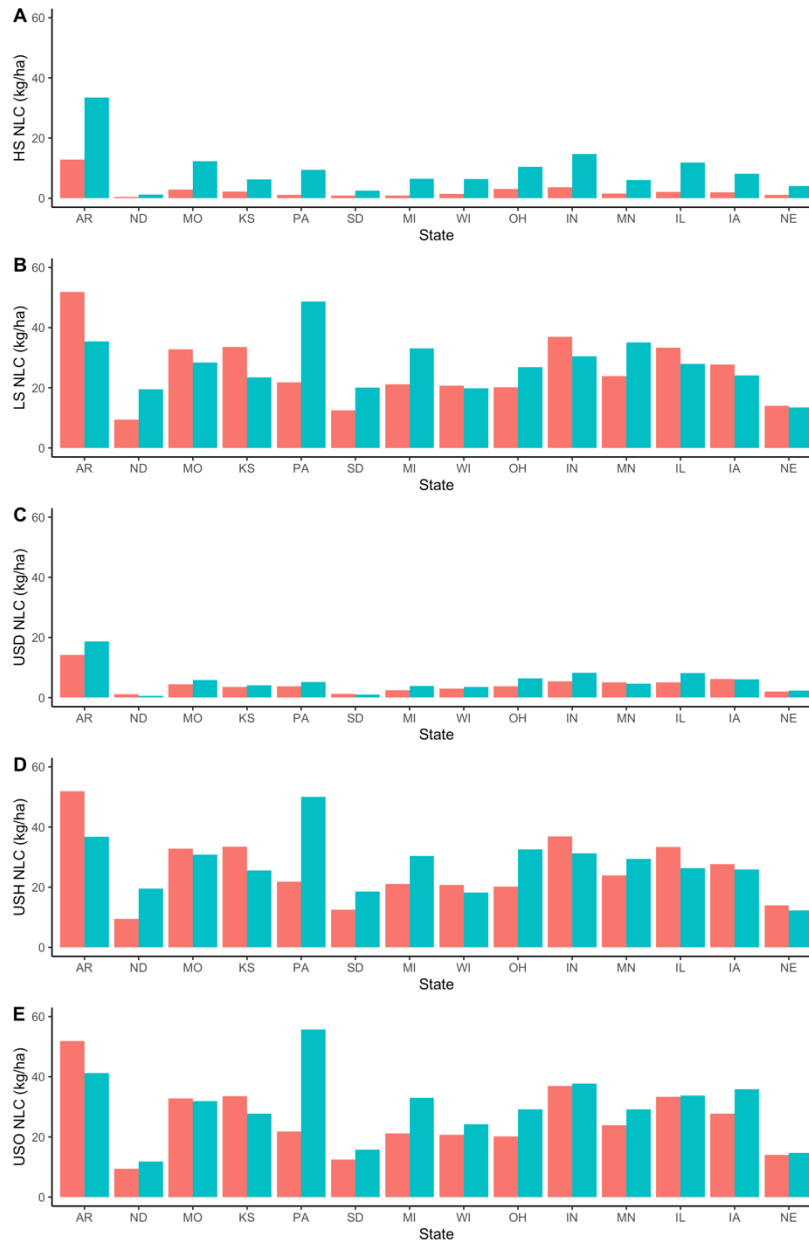
Figure 4.8 Bar plots of state level mean NLC for corn (red) and soy (blue) from 1989-2019. Order of states on the x-axis is in the same order as Figure 4.6.

121

a

b
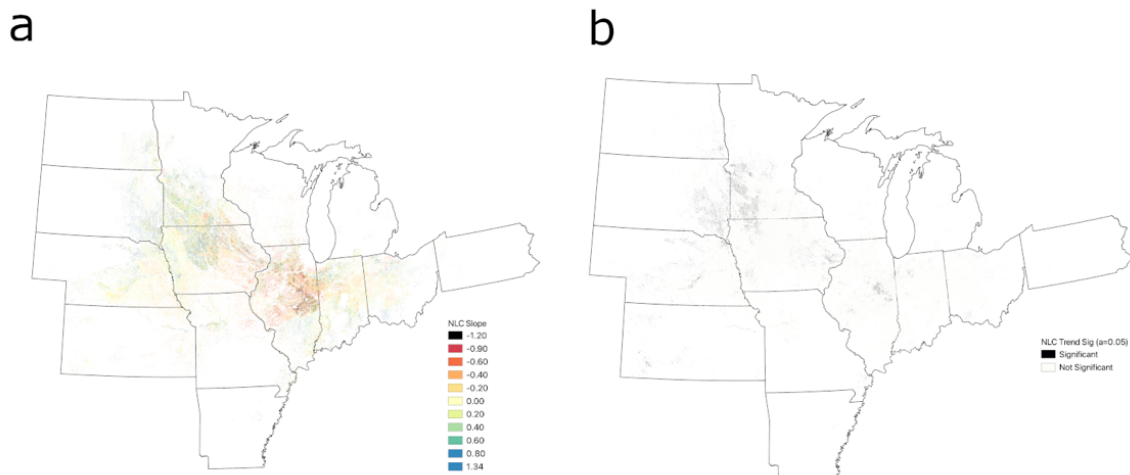
Figure 4.9 a). Sen's slope (kg ha^{-1}yr^{-1}) of seasonal Nitrate Leaching (NLC) over the period from 1989-2019 for all crops. b). Statistically significant trends (p <0.05, Mann-Kendall, two-tailed) in NLC.
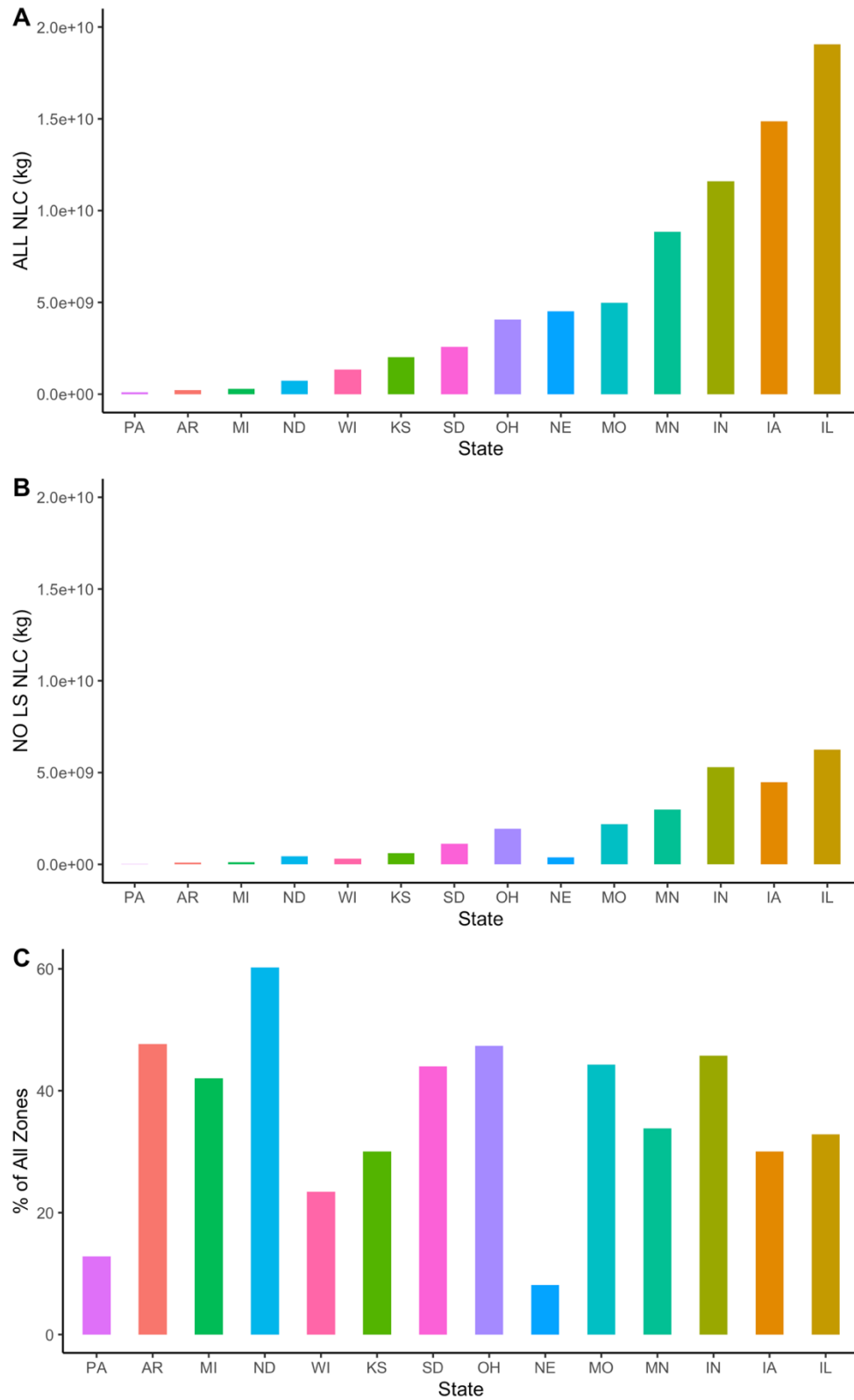
Figure 4.10 a) Bar plot of total simulated NLC (kg) from 1989-2019 with all YSZ considered. b). Bar plot of total simulated NLC (kg) from 1989-2019 with all zones except LS included. c). Percentage of NLC from all zones besides LS.

