# Cereal grain yield responses to fertilizer management in sandy soil in a long-term fertilizer experiment in Northeast Germany

## DISSERTATION

Zur Erlandung des akademischen Grades

Doctor rerum agriculturarum

(Dr.rer.agr)

**Eingereich an der**

Lebenswisseschaftlichen Fakultät der

Humboldt-Universität zu Berlin

**Von**

M.Sc. Thai, Thi Huyen

Präsident der Humboldt-Universität zu Berlin

Prof. Dr. Julia von Blumenthal

Dekan der Lebenswissenschaftlichen Fakultät

Prof. Dr. Christian Ulrichs

<u>Gutachter</u>

1. Prof. Dr. Sonoko D. Bellingrath-Kimura
2. Prof. Dr. Timo Kautz
3. Prof. Dr. Huynh Van Chuong

Tag der mündlichen Prüfung: 14. 11. 2022

# Acknowledgements

The accomplishment of this thesis involved many people and organizations, who directly or indirectly helped me during the last years. Thus, I would like to express my gratefulness to all those involved.

Acknowledgements go out to all the organizations and individuals involved in and supported my PhD project. My special thankfulness goes out to the Vietnamese government for the financial support through the 911 project. I would like to thank the Provisioning of Ecosystem Services in Agricultural Systems Group, Research Area 2 "Land Use and Governance", Leibniz Centre for Agricultural Landscape Research (ZALF), where the study was conducted, for her provision of facilities as well as the enthusiasm and helpfulness of her staff. I express my sincere thanks to the Müncheberg Experimental Station, ZALF, for providing the data that were used in this thesis. Special thanks go to the current manager of the Müncheberg Experimental Station and several generations of technical staff for the long-term maintenance of the experiment. I acknowledgements to Faculty of Life Science, Humboldt-Universität zu Berlin, where educated me for PhD; and to Department of Crop Science, Faculty of Agronomy, University of Agriculture and Forestry, Hue University facilitate me in terms of time to do study in Germany.

"You cannot make your name without the help of your teachers". My sincere acknowledgements are to my supervisor, Prof. Dr. Sonoko Dorothea Bellingrath-Kimura, your great and patient supervision during the last years helped me to accomplish this thesis. You not only helped me with scientific supervision but also in my life in Germany. You always stood behind me to help me in discussing and completing research proposals and all chapters of this thesis; to help in solving difficulties and challenges in PhD life, and to share great times during my working time at ZALF. Your bright ideas during discussions kept the project running on a track, sharpened my research questions, and research papers. I am especially impressed and inspired by your creative ideas, liberal writing style, encouragement, and patience. I learned a lot from you, not only with respect to scientific knowledge and critical thinking but also regarding your very friendly behavior. I am highly indebted to you.

I cannot express my sincere thanks enough to the people, who helped me a lot over the past more than four years in Müncheberg. I would like to express my appreciation to Dr. Dietmar Barkusky, Dr. Carsten Hoffmann, Dr. Richard Ansong Omari, and Sravya Mamidanna for your constructive comments and suggestions for my work. I would like to give my thanks to great friends at ZALF, and at WBZ, especially Annemarie Winter, Viola Kranich, and Annett Jahn who helped me a lot in my life at ZALF.

Deep appreciation goes to all friends and relatives in Vietnam and overseas, especially to those in Germany for their direct or indirect help. Several sentences could not list all these people. However, I would like to give my thanks to my colleagues, (and) my friends at

ZALF, Müncheberg for their hospitality and the family atmosphere dinners. Thanks are also due to all my Vietnamese friends studying in Berlin, for your help and encouragement; for all activities, and for the moments we shared together during the last four years.

Most sincere gratefulness goes to my beloved mum and dad who gave their utmost love to me and always support my ways. Kind appreciation is for my relatives' contribution. In particular, I would like to give my thanks to my three younger sisters, Thu Thai, Thuy Thai, and Hang Thai. You are always there with me when I am happy and when I am down. Sincere thankfulness to my mother-in-law for her love and her support to me.

I would like to give my greatest thanks to my beloved late husband Le Nhu Cuong, my daughter, Le Nhu Bang Chau, and my son, Le Bao Chau for their sacrifice; living without the help of his wife and their mum. I am blessed with love, sympathy, and encouragement from my late husband and my children whose are my great source of inspiration and motivation. Cuong and Bang Chau, Bao Chau, you are always my pride.