Nitrate below ground (Nitrate_Bl) at harvest is not a variable often measured across wide spatial scales for comparison. However, simulated model outputs of such variables can be useful for identifying the potential for subsequent nitrogen losses. On average across the region, the areas with the most nitrate remaining in the soil profile at the end of the corn production season were located in North Dakota, South Dakota, Minnesota, north central Iowa, and east central Illinois (Figure 4.11). The YSZs exhibit different distributions of mean Nitrate Bl depending on crop type. Under corn, USO zones have the highest mean Nitrate_Bl, followed by USH, LS, USD, and HS. Under soy, which is not unfertilized with nitrogen, USD has the highest mean Nitrate_Bl, followed by HS, USO, USH, and LS. Since Nitrate Bl is related to the depth of the profile, the impact of fertilization and management on these zones becomes more apparent in that deeper soil profiles contain more residual nitrate than those located on shallower soils, which lose nitrate more quickly through leaching (Zotarelli et al., 2007).
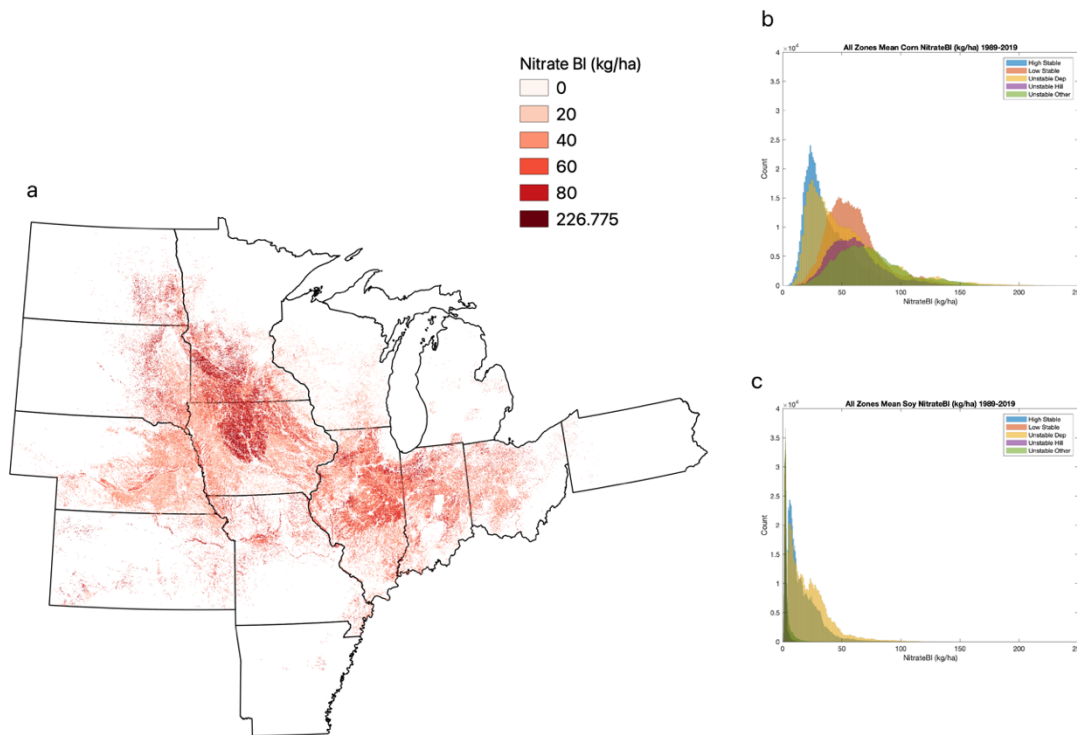
Figure 4.11 a) Map of simulated mean corn Nitrate_Bl from 1989-2019 for all YSZs, b) histograms of mean corn Nitrate_Bl from 1989-2019 by YSZ, and c) histograms of mean soy Nitrate_Bl from 1989-2019 by YSZ.

Combining NLC and scaling it per unit of yield illustrates the differing potential economic impacts among the zones as related to leaching, particularly for corn (Table 4.5). In terms of NLC per GWAD, USH, LS, and USO exhibit higher values of NLC per GWAD than USD or HS zones. Based on state level simulation means, it suggests that Arkansas, Pennsylvania, and Michigan produce more NLC per GWAD (e.g., Arkansas LS NLC per GWAD was 0.021). The three states with lowest NLC per GWAD were South Dakota, North Dakota, and Nebraska (e.g., South Dakota LS NLC per GWAD was 0.005) Simulated N_Plant values were largely reflective of yield and show similar spatial patterns to GWAD by YSZ.

DRN, ROF, and EOA variables differed little by YSZ and are largely reflective of the

background climate of the region on a seasonal basis.
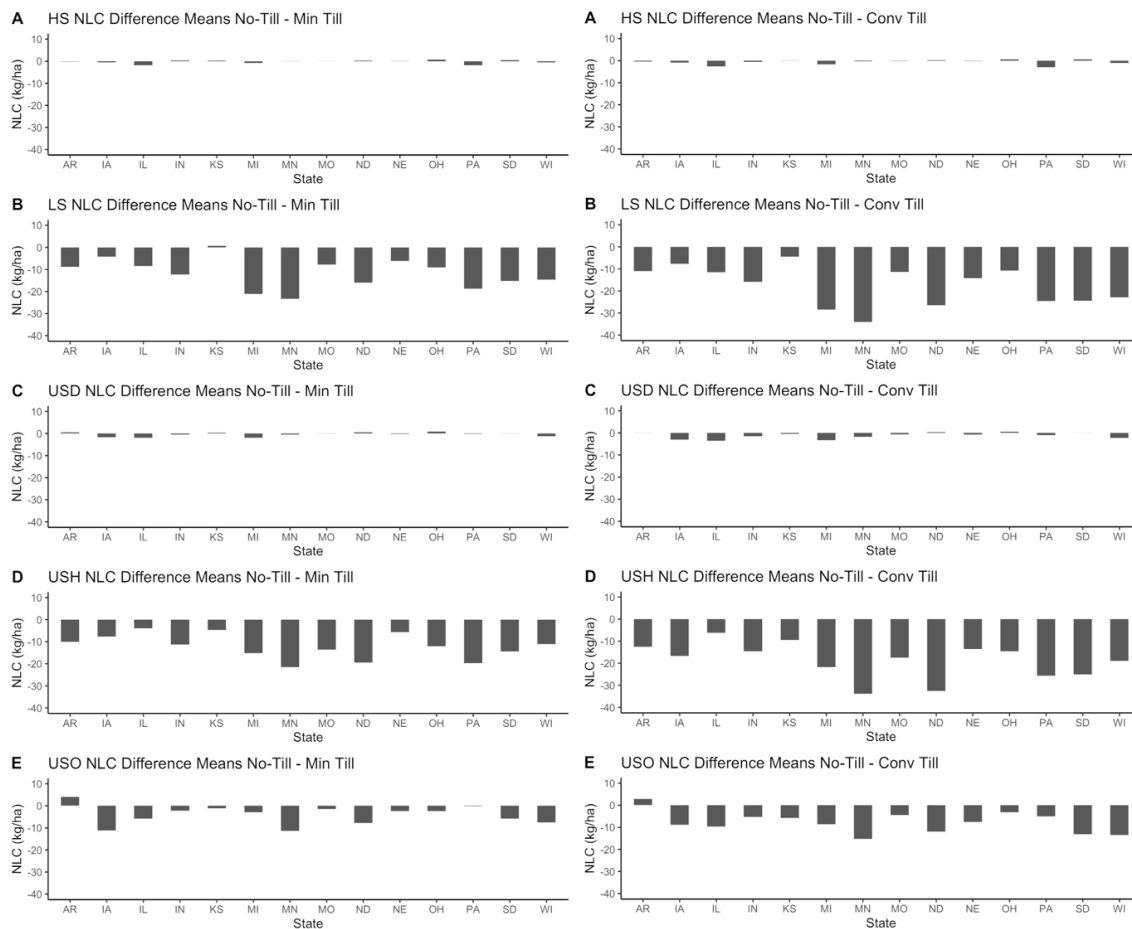
Figure 4.12 The impact of minimum spring tillage (10 cm) and conventional fall tillage (20 cm) on nitrate leaching expressed as the difference from the no-tillage treatment by yield stability zone: a) HS zones, b) LS zones, c) USD zones, d) USH zones, and e) USO zones.