This thesis is dedicated to my parents, and my children, especially to my husband's memory. Only for my beloved late husband, he passed away when I have done almost work in Germany and have just come back to Vietnam. My greatest appreciation to him who sacrificed and loved me and my children infinitely. Cuong, you are being forever in my mind, in my heart, in my soul. You are still here and always with me to share my success. I love you forever.

Last but not least, I would like to thank all those who helped me to prepare this treatise, one or another way.

Once again, thank you very much for all.

Thai, Thi Huyen

# Zusammenfassung

Dauerdüngungsversuche (LTFEs) sind für die landwirtschaftliche Forschung von hoher Bedeutung, da sie zum einen Ertragsschwankungen über lange Zeiträume dokumentieren und zum anderen zukünftige Ereignisse simulieren können. Die Untersuchung der Pflanzenreaktion auf landwirtschaftliche Bewirtschaftungspraktiken und Umweltveränderungen in der Vergangenheit liefern wichtige Erkenntnisse, um das zukünftige Pflanzenwachstum unter veränderten Klimabedingungen abzuschätzen. Vor diesem Hintergrund zielte die aktuelle Studie darauf ab, Ertragsreaktionen von Getreide auf das Düngemanagement in einem LTFE (1971 bis 2016) mit sandigem Boden in Nordostdeutschland, zu analysieren. Die konkreten Ziele lagen a) in der Reaktion des Getreideertrags auf das Düngemittelmanagement, b) die Bestimmung des Einflusses unterschiedlicher Witterungsphasen und c) der Vergleich verschiedener Analysemodelle. Es wurden die Getreidearten Sommergerste (SB), Winterroggen (WR) und Winterweizen (WW) betrachtet, und deren die Kornerträge auf dem LTFE analysiert.

Die Studie zeigte, dass der Getreideertrags neben dem Düngemittelmanagement auf eine komplexe Beziehungen zwischen klimatischer Abhängigkeit, vorausgehender Ernte und Bodeneigenschaften reagierte. Düngung war der wichtigste Bestimmungsfaktor für den WR-Ertrag (48 %, Kapitel 3), während der Einfluss auf den WW-Ertrag 34 % (Kapitel 4) und auf den SB-Ertrag 11 % (Kapitel 2) betrugen. Zunächst wurden düngebedingte Ertragsschwankungen von SB unter jährlichen Witterungsschwankungen auf sandigen Böden untersucht. Die kombinierte Düngung mit chemischem Stickstoffdünger (NF) verbesserte die Ertragsstabilität des Getreides. Eine geeignete kombinierte Anwendung von NF und organischem Dünger führte zu einem besseren Getreideertrag als einzelne Anwendungen von ausschließlich NF oder organischem Dünger (Kapitel 2, 3 und 4). Für Wintergetreide (WR und WW) wurde die NF-Ausbringung als Hauptdeterminante des Kornertrags identifiziert (Kapitel 3, 4). Es wurde geschlussfolgert, dass Düngemittelgaben weniger Auswirkungen auf den Sommergerste-Ertrag haben als auf die Kornerträge von Winterroggen und Winterweizen.

Die Witterungsbedingungen bei der Aussaat und im frühen Wachstumsstadium des Getreides (Frühling für SB, Herbst für WR und WW) beeinflussen maßgeblich die Ertragswerte (Kapitel 2, 3 und 4). Für Wintergetreide war die Intensität und Dauer extremer Temperaturen im Sommer (Erntejahr), insbesondere die Anzahl der Tage mit einer Höchsttemperatur über 30 °C im Juli (kumulierte Tage Tmax > 30 °C im Juli) eine wichtige Variable für den Ertrag (Kapitel 3, 4). Der meteorologische Jahresverlauf war der wichtigste Faktor, der den SB-Ertrag (55 %, Kapitel 2) und den WW-Ertrag (42 %, Kapitel 4) bestimmte, während Witterungsfaktoren den WR-Ertrag um 32 % beeinflussten (Kapitel 3). Bei Sommergerste wirkten sich die Niederschlagsmenge im März und die Temperatur im April negativ auf den Kornertrag aus. Zudem wirkte sich der Gesamtniederschlag während der Vegetationsperiode (April-Juli) bei gleichzeitiger mineralischer Stickstoffdüngung

positiv auf den SB-Ertrag aus (Kapitel 2). Für WR waren die Temperatur im September und Oktober, der Niederschlag im November, die Temperatur im Dezember und Mai und die kumulativen Tage Tmax > 30 °C im Juli wichtige Witterungsvariablen, die Schwankungen des Gewinnertrags erklärten. Unter den genannten waren die kumulierten Tage Tmax > 30 °C im Juli die wichtigste Einflussgröße für den WR-Ertrag (Kapitel 3).

Bei WW waren die Temperatur im Oktober, die kumulative Anzahl der Frosttage im Dezember und Februar, die Niederschläge im Juni und die kumulativen Tage Tmax > 30 °C im Juli wichtige Variablen, die Schwankungen in den Kornerträgen erklärten. Die kumulative Anzahl der Frosttage im Dezember war dabei die wichtigste Variable für den WW-Ertrag (Kapitel 4). Neben Düngung und Witterungsbedingungen wurden weitere agronomisch relevante Faktoren wie Bodenparameter und Vorfrucht berücksichtigt. Die Vorfrucht- und Bodenparameter könnten einen Einfluss auf den Kornertrag von Sommergerste haben (Kapitel 2), während die Vorfruchtart und der Vorfruchtertrag, der Gesamt-N im Boden und der organische Kohlenstoff im Boden Variablen sind, die den Kornertrag von WR (Kapitel 3) und WW (Kapitel 4) beeinflussen.

Ein Vergleich verschiedener Analysemethoden in der Studie verstärkt die aufgeführten Ergebnisse. Insbesondere zeigten Untersuchungen, dass das ANOVA-Ergebnis und das Modell GLM nur den Zielfaktor lieferten, der den Getreideertrag beeinflusste (Kapitel 2, 3 und 4). Währenddessen zeigte BMA die Witterungsvariablen als Hauptfaktor für den SB-Ertrag, unterschätze hingegen Vorfrucht- und Bodenvariablen (Kapitel 2). Das M5P-Modell zeigte eine gute Vorhersageleistung als weitere Analyse nach GLM, um (i) lineare, nichtlineare und kombinierte Wechselwirkungen auf den Wintergetreideertrag aufzudecken und (ii) kritische Schwellenwerte für die Erklärung der Variablen und ihren Einfluss auf den Wintergetreideertrag zu identifizieren. Diese Analysen setzen jedoch Modellanpassung und anschließenden Interpretation voraus (Kapitel 3, 4). LMM zeigte eine bessere Vorhersageleistung im Vergleich zu M5P (Kapitel 4). Statistische Methoden wie LMM konzentrieren sich seit langem auf die Inferenz aus einer Stichprobe, während sich maschinelle Lernmodelle wie M5P auf die Vorhersage konzentrieren, um verallgemeinerbare Vorhersagemuster zu finden. Daher wurde die gemeinsame Verwendung verschiedener Analysemodelle wie ANOVA/GLM, M5P und LMM-Modell untersucht (Kapitel 4), um die Inferenz- und Vorhersageertragsreaktionen von Getreide im LTFE zu bewerten. Diese Kombination könnte dazu beitragen, methodische Mängel in Zukunft zu reduzieren oder zu beheben.

Diese Studie kommt zu dem Schluss, dass saisonale Wettervorhersagen und dazu passende Aussaattermine wichtige Faktoren sind, um die Erträge zu verbessern und die Ertragsvariabilität in SB und WW in sandigen Böden zu verringern. Die Bereitstellung nennenswerter Mengen an mineralischem NF und Stalldünger sind wichtige Einflussgrößen für einen erhöhten Kornertrag für WR. Es ist daher wichtig, die Aussaattermine auf geeignete Zeiten zu legen, um ein optimales Wachstum des Getreides im Frühjahr zu

gewährleisten. Daneben beeinflussten auch Extremwitterungsereignisse im Winter und Sommer das Wachstum, die Entwicklung und den Ertrag des Wintergetreides. Daher ist es notwendig, das Management geeigneter Vorfrüchte und/oder die Verwendung geeigneter Weizen- und Roggensorten anzupassen, um auf jährliche Wetteränderungen zu reagieren. Darüber hinaus wird empfohlen, die Bewässerung für einige Getreidesorten an trockenen Sommertagen in Betracht zu ziehen, insbesondere während Dürreperioden. Diese Studie zeigt die Notwendigkeit eines geeigneten Düngemanagement für Getreide auf einer sandigen LTFE-Versuchsfläche in Nordost-Deutschland. Eine an die Getreidearten angepasste Menge und Kombination von mineralischem NF und Wirtschaftsdünger sollte ausgewählt werden, um die Erträge zu optimieren. Die Ergebnisse dieser Analyse tragen dazu bei, die in der Literatur beschriebenen Strategien für eine nachhaltige Pflanzenproduktion in Zeiten des Klimawandels zu ergänzen.