Table 4.5 NLC per GWAD state level means for corn years by state. Unitless.

State	HS	LS	USD	USH	USO
Arkansas	1.158E-03	0.021	1.520E-03	0.022	0.017
Pennsylvania	2.167E-05	0.011	2.233E-04	0.011	0.007
Michigan	2.238E-05	0.011	2.243E-04	0.009	0.007
Missouri	9.854E-05	0.010	3.723E-04	0.012	0.009
Indiana	1.153E-04	0.010	4.607E-04	0.010	0.008
Kansas	9.604E-05	0.009	3.124E-04	0.011	0.010
Illinois	5.608E-05	0.008	3.040E-04	0.008	0.007
Minnesota	7.313E-05	0.007	5.394E-04	0.008	0.006
Ohio	3.123E-05	0.006	1.719E-04	0.007	0.005
Wisconsin	3.415E-05	0.006	2.208E-04	0.007	0.006
Iowa	7.245E-05	0.005	5.624E-04	0.006	0.006
South Dakota	1.154E-05	0.005	1.129E-04	0.006	0.004
North Dakota	1.397E-05	0.005	2.776E-03	0.096	0.015
Nebraska	2.112E-05	0.003	1.055E-04	0.003	0.003

Impact of Tillage

Under the three separate tillage scenarios, the impacts of tillage on the seasonal results are striking. SALUS simulates the breakdown of soil organic matter (SOM) and increased saturated hydraulic conductivity near the surface of the soil profile (Basso and Ritchie 2015). Over time, the impacts of tillage result in slightly lower yields, with generally larger yield reductions with increased intensity of tillage. The impacts of tillage on nitrogen-related variables are more complex. In general, there was more Nitrate_Bl at the end of the growing season under the tilled treatments. N_Plant is greater under the two tilled options than under no till and greatest under conventional tillage, but was not great enough to offset the increased Nitrate_Bl under the tilled treatments. As a result of relatively unstable yields, somewhat increased N_Plant and Nitrate_Bl, NLC also increased under both tillage scenarios with the largest changes under

conventional tillage (Figure 4.12). States with the largest increases in simulated leaching with increasing tillage were North Dakota, Minnesota, and Iowa.

Comparison with Observed Stream Concentration Data

Observed stream nitrate totals were plotted on top of 8-digit Hydrologic Unit Code (HUC) maps of total simulated nitrate leached within each unit from harvest 2012 through harvest 2013 (Figure 4.13 a). Despite the simplifications in our model regarding crop rotation, the inability to simulate all fields and sources of environmental nitrogen, and other unaccounted-for processes there is a strong visual between both variables with agricultural intensity. There is also strong empirical evidence for a linkage between WWI and nitrate loss in the western Corn Belt. Loecke et al (2017) examined nitrate levels in streams following the 2012 Midwestern Drought, particularly in locations that experienced a large negative WWI over that period (from abnormally dry to abnormally wet). In our simulations, we see a similar response in the western HUC-8 zones in terms of nitrate leached and how it corresponds to measured stream nitrogen levels (Figure 4.13 b). In eastern sections of the study region (OH, eastern IN, and PA), leaching in the 2013 growing season appeared to be more related to the total movement of water, as these regions were not as affected by the 2012 drought. Overall correlation between simulated HUC-8 total nitrate and observed stream nitrate is moderate across the region (Pearson's r=0.34, not shown), though the complexity of the actual agricultural landscape versus the simulated landscape likely accounts for at least some of the weakness in the statistics.
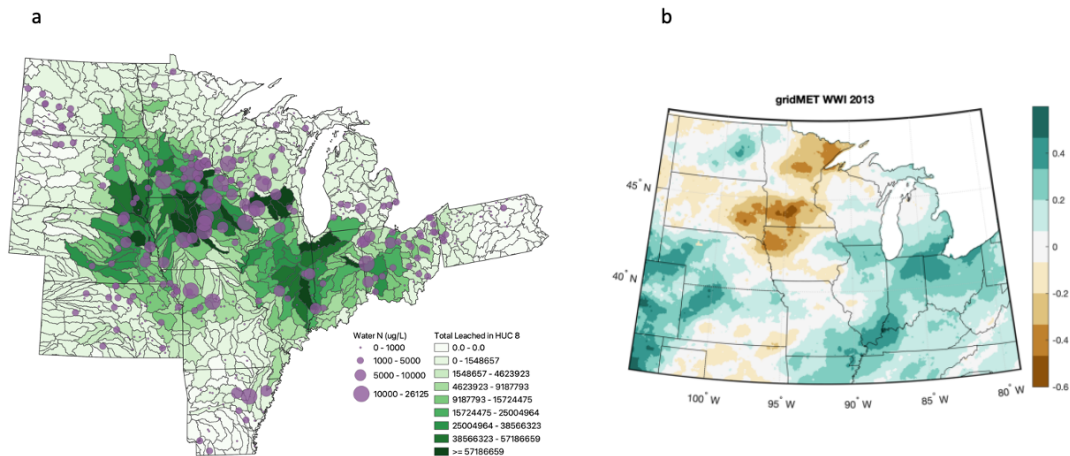
Figure 4.13 a) Total NLC from harvest 2012 through harvest 2013 with each HUC-8 watersheds (green fill), Scaled purple dots indicate stream nitrate concentration. b) WWI for the study region for 2013.

County Level Daily Results

At the county level, the relationships between the YSZs behave similarly to the broader region in terms of relative yield and leaching differences (e.g. HS = low NLC, high GWAD; LS= high NLC, low GWAD) when the seasonal results are examined. However, daily plots of total NLC suggest that the timing of when leaching occurs has a geographic component associated with the character of precipitation within that given county (Figure 4.14). In western and northern locations in the region, the majority of the nitrate was leached between April and October, even in LS, USH, and USO zones. In the two eastern locations displayed, NLC is more consistent throughout the year, suggesting a linkage with the annual cycle of precipitation. All zones and regions show a decrease in NLC on days of the year that correspond to peak crop N-uptake (generally following side-dress nitrogen application 45 days after planting through days 70-80 after planting), followed by an increase after this period. The large individual peaks within the histograms are reflective of individual heavy rain events, whereas the general shape displayed by the bar plots reveals the seasonal nature of leaching punctuated by large weather