**Keyword:** Langzeitversuche, Getreideertrag, Sommergerste, Winterroggen, Winterweizen, Düngung, Klimavariabilität, Nordostdeutschland

# Summary

Long-term fertilizer experiments (LTFEs) are vitally important in agricultural research as they can document, monitor, learn and demonstrate what happened in the past as well as predict and simulate what will happen in the future. By investigating the plant response to agricultural management practices and environmental changes in the past, these provide important knowledge to estimate future plant growth under climate change. Against this backdrop, the current study aimed to analyze cereal grain yield responses to fertilizer management in sandy soil in a long-term (1971 to 2016) fertilizer experiment in Northeast Germany. The objectives of this study were to a) analyze cereal grain yield responses to fertilizer management, b) analyze sensitivity timing of weather events, and c) compare different analysis models. Spring barley (SB), winter rye (WR), and winter wheat (WW) were considered as cereals to analyze the grain yield responses in the LTFE.

The study revealed that cereal yield response to fertilizer management involved complex relationships among climatic dependence, preceding crop, and soil characteristics. Fertilizer was the most important factor determining WR yield (48%, chapter 3), while the rates for WW yield and SB yield were 34% (chapter 4) and 11% (chapter 2). It was suggested that choosing SB as the first sample cereal among cereals planting in the LTFE to test the grain yield was influenced by annual weather condition in sandy soil and dry region as the experimental site. The combined fertilizer application with chemical nitrogen fertilizer (NF) input enhanced the yield stability of cereal. A suitable combined application of NF and organic fertilizer produced a better cereal yield than individual applications of either NF or organic fertilizer (chapters 2, 3 &4). For winter cereals (WR and WW), NF application was identified as the main determinant of the grain yield (chapters 3, 4). It can be explained that fertilizer applications have fewer effects on SB yield than on grain yield of winter rye and winter wheat.

Weather condition at seeding and early growth stage of cereal (springtime for SB, autumn for WR, WW) were found to be the sensitive timing that influence the grain yield (chapter 2, 3 &4). For winter cereals, the intensity and duration of extreme temperatures in the summertime (harvest year), especially the number of days recorded with a maximum temperature above 30°C in July (cumulative days Tmax > 30°C in July) was an important variable for the yield (chapter 3, 4). Annual weather condition is the most important factor determining SB yield (55%, chapter 2) and WW yield (42%, chapter 4), while the weather condition influence WR yield by 32% (chapter 3). For spring barley, the precipitation rate in March and temperature in April negatively affected the grain yield. Meanwhile the total precipitation during the growing season (April-July) positively affected SB yield when high mineral NF application was supplied (chapter 2). For WR, important weather variables explaining the gain yield variation were temperature in September and October, precipitation in November, temperature in December and May, and cumulative days Tmax > 30°C in July. Among these, cumulative days Tmax > 30°C in July was the most important weather variable

that influenced the WR yield (chapter 3). For WW, temperature in October, cumulative number of freezing days in December and February, precipitation in June, and cumulative days Tmax > 30°C in July were important variables explaining the grain yield. Among these variables, the cumulative number of freezing days in December was the most important weather variable influencing the WW yield (chapter 4). Along with fertilizer and weather condition, other agronomic factors such as soil parameter and preceding crop were also considered as factors affecting the grain yield variation of cereals. The preceding crop and soil parameter could have an impact on  grain yield of spring barley (chapter 2), while the preceding crop type and the preceding crop yield, the total N in the soil, and the soil organic carbon are variables that influenced grain yield of winter rye (chapter 3), winter wheat (chapter 4).