events and appears to be more related to the annual cycle of total precipitation and the soil characteristics at a particular location. For example, Auglaize County, Ohio receives more precipitation on average (mean: 1009 mm/yr, stdev: 161 mm), while Boone County, Iowa receives less (mean: 908 mm/yr. stdev: 192 mm) precipitation but precipitation is more variable year to year and more concentrated (Figure 4.14 a) during the growing season (July mean: 137mm & December mean: 32 mm) than Auglaize County, Ohio (July mean: 109 mm & December mean: 66 mm) (Figure 12e). The more consistent supply of rainfall throughout the year suggests a longer potential leaching season and more consistent leaching throughout the year and year to year than the more variable precipitation regime in Boone County, IA. The behavior of the individual zones at these locations, based on total NLC is: HS and USD zones exhibited few "spikes" on a given day of the year with a smoother, less variable seasonal cycle. In contrast, weather-related "spikes" were abundant across all locations in the LS, USH, and USO zones. At all locations displayed, the largest individual NLC events occurred during corn years following fertilization, again displaying the sensitivity of the LS, USH, and USO zones to weather. In soy years (not shown), NLC between the period of crop establishment and harvest is minimal when compared to corn years where fertilizer is applied. Attempts to parse out response by soil textural class, yielded little improvement in results, illustrating the complexity of interactions between physical and managed aspects of the agricultural system even at the county scale.
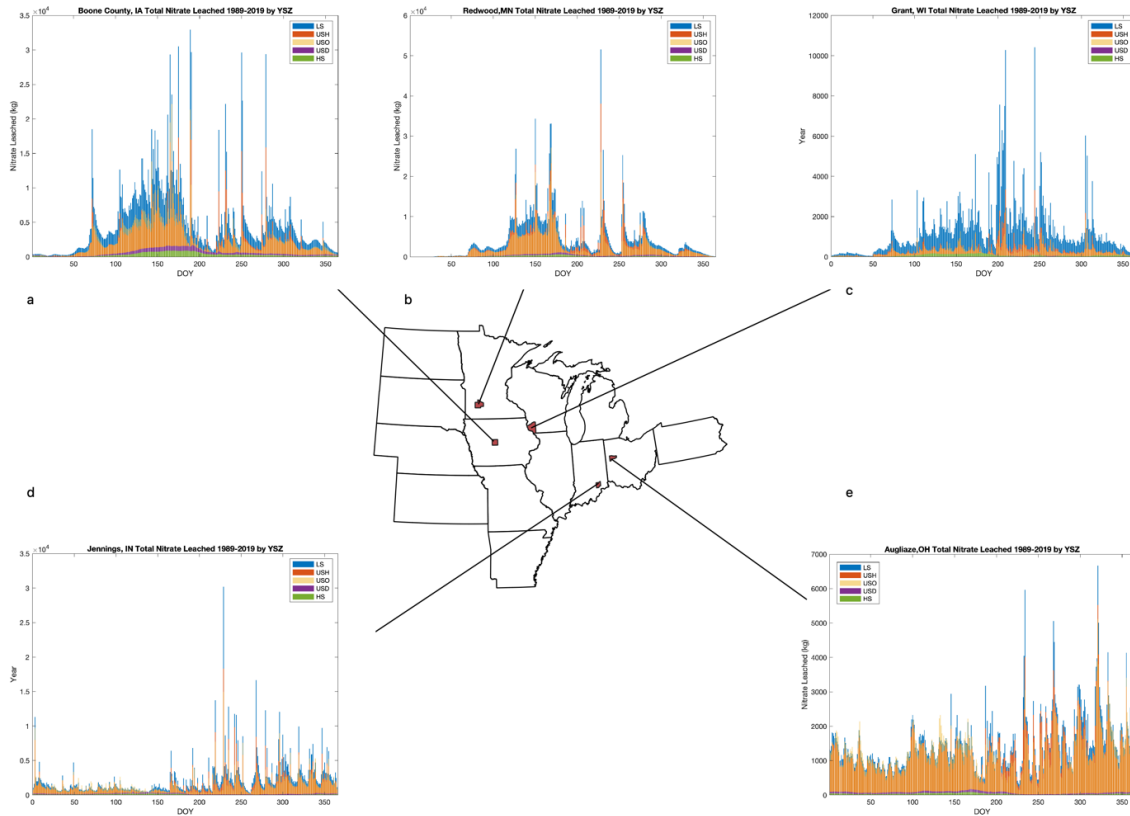
Figure 4.14 Total simulated NLC (kg) by numerical Day of Year (DOY) from 1989-2019 by YSZ for five counties: a) Boone County, Iowa; b) Redwood County, Minnesota; c) Grant County, Wisconsin; d) Jennings County, Indiana; and e) Auglaize County, Ohio. Blue bars represent LS zones, red bars USH, yellow bars USH, purple USD, and green HS.

Single Site (CLU) Results

Within the context of the effect of weather and climate directly on YSZs, it was helpful to conduct analyses of singular fields as the soil classification within our framework was assigned to the majority value for that CLU (e.g. soils are the same). This has the effect of eliminating variability based on soil type, aside from modifications due to the YSZ parameterization. The 3 highest correlated driving variable/crop/lag combinations at single CLU locations are shown in Table 4.6.

131

Table 4.6 Three leading seasonal correlation coefficients (r) between SALUS model outputs and driving hydroclimatic variables at select single field locations. Current refers to driving variables in the current season, Lag1 refers to lagging the driving variable by one season when correlated with SALUS output. Below the count and state name for each location is a cell displaying YSZs present.

County, ST	NLC Combination	r	GWAD Combination	r	NitrateBl Combination	r	EOA Combination	r	DRN Combination	r
Clay, NE	LS WWI Soy Lag1	-0.61	HS ET/PET All Current	0.78	LS ET/PET All Current	0.69	HS PET WAT YR Soy Current	0.97	LS ET/PET All Current	0.69
HS, LS	LS WWI CORN Current	0.54	LS_MAXT_GS Soy Current	-0.74	HS MAXT_GS Corn Current	0.56	LS PET WAT YR Soy Current	0.96	HS MAXT_GS Corn Current	0.56
	LS ET/PET Corn Lag1	-0.49	LS ET/PET All Current	0.70	LS WWI CORN Current	-0.54	LS MAXT_GS Soy Current	0.89	LS WWI Corn Current	-0.54
Worth, IA	USD PRCP GS Corn Current	0.77	HS MAXT_GS Soy Current	-0.84	USO PET WAT YR Soy Current	0.95	HS PET WAT YR Corn Current	0.94	USO PRCP GS Corn Current	0.93
HS, LS, USD, USO	USO PRCP GS Corn Current	0.75	HS_ET/PET All Current	0.84	USO MAXT_GS Soy Current	0.92	USO PET WAT YR Soy Current	0.93	LS PRCP GS Corn Current	0.92
	USO PRCP GS Soy Lag1	-0.70	HS PET WAT YR Soy Current	-0.81	LS PET WAT YR Soy Current	0.81	LS PET WAT YR Soy Current	0.93	HS PRCP GS Corn Current	0.91
Grant, WI	LS PET WAT YR Soy Current	0.70	USO_ET/PET All Current	0.81	HS MAXT_GS Corn Current	0.88	USD PET WAT YR Soy Current	0.95	USD PRCP GS Corn Current	0.95
HS, LS, USD	LS ET/PET Soy Lag1	0.63	HS ET/PET All Lag1	0.85	USD MAXT_GS Corn Current	0.85	HS PET WAT YR Soy Current	0.94	HS PRCP GS Corn Current	0.95
	LS_MAXT_GS Soy Lag1	0.61	LS ET/PET All Current	0.78	HS PET WAT YR Corn Current	0.82	LS MAXT_GS Soy Current	0.94	LS PRCP GS Corn Current	0.94
Henry, OH	USH ET/PET Soy Lag1	0.64	HS PET WAT YR Soy Current	-0.67	HS ET/PET Corn Current	0.75	USH MAXT_GS Soy Current	0.93	LS ET/PET Soy Current	-0.73
HS, LS, USH, USO	LS ET/PET Soy Lag1	0.53	USH MAXT_GS Soy Current	-0.67	USH ET/PET Corn Current	0.72	LS_MAXT_GS Current Soy	0.92	USH ET/PET Soy Current	-0.73
	USH MAXT_GS Soy Lag1	0.52	LS MAXT_GS Soy Current	-0.66	HS ET/PET All Current	0.72	USO MAXT_GS Soy Current	0.92	USO ET/PET Soy Current	-0.73

Across the selected sites, the strength of the correlations varied substantially by location, with the highest correlation values between the modeled output of EOA and DRN and driving variables, which is not surprising as these variables respond directly to weather within the model framework. The low correlation values present at the Clay County, NE CLU illustrate the decoupling effect of irrigation on these relationships. This holds across correlations with all modeled outputs with respect to irrigation, indicating lower correlations between hydroclimatic driving variables and model outputs. For EOA and DRN, YSZs display less sensitivity to the driving hydroclimatic variable than the other variables more directly related to management (e.g. NLC, GWAD, and NitrateBl). Besides EOA, the highest correlated driving variables across the region were mostly PET WAT YR, and MAXT GS mixed among the various YSZs and crops. Precipitation related variables played a more direct role for DRN with the highest correlations being ET/PET, MAXT GS, WWI, and PRCP GS. For both variables, all of the top 3 leading correlation coefficients were in the current season for both crop types, suggesting little carryover effect from the previous season on EOA or DRN.

For GWAD and NitrateBl, most of the highest correlations were for the current season, suggesting again that there is little carryover from the previous crop that affects this outcome. Across all locations shown, HS zones were the most correlated with GWAD For NitrateBl, and correlations were strongest in zones other than HS. The LS, USH, and USO zones exhibited stronger correlations with driving variables that HS or USD zones. GWAD appears to be most correlated with ET/PET, MAXT GS, and PET WAT YR variables, illustrating the importance of moisture availability and supportive temperatures for higher crop yields. There is a notable drop-off of the strength of correlations as one moves east in the region (from Worth County, IA to Grant County, WI, to Henry County, OH), suggesting that western locations, with less annual

133

and growing season precipitation and higher interannual variability in precipitation are more closely tied to these modeled outputs. For NitrateBl, the situation is similar but notably different. The LS, USO, and USH zones were predominant rather than HS YSZs in terms of strength of correlations with NitrateBl. With ET/PET, MAXT GS, PET WAT YR, and WWI being predominant among leading correlation coefficients and all were for corn seasons suggesting a relationship to fertilization. NitrateBl displayed similar constraints to GWAD though variables suggesting precipitation variability plays a more direct role, with WWI showing up as a leading coefficient at some locations (Table 4.6).