The comparison of different analytical methods in the study strengthen the statement of the analysis. In particular, the study indicates that the ANOVA result and the model GLM provided only the target factor affecting cereal yield (chapters 2, 3 &4). Meanwhile, BMA quantified in detail weather variables (as main factor) influence SB yield, however it missed preceding crop, and soil variables in the model (chapter 2). M5P model has well predictive performance as a further analysis after GLM to (i) unravel linear, non-linear interactions and combined effects on winter cereal yield, and (ii) identify critical thresholds of explanatory the variables and their influence the winter cereal yield, but challenged to model fitting and subsequent interpretation (chapters 3, 4). LMM showed a higher predictive performance compared to the M5P (chapter 4). However, statistical methods such LMM have a long-standing focus on inference from a sample, whereas machine learning models such as M5P concentrates on prediction to find in generalizable predictive patterns. Therefore, there was a comprehensive research (chapter 4) to co-use of different analysis models such as ANOVA/GLM, M5P, and LMM model, to investigate the inference and prediction yield responses of cereal in the LTFE. This combination could help to address all those methodological shortcomings.

This study concludes that seasonal weather forecasts and suitable sowing dates are important factors to consider for improving yields and reducing yield variability in SB and WW in sandy soil. Meanwhile, supplying appreciable amounts of mineral NF and farmyard manure are important considerations for increased grain yield in WR. It is thus essential to adjust the sowing dates to suitable times to ensure optimum growth of the cereals in spring. Besides the extreme weather in winter and summer also influenced the growth, development, and yield of the winter cereals. Therefore, it is necessary to adjust the management of appropriate preceding crops and/or the usage of appropriate wheat, and rye cultivars to adapt to year-to-year weather changes. Additionally, it is necessary to consider irrigation for tested cereals during dry summer days, especially during droughts. Regarding fertilizer management for the cereal in the LTFE, the current study highlighted the need to consider the role and amount of nitrogen sources and to choose the optimal amount and combination of mineral NF and farmyard manure in order to get a higher yield for each tested cereal in the sandy soil and dry region as the experimental site. Overall, the findings of this analysis

contribute to the existing literature contribute to comprehensive strategies for sustainable crop production with regard to climate change in the future.

**Keyword:** Long-term experiments, cereal grain yield, spring barley, winter rye, winter wheat, fertilizer, climate variability, northeast Germany

# Table of content

# List of tables

# List of figures

## List of publications

This thesis is based on work in the following two papers published in peer-reviewed journals and one submitted manuscript. The numbering order reflects the order of the papers appears in the text.

1. **Thai, T. H.**, Bellingrath-Kimura, S. D., Hoffmann, C., and Barkusky, D., 2020. Effect of long-term fertilizer regimes and weather on spring barley yields in sandy soil in North-East Germany. *Archives Agronomy and Soil Science*, 66 (13), 1812-1826.
   **Research paper: published**
   **Doi:** https://doi.org/10.1080/03650340.2019.1697436
   **(Chapter 2)**

2. **Thai, T. H.**, Mamidanna, S., Barkusky, D., and Bellingrath-Kimura, S. D., 2020. MP5 machine learning algorithm for analysis of winter rye yield in a long-term field experiment in Müncheberg, Germany. *European Journal of Agronomy*.
   **Research paper (submitted manuscript): under review**
   **(Chapter 3)**

3. **Thai, T. H.**, Omari, RA., Barkusky, D., and Bellingrath-Kimura, S. D., 2020. Statistical analysis versus the M5P machine learning algorithm to analyze the yield of winter wheat in a long-term fertilizer experiment. *Agronomy,* 10 (11), 1779.
   **Research paper: published**
   **Doi:** https://doi.org/10.3390/agronomy10111779
   **(Chapter 4)**

# Abbreviations and acronyms

| | |
|---|---|
| ANOVA | Analysis of variance |
| BMA | Bayesian model averaging |
| CART | Classification and regression trees |
| CEC | Cation exchange capacity |
| CV | Coefficient of variation |
| Days Tmax > 30°C | Cumulative number of days recorded having mean temperatures above 30°C in every month |
| DM | Dry mass |
| DWD | German Meteorological Service |
| FAO | The Food and Agriculture Organization of the United Nations |
| FAOSTAT | Food and Agriculture Organization Statistical Databases of the United Nations |
| FYM | Farmyard manure |
| GDD | Growing degree days |
| GLM | General linear model |
| LMM | Linear mixed-effects model |
| LTFE | Long-term field experiment |
| M5P | M5P machine learning algorithm/ M5P model |
| MAE | Mean absolute error |
| MF | Mineral fertilizer application |
| Mg DM | Mega gram dry mass |
| ML | Machine learning |
| MN | Mineral nitrogen fertilizer |
| MRM | Multiple linear regression model |
| NF | Mineral nitrogen fertilizer |
| OR | Organic fertilizer |
| ORF | Organic fertilizer |
| $R^2_c$ | Conditional coefficient of determination for both fixed and random factors |
| $R^2_m$ | Marginal coefficient of determination for fixed factors alone |
| RCBD | Randomized complete block design |
| RMSE | Root mean square error |
| SB | Spring barley |
| SD | Standard deviation |
| Se | Standard error |
| SOC | Soil organic carbon |
| SOM | Organic carbon matter |
| WEKA | Waikato Environment for Knowledge Analysis |
| WR | Winter rye |
| WW | Winter wheat |
| ZALF | Leibniz Zentrum für Agrarlandschafts Forschung (Leibniz Centre for Agricultural Landscape Research |