The results for NLC are more nuanced than those for the other modeled variables presented. LS, USO, USH, and USD were the predominant YSZs present in the leading correlation coefficients. Soy and corn crops display dramatically different behavior than the other variables presented in that the previous year (e,g. Lag 1) plays a dominant role in the magnitude of nitrate leaching, particularly in true soy seasons. This relationship is dependent on the management of the crops and crop rotations present. As soy crops were unfertilized, NLC is dependent on the amount of residual NitrateBl from the previous corn crop, which is in turn related to the GWAD and individual YSZ behavior. The absence of the HS zone in leading correlation coefficients, illustrates the minimal contribution of HS zones (and to a lesser extent, USD zones) to NLC regardless of location or crop type. At the western locations, WWI and Growing Season Precipitation (PRCP_GS) were the leading driving variables related to NLC. Under rainfed conditions (Worth County, IA), PRCP_GS in the current season for corn or previous season (Lag1) for soy is the leading driving variable for NLC. Under irrigation (Clay County, NE), PRCP_GS isn't as limiting since it's supplemented by irrigation. At this location, WWI (which is relevant outside of the growing season, when no irrigation is applied) appears to

be the primary driver of NLC. At the two more eastern locations, the PET WAT YR, ET/PET, and MAXT_GS variables are the leading correlated driving variables suggesting that air temperature is the primary determinant in NLC for the current season. The strength of correlations between driving variables and model output also decreases as one moves east across the region or when supplemental irrigation is assumed (i.e. Clay County, NE), suggesting that precipitation variability at the eastern locations is less important for NLC and more dependent on the supportive thermal environment for the previous and current crop (Figure 4.15).
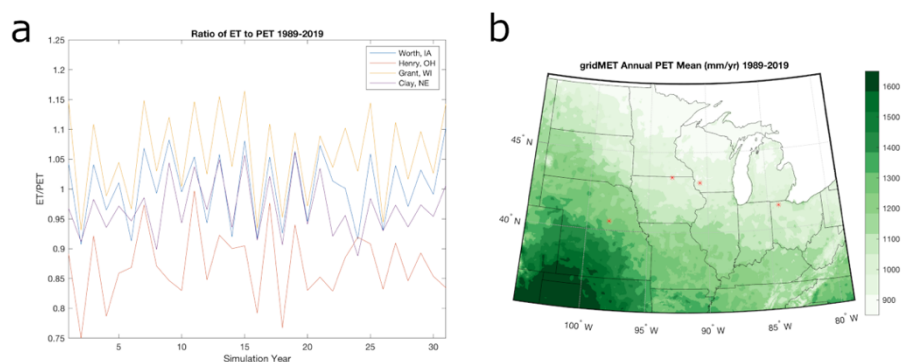
Figure 4.15 a) Ratio of crop ET to PET for 4 single CLUs across the study region: Worth County, IA (blue), Henry County, OH (orange), Grant County, WI (yellow), and Clay County, NE (purple). b) Contour plot of mean annual PET from 1989-2019, red * on map indicate the location of the individual CLUs plotted in a.

An example of this behavior is presented for Worth County, IA from 1989-2019 (Figure 4.16). The maximum NLC year for this location occurred in 1990, which was planted as soy. The current growing season precipitation (Figure 4.16 a) had the leading correlation with NLC for this location, while WWI was also strongly correlated at the site. Fertilizer was applied in 1989 (180 kg/ha) followed by the second driest growing season in the simulation timeline (Figure 4.16

a). which resulted in high NitrateBl at the end of the 1989 growing season. Given a relatively normal amount of GS_PRCP in 1990 (Figure 4.16 c), NLC was the largest seasonal total for this location (HS: 7.2 kg/ha, LS: 144 kg/ha, USD: 13.6 kg/ha, and USO: 167.9 kg/ha) over the simulation period. Subsequent droughts at this location, such as the 2012 Midwestern Drought, provide an example of the modulating effects of crop type and management. 2012 was planted as soy with no fertilizer applied, and while the 2013 season that followed was wet, the 2013 corn crop wasn't fertilized until late-spring. This resulted in less overall NLC in the 2013 season, due to crop rotation and management, despite the drought being more severe at this location in 2012 than in 1989. Scatter plots showing the relationship between PRCP_GS and NLC in the current growing season (Figure 4.16 b) and DRN and NLC in the current season (Figure 4.16 d) highlight the different strength of the correlations by model output variable, in this case NLC and DRN.
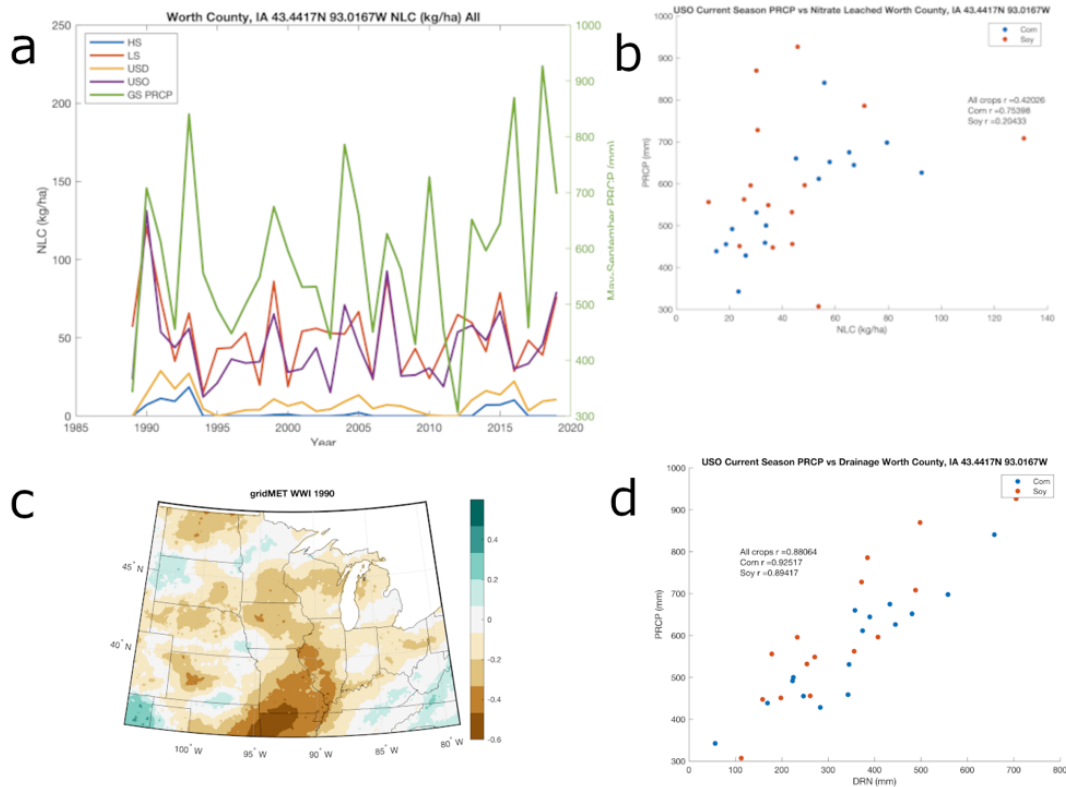
Figure 4.16 a) Time series of simulated NLC (left-axis) and observed PRCP_GS (green, right-axis) by YSZ for single CLU in Worth County, IA from 1989-2019, b) scatter plot for USO YSZ for current season PRCP_GS and NLC (orange dots represent soy years and blue dots represent corn years), c) Observed WWI index for 1990, d) scatter plot of USO YSZ for current season PRCP_GS and DRN (orange dots represent soy years and blue dots represent corn years).

Discussion/Conclusion

On the regional level, HS YSZs represent the majority (54%) of pixels simulated across the region. These zones are characterized by above field-average grain production, high plant nitrogen uptake, and typically little nitrate left in the soil profile at harvest. As a result of this good match between crop requirements and needs, there is a relatively low amount of leaching from these zones in general and especially relative to the other YSZs. This is not to say these zones are not susceptible to leaching and do not contribute to excess nitrate present on the landscape. But our modeling framework suggests that their contribution is small relative to other YSZs at the same time and place. USD YSZs are low spots in the landscape and are a relatively

137

small portion of the pixels simulated (4%). These zones are characterized by deep soils and additional water supply during wet years. Due to the empirical yield penalty imposed for these zones during wet years (Table 4.1), these zones are characterized by more variable yields and on average are lower than HS but higher than other zones. They are responsible for more leaching than HS zones, however their contribution is also small relative to other YSZ and due to the small proportion of USDs present on the landscape do not contribute substantially to the regional excess of nitrogen in the system. For individual producers, the relatively unstable nature of these zones suggest opportunities for precision management (e.g. Paiao et al. 2020; Jin et al. 2019; Basso et al. 2011) for producers managing fields with a large presence of USD YSZs.