# Chapter 1

# General Introduction

## 1.1. Challenges in agricultural production under climate change

Climate change is recognized as one of the most serious environmental threats facing mankind worldwide and will get more serious in the future. It is well proven that climate change affects agriculture in several ways, including its direct impact on food production. The changing climatic conditions are strongly impacting crop farming, ultimately causing a strong reduction in crop yields. It is projected that climate change would affect 15-30% of global agricultural food production by 2080-2100 FAO (2022). It has also been documented that climate change is to end with significant losses in yield levels of various crops and the speed of these losses is predicted to be fastened with the passage of time (Challinor et al., 2014; Gourdji et al., 2013; Rosenzweig et al., 2014).

Generally, climate change influences agriculture and global food security through changes in agroecological conditions (Schmidhuber and Tubiello, 2007). Climate change, which is attributable to the natural climate cycle and human activities, has adversely affected agricultural productivity globally (Ziervogel et al., 2006). Temperature and precipitation are two important parameters of climate. Anomalies and abrupt changes in temperature and precipitation parameters can pose a severe impact on crop farming. As the planet warms, precipitation patterns vary, and extreme events such as droughts, floods, and forest fires become more frequent (Zoellick, 2009), which results in poor and unpredictable yields. For instance, wheat production in Pakistan is under heavy threat due to rising temperatures. It has been predicted that yield will decline by 6-9% from 1°C increase in temperature in dry areas between 2040 to 2069 (Ahmad et al., 2019; Sultana et al., 2009). Also, drought intensity in Germany reduced the mean grain yield of winter rye by 16% in spring 2007 compared to the mean of 2000-2009 (DWD, 2020; Statistisches, 2019). In the same year, the grain yields of wheat, barley, and triticale were reduced by 6-9%. The global air temperature over the last 30 years has been increasing faster than any other period over the last 150 years, with night-time temperatures rising at a faster rate than daytime temperatures (Hartmann et al., 2013). Along with temperature, precipitation has a very strong impact on final crop yield as it is related to water availability and soil moisture, which crops directly use for their growth.

One human contributing factor that is linked to climate change and food production is changes in soil. Changes in soil properties tend to occur slowly. It may take decades to indicate an ongoing negative depletion trend in a soil's productivity and it can take an equally long time to restore it (Kirchmann, 2007). On the other hand, the effects of other agricultural management practices (e.g., fertilization, tillage, crop rotation) can immediately influence on crop yield and soil characteristics. Humans are becoming increasingly aware that their practice is having a great impact on the sustainability of the agricultural system and the earth (Vitousek et al., 1997). The agricultural system should meet the needs of the present without compromising the ability of future generations. In order to develop comprehensive strategies for sustainable crop production, needed to understand what happened in the past and predict trends that will happen in the future. Thus, for the identification of long-term trends, experiments with a long duration are needed.

Long-term field experiments (LTFEs) offer the best practical means of studying the effects of management or anthropogenic global change on soil fertility, sustainability of yield, or wider environmental issues (Grosse et al., 2020; Johnston and Poulton, 2018). LTFEs could test what happened in the past and predict trends that will happen in the future and is thus a useful method for quantifying small changes in soil properties over time. Such experiments, although not perfect, are the only practical ways of assessing the long-term sustainability and productivity of husbandry systems within an agro-ecological zone in which they exist. LTFEs also could monitor the effects of climate change, including increasing atmospheric temperature, and make changes, if needed, to maintain the sustainability of farming in an agroecological system over a long-term span. They could provide data to improve best husbandry practices to benefit farmers, local ecology, and the wider environment. They allow a realistic assessment of the effect of agricultural processes on the environment and of non-agricultural anthropogenic activities on soil fertility and plant quality. LTFEs provide long-term datasets that can be used to develop mathematical models to describe a range of agricultural practices that could be used to predict the effects of climate change on soil properties and the productive capacity of soils.