The LS, USH, and USO YSZs in aggregate represent a substantial proportion of pixels simulated (42%) across the study region. These zones are characterized by lower yields than HS or USD YSZs, though the USH and USO YSZs display more variability than LS due to generally higher yields and plant populations, fewer shallower/compacted soils, greater plant nitrogen uptake, and less nitrate remaining in soil at harvest. This results in much less predictable and greater nitrate leaching than HS and USD YSZs (96% vs. 4%) under uniform management for the entire field. The large mismatch in nitrogen use in the LS, USH, and USO YSZs presents both a challenge and opportunity for producers. Our modeling results suggest that the USH and USO YSZs are prime targets for precision management in terms of nutrients applied due to their higher correlations with hydroclimatic drivers. The LS YSZs produce poorly in most years, regardless of the weather conditions in a given growing season. Our results suggests these zones may be better utilized both economically and environmentally by other crops or land uses rather than grain production (e.g. Basso et al. 2019). Though the nature and types of these adaptions/production changes is beyond the scope of this study, recent research suggests that

producer's perceptions of profitability and environmental benefits are key to the adoption of precision agricultural practices, and more research is needed particularly on the perceptions of producers regarding environmental benefits in the Midwest (Kolady et al. 2021).

Our results suggest that tillage across the region has a negligible to negative effect on mean aggregate yields in most states, resulting in greater plant nitrogen uptake, higher nitrate below ground, and higher nitrate leaching overall. This generally increases with the invasiveness of the tillage and reflects the increased breakdown of soil organic matter and increased hydraulic conductivity in the surface layers of the soil profile (Maharjan et al. 2018). The greater plant nitrogen uptake is consistent with field studies but is not enough to offset the increased supply of nitrate sources from fertilizer and SOM. Nitrate leaching has also been shown in field studies in the Midwest to increase under tillage with increased rain intensification (Hess et al. 2020), suggesting that our model is capable of capturing these tillage dynamics. This also highlights the fact that the results presented for the no-tillage system in our case are, at best, a conservative estimate of the dynamics of the nitrogen balance on a regional scale. Correlation analyses similar to those conducted at the individual CLU sites were attempted for county and state level results, however the heterogeneity of soil type, weather, and crop response resulted in generally poor correlations. Hess et al. (2020) found that intensified rainfall led to increased extractable soil nitrate in soil when compared to a control plot, suggesting increased rainfall over the region (e.g. Baule et al. 2022; Kunkel et al. 2020; Hayhoe et al. 2018) may play a variable role over the time in regional nitrogen dynamics. Further study and exploration of our modeling results is needed to further understand this linkage and representation of this effect within our modeling framework. While nitrogen dynamics are complex and non-linear, our results in terms of leaching amounts compare well with plot studies and the ranges presented in the literature for the region for nitrate

leaching (Hess et al. 2020; Syswerda et al. 2012; Zhou and Butterbach-Bahl 2014; Basso and Ritchie 2005), even for the low stable and unstable YSZs.

When soil variability is constrained, as is the case with single CLUs within our framework, the dynamics of the soil-plant-atmosphere continuum become clearer. Regionally there appears to be a spatial pattern in our results, with nitrate leaching at locations in western and northern sections of the study domain more strongly associated with precipitation variability, in the sense that enough water to move nitrate through the soil profile is present and suggests a tie to water stress being a driver for conditions susceptible to nitrate loss. In eastern sections of the region where precipitation is more consistent year to year, throughout the season, and often exceeds crop water requirements (Figure 4.15), leaching values over time are more consistent with less variability and fewer heavy precipitation-driven events. This suggests that precipitation variability isn't the limiting constraint in these areas, but rather the availability of water and supportive thermal environment within the growing season that are better correlated with nitrate leaching. The unstable YSZs were better correlated overall with hydroclimatic drivers of nitrate leaching than the stable zones.

In these simulations, management played a decisive role. The highest correlating variables with soy crops were often those from the season prior (e.g. Lag1), which were fertilized corn crops. This suggests that the management and the weather during the prior growing season has the largest effect on nitrate lost in years when soy is grown. With differing crop rotations and management practices (e.g. nutrient applications, tillage) the response of the system may be substantially different in magnitude and temporal characteristics than those presented here, illustrating the role of the producer in management and outcomes. Given the closer ties to in-season weather, precision application of nutrients to unstable zones is one potential strategy for

controlling excessive nutrient losses from these zones. With more stable behavior, the HS and LS zones are likely appropriate for longer-term strategic management decisions. While the relative behavior of the zones and their differing responses is interesting and potentially valuable for producers in those locations, the role of production intensity in our simulations is also clear. While the largest individual seasonal totals of nitrate leaching were simulated in western sections of the region, the largest contributors of nitrate leached (60%) in bulk are the traditional corn belt states of Illinois, Indiana, and Iowa, where the density of production is most intense on the state level.

This study presents several avenues for future research and follow up studies. Given computational constraints at the time of simulation, the uniform crop rotation by year could likely be improved upon, to improve the fidelity of the simulation when compared to the actual agricultural landscape. The current set of simulation uses only a corn-soy rotation, in addition to randomization of starting crops, simulation of more complex or simpler (i.e., continuous corn) crop rotations. One potentially valuable addition would be the inclusion of simulations with a cover crop, which have been shown to be an effective approach to reduce nitrogen loss during the non-growing season (Christopher et al, 2021). Simulations here were also conducted using a split nitrogen application, which is an approach designed to reduce nitrogen loss, which could suggest that our results are conservative in nature. Additional simulations conducted with other fertilizer application timings, such as a fall bulk application, could shed more light on sources and timing of nitrate leaching under more traditional management practices. These differences in rotation and fertilizer timing could result in substantially different results than those presented here, illustrating the essential role of management in management nitrogen leaching and achieving environmentally and economically sustainable yields.

141

REFERENCES

REFERENCES

Abatzoglou, J. T., 2013: Development of gridded surface meteorological data for ecological applications and modelling. *Int. J. Climatol.*, **33**, 121–131, https://doi.org/10.1002/joc.3413.

Ballard, T. C., E. Sinha, and A. M. Michalak, 2019: Long-Term Changes in Precipitation and Temperature Have Already Impacted Nitrogen Loading. *Environ. Sci. Technol.*, acs.est.8b06898, https://doi.org/10.1021/acs.est.8b06898.

Basso, B., and J. T. Ritchie, 2005: Impact of compost, manure and inorganic fertilizer on nitrate leaching and yield for a 6-year maize–alfalfa rotation in Michigan. *Agric. Ecosyst. Environ.*, **108**, 329–341, https://doi.org/10.1016/j.agee.2005.01.011.

Basso, B., and J. Ritchie, 2015: Simulating crop growth and biogeochemical fluxes in response to land management using the SALUS model. *The Ecology of Landscapes: Long-Term Research on the Path to Sustainability*, S. Hamilton, J. Doll, and G.P. Roberston, Eds., Oxford University Press, 252–274.

Basso, B., M. Bertocco, L. Sartori, and E. C. Martin, 2007: Analyzing the effects of climate variability on spatial pattern of yield in a maize–wheat–soybean rotation. *Eur. J. Agron.*, **26**, 82–91, https://doi.org/10.1016/J.EJA.2006.08.008.

——, J. T. Ritchie, D. Cammarano, and L. Sartori, 2011: A strategic and tactical management approach to select optimal N fertilizer rates for wheat in a spatially variable field. *Eur. J. Agron.*, **35**, 215–222, https://doi.org/10.1016/J.EJA.2011.06.004.

——, and Coauthors, 2016: Tradeoffs between Maize Silage Yield and Nitrate Leaching in a Mediterranean Nitrate-Vulnerable Zone under Current and Projected Climate Scenarios. *PLoS One*, **11**, e0146360, https://doi.org/10.1371/journal.pone.0146360.

——, G. Shuai, J. Zhang, and G. P. Robertson, 2019: Yield stability analysis reveals sources of large-scale nitrogen loss from the US Midwest. *Sci. Rep.*, **9**, 5774, https://doi.org/10.1038/s41598-019-42271-1.

Baule, W. J., J. A. Andresen, and J. A. Winkler, 2022: Trends in Quality Controlled Precipitation Indicators in the United States Midwest and Great Lakes Region. *Front. Water*, **4**, 8, https://doi.org/10.3389/FRWA.2022.817342/BIBTEX.

Congreves, K. A., B. Dutta, B. B. Grant, W. N. Smith, R. L. Desjardins, and C. Wagner-Riddle, 2016: How does climate variability influence nitrogen loss in temperate agroecosystems under contrasting management systems? *Agric. Ecosyst. Environ.*, **227**, 33–41, https://doi.org/10.1016/J.AGEE.2016.04.025.