## 1.2. Long-term field experiments (LTFEs)

### 1.2.1. Background of LTFEs

"Agricultural long-term field experiments (LTFEs) are defined as field experiments with a minimum duration of 20 years and a static design" (Grosse et al., 2019). LTFEs are vitally important sources of knowledge (Debreczeni and Körschens, 2003) for agriculture, nutritional and environmental research (Körschens, 2006). They provide one of the means to measure sustainable management systems in agriculture (Rasmussen et al., 1998). They are records of the past and may serve as early warning systems for the future (Dawe et al., 2000).

The LTFEs offer the possibility to analyze, recognize and document the gradual long-term changes in soil, crop production system and ecosystems occurring as a result of long-term agrotechnical operations including that of fertilization (Debreczeni and Körschens, 2003). They are also vitally important in monitoring, understanding and proving the changes in the agricultural ecosystems such as crop yield and soil fertility. The crop yield and soil characteristics are affected by agricultural management practices (e.g., fertilization, tillage, crop rotation) and can be associated with the inter annual variability of weather (Körschens, 2006; Merbach and Deubel, 2008). The information gained from LTFEs cannot be replaced by other methods, and their scientific value is immeasurable for environmentally friendly land use and sustainable crop production (Merbach and Deubel, 2007; Merbach and Deubel, 2008; Richter et al., 2007). The data from LTFEs can inform the validation of the simulation and prediction data collected under changing soil, managements and climate conditions. They enable to estimate the consequences of current land use management and evaluate future developments and may provide support for decisions about agricultural and environmental policy (Brentrup et al., 2004; Merbach et al., 2013; Willocquet et al., 2008). They allow assessing the nutrient management strategies that sustain crop yield, maintain soil fertility and preserve the environment under varying weather conditions. Moreover, there are also other issues, such as crop rotation and tillage which are conducted in LTFEs. Therefore, LTFEs should be maintained as a scientific heritage for future generations (Körschens, 2006).

Agricultural scientists are interested in studying in LTFEs and analyzing data from the LTFEs (Debreczeni and Körschens, 2003; Körschens et al., 2013) in order to answer the research questions that relate to essential topics such as the effect of fertilizer and management on yield, soil fertility, food quality, pest and pathogens and the interaction between climate, soil and plant. Further, research on LTFEs assist in studies of soil elements such as the development of humus balance methods, soil carbon element research for dynamics, sinks and sources and deducing optimal contents, nitrogen dynamics and nitrogen cycle, the quantification of trace gas emissions or elements related to environmental research and verification of models for applied and environmental research.

In the world, the oldest LTFEs were established at Rothamsted in England by the Rothamsted experimental station's founder J. B. Lawes and J. H. Gilbert in 1843 (Debreczeni and Körschens, 2003). They have run continuously since their foundation for 179-years. Later, different LTFEs were set up in other lands too. A recording done by Debreczeni and Debreczeni and Körschens (2003) indicated that 25 LTFEs in the world are already older than 100 years. In which, 11 of them can be found in England, three in Denmark, two in France, two in Germany, two in Ukraine and five in America. Remaining LTFEs that are younger than 100 years number nearly 600 in different countries of the world, most of them in Europe (Debreczeni and Körschens, 2003).

In Germany, study of Grosse et al. (2020) reported that a total of 205 LTFEs were identified with a minimum duration of 20 years, of which 140 LTFEs are ongoing and 65 are terminated. The of age 22 terminated LTFEs is unknown since the exact ending year is unknown, only the starting date of the LTFEs is known. Three LTFEs run for more than 100 years, 50 trials have a duration between 50 and 99 years, 124 LTFEs have a duration between 20-49 years. The oldest LTFE in Germany were established in 1878 at the Julius- Kühn-Field in Halle.