Deines, J. M., R. Patel, S. Z. Liang, W. Dado, and D. B. Lobell, 2021: A million kernels of truth: Insights into scalable satellite maize yield mapping and yield gap analysis from an extensive ground dataset in the US Corn Belt. *Remote Sens. Environ.*, **253**, 112174, https://doi.org/10.1016/J.RSE.2020.112174.

Gerakis, A., D. P. Rasse, Y. Kavdir, A. J. M. Smucker, I. Katsalirou, and J. T. Ritchie, 2006: Simulation of Leaching Losses in the Nitrogen Cycle. *Commun. Soil Sci. Plant Anal.*, **37**, 1973–1997, https://doi.org/10.1080/00103620600767462.

Gijsman, A.J., G. Hoogenboom, W.J. Parton, and P.S. Kerridge, 2002: Modifying DSSAT crop models for low-input agricultural systems using a soil organic matter-residue module from CENTURY. Agron. J. **94**, 462–474

Hatfield, J., 2012: Agriculture in the Midwest. *U.S. National Climate Assessment Midwest Technical Input Reports*, J.A. Winkler, J. Andresen, J. Hatfield, D. Bidwell, and D. Brown, Eds., Great Lakes Integrated Sciences + Assessments.

Hayhoe, K., and Coauthors, 2018: *Chapter 2 : Our Changing Climate. Impacts, Risks, and Adaptation in the United States: The Fourth National Climate Assessment, Volume II*.

Hess, L. J. T., E. L. S. Hinckley, G. P. Robertson, and P. A. Matson, 2020: Rainfall intensification increases nitrate leaching from tilled but not no-till cropping systems in the U.S. Midwest. *Agric. Ecosyst. Environ.*, **290**, 106747, https://doi.org/10.1016/J.AGEE.2019.106747.

Iqbal, J., and Coauthors, 2018: Extreme weather-year sequences have nonadditive effects on environmental nitrogen losses. *Glob. Chang. Biol.*, **24**, e303–e317, https://doi.org/10.1111/gcb.13866.

Jin, Z., S. V. Archontoulis, and D. B. Lobell, 2019: How much will precision nitrogen management pay off? An evaluation based on simulating thousands of corn fields over the US Corn-Belt. *F. Crop. Res.*, **240**, 12–22, https://doi.org/10.1016/J.FCR.2019.04.013.

Johnston, C. A., 2014: Agricultural expansion: land use shell game in the U.S. Northern Plains. *Landsc. Ecol.*, **29**, 81–95, https://doi.org/10.1007/s10980-013-9947-0.

Jones, J. W., and Coauthors, 2016: Brief history of agricultural systems modeling. https://doi.org/10.1016/j.agsy.2016.05.014.

Kalkhoff, S. J., L. E. Hubbard, M. D. Tomer, and D. E. James, 2016: Effect of variable annual precipitation and nutrient input on nitrogen and phosphorus transport from two Midwestern agricultural watersheds. *Sci. Total Environ.*, **559**, 53–62, https://doi.org/10.1016/j.scitotenv.2016.03.127.

Kiefer, M. T., J. A. Andresen, D. Doubler, and A. Pollyea, 2019: Development of a gridded reference evapotranspiration dataset for the Great Lakes region. *J. Hydrol. Reg. Stud.*, **24**, 100606, https://doi.org/10.1016/j.ejrh.2019.100606.

Kolady, D. E., E. Van der Sluis, M. M. Uddin, and A. P. Deutz, 2021: Determinants of adoption and adoption intensity of precision agriculture technologies: evidence from South Dakota. *Precis. Agric.*, **22**, 689–710, https://doi.org/10.1007/S11119-020-09750-2/TABLES/10.

Kunkel, K. E., T. R. Karl, M. F. Squires, X. Yin, S. T. Stegall, and D. R. Easterling, 2020a: Precipitation Extremes: Trends and Relationships with Average Precipitation and Precipitable Water in the Contiguous United States. *J. Appl. Meteorol. Climatol.*, **59**, 125–142, https://doi.org/10.1175/JAMC-D-19-0185.1.

——, S. E. Stevens, L. E. Stevens, and T. R. Karl, 2020b: Observed Climatological Relationships of Extreme Daily Precipitation Events With Precipitable Water and Vertical Velocity in the Contiguous United States. *Geophys. Res. Lett.*, **47**, e2019GL086721, https://doi.org/10.1029/2019GL086721.

Lark, T. J., J. Meghan Salmon, and H. K. Gibbs, 2015: Cropland expansion outpaces agricultural and biofuel policies in the United States. *Environ. Res. Lett.*, **10**, 044003, https://doi.org/10.1088/1748-9326/10/4/044003.

Li, C. S., 2000: Modeling trace gas emissions from agricultural ecosystems. *Methane Emissions from Major Rice Ecosystems in Asia*, Springer Netherlands, 259–276.

Li, Z. T., J. Y. Yang, C. F. Drury, and G. Hoogenboom, 2015: Evaluation of the DSSAT-CSM for simulating yield and soil organic C and N of a long-term maize and wheat rotation experiment in the Loess Plateau of Northwestern China. *Agric. Syst.*, **135**, 90–104, https://doi.org/10.1016/J.AGSY.2014.12.006.

Lin, M., and M. Henry, 2016: Grassland and Wheat Loss Affected by Corn and Soybean Expansion in the Midwest Corn Belt Region, 2006–2013. *Sustainability*, **8**, 1177, https://doi.org/10.3390/su8111177.

Liu, H. L., and Coauthors, 2011: Using the DSSAT-CERES-Maize model to simulate crop yield and nitrogen cycling in fields under long-term continuous maize production. *Nutr. Cycl. Agroecosystems*, **89**, 313–328, https://doi.org/10.1007/s10705-010-9396-y.

Loecke, T. D., A. J. Burgin, D. A. Riveros-Iregui, A. S. Ward, S. A. Thomas, C. A. Davis, and M. A. St. Clair, 2017: Weather whiplash in agricultural regions drives deterioration of water quality. *Biogeochemistry*, **133**, 7–15, https://doi.org/10.1007/s10533-017-0315-z.

Maharjan, G. R., A. K. Prescher, C. Nendel, F. Ewert, C. M. Mboh, T. Gaiser, and S. J. Seidel, 2018: Approaches to model the impact of tillage implements on soil physical and nutrient properties in different agro-ecosystem models. *Soil Tillage Res.*, **180**, 210–221,

https://doi.org/10.1016/J.STILL.2018.03.009.

Mladenoff, D. J., R. Sahajpal, C. P. Johnson, and D. E. Rothstein, 2016: Recent Land Use Change to Agriculture in the U.S. Lake States: Impacts on Cellulosic Biomass Potential and Natural Lands. *PLoS One*, **11**, e0148566, https://doi.org/10.1371/journal.pone.0148566.

Paiao, G. D., F. F. Fernández, J. A. Spackman, D. E. Kaiser, and S. Weisberg, 2020: Ground-based optical canopy sensing technologies for corn–nitrogen management in the Upper Midwest. *Agron. J.*, **112**, 2998–3011, https://doi.org/10.1002/AGJ2.20248.

Pryor, S. C., J. A. Howe, and K. E. Kunkel, 2009: How spatially coherent and statistically robust are temporal changes in extreme precipitation in the contiguous USA? *Int. J. Climatol.*, **29**, 31–45, https://doi.org/10.1002/joc.1696.

Rosenzweig, C., and Coauthors, 2018: Coordinating AgMIP data and models across global and regional scales for 1.5C and 2.0C assessments. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.*, **376**, https://doi.org/10.1098/RSTA.2016.0455.

Sinha, E., and A. M. Michalak, 2016: Precipitation Dominates Interannual Variability of Riverine Nitrogen Loading across the Continental United States. *Environ. Sci. Technol.*, **50**, 12874–12884, https://doi.org/10.1021/acs.est.6b04455.

Smith, W., and Coauthors, 2019: Assessing the Impacts of Climate Variability on Fertilizer Management Decisions for Reducing Nitrogen Losses from Corn Silage Production. *J. Environ. Qual.*, **48**, 1006, https://doi.org/10.2134/jeq2018.12.0433.

Syswerda, S. P., B. Basso, S. K. Hamilton, J. B. Tausig, and G. P. Robertson, 2012: Long-term nitrate loss along an agricultural intensity gradient in the Upper Midwest USA. *Agric. Ecosyst. Environ.*, **149**, 10–19, https://doi.org/10.1016/j.agee.2011.12.007.

USDA, 2014: Gridded Soil Survey Geographic (gSSURGO) Database for the United States of America and the Territories, Commonwealths, and Island Nations Served by the USDA-NRCS.

Van Metre, P. C., J. W. Frey, M. Musgrove, N. Nakagaki, S. Qi, B. J. Mahler, M. E. Wieczorek, and D. T. Button, 2016: High Nitrate Concentrations in Some Midwest United States Streams in 2013 after the 2012 Drought. *J. Environ. Qual.*, **45**, 1696–1704, doi:10.2134/jeq2015.12.0591. http://doi.wiley.com/10.2134/jeq2015.12.0591

Wallach, D., D. Makowski, J. W. Jones, and F. Brun, 2019: *Working with Dynamic Crop Models : Methods, Tools and Examples for Agriculture and Environment.* 3rd ed. 616 pp.

Wimberly, M. C., L. L. Janssen, D. A. Hennessy, M. Luri, N. M. Chowdhury, and H. Feng, 2017: Cropland expansion and grassland loss in the eastern Dakotas: New insights from a farm-level survey. *Land use policy*, **63**, 160–173, https://doi.org/10.1016/J.LANDUSEPOL.2017.01.026.

Zhou, M., and K. Butterbach-Bahl, 2014: Assessment of nitrate leaching loss on a yield-scaled basis from maize and wheat cropping systems. *Plant Soil*, **374**, 977–991, https://doi.org/10.1007/s11104-013-1876-9.

Zotarelli, L., J. M. Scholberg, M. D. Dukes, and R. Muñoz-Carpena, 2007: Monitoring of Nitrate Leaching in Sandy Soils. *J. Environ. Qual.*, **36**, 953, https://doi.org/10.2134/jeq2006.0292.

CHAPTER 5. CONCLUSIONS

Summary of Results

The previous three chapters present several new findings and approaches to the treatment, analysis, and impact modeling of climate, its inherent variability, and change over observational record.

The first two chapters are more directly focused on the climate system and the statistical behavior of climatic variables. Chapter 2 highlights the development of a method to bias-correct reanalysis-based solar radiation data over a 13-state region of the Midwestern United States during the April-September growing season. The impacts of the correction are illustrated through process-based modeling and illustrate the spatially variable nature of the bias. Chapter 3 illustrates the effect of data quality and statistical confidence level on the spatial interpretation of precipitation indicators, using a novel multitiered approach to assess data completeness, detect observer bias, and unreported break points.

Chapter 4 presents a novel approach to simulating sub-field yield variability and associated nitrate leaching using process-based crop simulation models. Analysis of nitrate leaching output suggests that drivers of nitrate leaching at a specific location (i.e., individual field) vary across a 14-state region of the broader Midwest, ranging from growing season precipitation to the frequency of water stress, and within the same field by YSZ. In aggregate, interaction between topography, soil variability, SOM, and background hydroclimatic appear to be responsible for the characteristics described by larger spatial aggregation (i.e., county, state, region). Results from this modeling exercise are calibrated to grain yields across the region and results are in line with many field studies meta-analyses of field results related to nitrate leaching (e.g., Basso et al., 2019; Syswerda et al., 2012; Zhou and Butterbach-Bahl, 2014, Basso and Ritchie 2005). The relative results between YSZs across the region suggest that management

does and can play a critical role in managing nutrients, generally suggesting precision approaches to nutrients are advantageous both economically and environmentally. While the approach to yield stability zones and their simulation is novel, there have been recent studies that have taken both statistical (Deines et al. 2021) and process-based (Jin et al. 2019) approaches to spatial simulations of Midwestern field crop agriculture that incorporate sub-field variability. The work here is primarily focused on nitrate leaching, while Jin et al. (2019) and Deines et al. (2021) were focused primarily on yield. All three studies mentioned, illustrate the utility and potential for developing management plans that address sub-field variability utilizing publicly available resources.

This dissertation offers a collection of works each focused on a different aspect of the Midwestern agricultural-climate interface, rooted in classic geographical/climatological techniques and implementing cutting-edge simulation framework to increase understanding around the impacts fueled by both climate and management strategies. Suggestions for future research in the next section could solidify conclusions drawn from model outputs and offer insights into potential causes of observer bias in precipitation in the COOP network and increase understanding of the flexibility of the bias-correction method conducted in Chapter 2.

Recommendations for Future Research

Increased validation of different components of the YSZ framework would offer increased confidence in model outputs and potential for model improvements. In terms of field measurements, further field sampling of nitrate levels in various YSZs across the landscape or installation of lysimeters would increase confidence in the nitrate leaching outputs from Chapter 4. Recent analyses of model outputs suggest a temporal connection to the frequency of water stress and organic matter content that needs further exploration that may aid in interpretation of

correlations with driving hydroclimatic variables (e.g., Shuai and Basso, 2022). Additionally, application of the YSZ approach to climate projections could be valuable for the design of management plans and crop genetics. This will be essential to sustain yields as climate projections indicate future growing seasons will include conditions outside of the observed envelope of historical growing season conditions (e.g., Baule et al., 2017, Basso et al., 2021, Tang et al., 2020).

While the question of trends in precipitation in the Midwestern United States is a topic well covered in the literature, questions remain regarding the stability of observer bias over time and the causes of observer bias. That is, "Are sources of observer-bias linked to how observers are trained?" The geographical representation of stations that Pass or Failed the observer bias tests in Chapter 3 (Figure 1, Ch 3.) suggests that something other than the climate and weather of a region may be responsible. Further investigation into how observers have been trained over time to conduct observations and how this presents itself in the data would be an interesting multidisciplinary follow up to the results presented in Chapter 3.

Broadening the application of the bias correction technique to other gridded solar radiation products from other sources (e.g., Slater, 2016) would greatly aid in the assessment of the flexibility of the correction method documented in Chapter 2. Additionally, explorations of corrections of bias in climate model outputs from the Coupled Model Intercomparison Project Phase 6 (CMIP6) and their potential impact on process-based crop model simulations of yield and water-stress.

REFERENCES

REFERENCES

Basso, B., and J. T. Ritchie, 2005: Impact of compost, manure and inorganic fertilizer on nitrate leaching and yield for a 6-year maize–alfalfa rotation in Michigan. *Agric. Ecosyst. Environ.*, **108**, 329–341, https://doi.org/10.1016/j.agee.2005.01.011.

Basso, B., R.A. Martinez-Feria, L. Rill, and J.T. Ritchie, 2021: Contrasting long-term temperature trends reveal minor changes in projected potential evapotranspiration in the US Midwest. *Nature Communications*, **12**, 1476, doi:10.1038/s41467-021-21763-7

——, G. Shuai, J. Zhang, and G. P. Robertson, 2019: Yield stability analysis reveals sources of large-scale nitrogen loss from the US Midwest. *Sci. Rep.*, **9**, 5774, https://doi.org/10.1038/s41598-019-42271-1.

Baule, W., B. Allred, J. Frankenberger, D. Gamble, J. Andresen, K. M. Gunn, and L. Brown, 2017: Northwest Ohio crop yield benefits of water capture and subirrigation based on future climate change projections. *Agric. Water Manag.*, **189**, doi:10.1016/j.agwat.2017.04.019.

Deines, J. M., R. Patel, S. Z. Liang, W. Dado, and D. B. Lobell, 2021: A million kernels of truth: Insights into scalable satellite maize yield mapping and yield gap analysis from an extensive ground dataset in the US Corn Belt. *Remote Sens. Environ.*, **253**, 112174, https://doi.org/10.1016/J.RSE.2020.112174.

Jin, Z., S. V. Archontoulis, and D. B. Lobell, 2019: How much will precision nitrogen management pay off? An evaluation based on simulating thousands of corn fields over the US Corn-Belt. *F. Crop. Res.*, **240**, 12–22, https://doi.org/10.1016/J.FCR.2019.04.013.

Shuai, G. and B. Basso, 2022: Subfield maize yield prediction improves when in-season crop water deficit is included in remote sensing imagery. *Remote Sensing of the Environment*, **272**, 112938, doi:10.1016/j.rse.2022.112938

Syswerda, S. P., B. Basso, S. K. Hamilton, J. B. Tausig, and G. P. Robertson, 2012: Long-term nitrate loss along an agricultural intensity gradient in the Upper Midwest USA. *Agric. Ecosyst. Environ.*, **149**, 10–19, https://doi.org/10.1016/j.agee.2011.12.007.

Tang, Y., J.A. Winkler, A. Viña, F. Wang, J. Zhang, Z. Zhao, T. Connor, H. Yang, Y. Zhang, X. Zhang, X. Li, and J. Liu, 2020: Expanding ensembles of species present-day and future climatic suitability to consider the limitations of species occurrence data. *Ecol. Indic.*, **110**, doi: 10.1016/J.ECOLIND.2019.105891.

Zhou, M., and K. Butterbach-Bahl, 2014: Assessment of nitrate leaching loss on a yield-scaled basis from maize and wheat cropping systems. *Plant Soil*, **374**, 977–991, https://doi.org/10.1007/s11104-013-1876-9.