The study of Grosse et al. (2020) also indicated that LTFEs classification in Germany depend on various usage which covered essential topics and research questions. Usually, the LTFEs are classified according to their research themes, land use and farming system. For land use, the 168 LTFEs are arable field crops, 34 LTFEs are grassland, two LTFEs are for vegetables and one LTFE is for pomiculture. Further, the LTFEs research in the context of soil and yield are classified according to their research themes such as fertilization, tillage, crop rotation or other research themes and their combinations. Thus, most of LTFEs in Germany (n=191) are grouped into three classes of fertilization experiments, tillage experiments and crop rotation experiments. The majority of LTFEs have a research theme "fertilization" (n= 158) and are subdivided into field crops experiments (n=124) and grasslands experiments (n=34). This information is published in an online overview map (https://ltfe-map.bonares.de) which is created by BonaRes Data Centre (Leibniz Centre for Agricultural Landscape Research (ZALF)) (Figure 1.1). Furthermore, the information details of the LTFEs in Germany are also found on the BonaRes Data Portal (Grosse and Hierold, 2019) and in a study of Grosse et al. (2021).

Below is the online overview map of LTFEs in Germany (Figure 1.1). The map content can be displayed according to different categories e.g., the research themes of the LTFEs, land use or farming category. In addition, the overview information details about every single LTFE are provided in a pop-up window. Thus, information of a long-term fertilizer experiment is called V140 which can be located in the map and its data is used for this thesis.

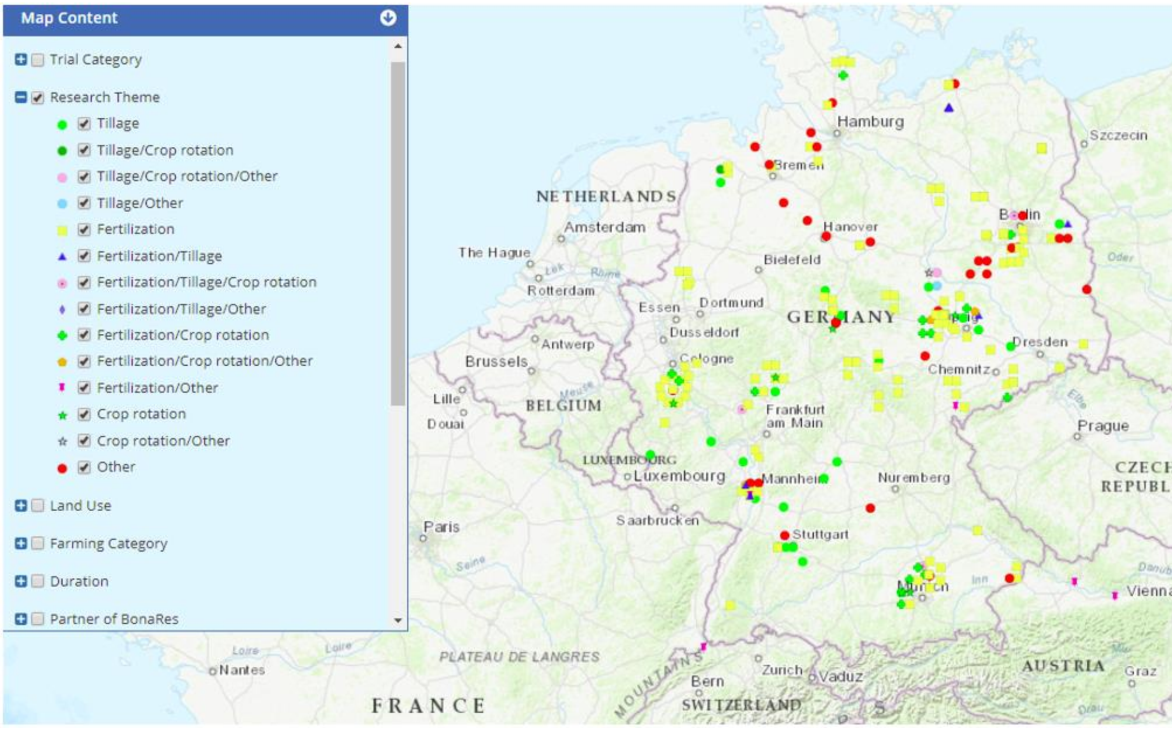



**Figure 1.1.** Map of Long-term field experiment in Germany

### 1.2.2. Description of the LTFE "V140"study site

V140 is an agricultural long-term field experiment which was established in 1963 by the German Academy of Agriculture Institute of Agriculture and Crop Production in Müncheberg, Germany (latitude: 52° 30' N; longitude: 14° 8' E; altitude: 62 m a.s.l.). The present LTFE location is at the experimental station of the Leibniz Centre for Agricultural Landscape Research (ZALF), about 50 km east of Berlin in the district Märkisch-Oderland

(Figure 1.2). The area is characterized by dry periods, particularly during early summer (Barkusky, 2018).

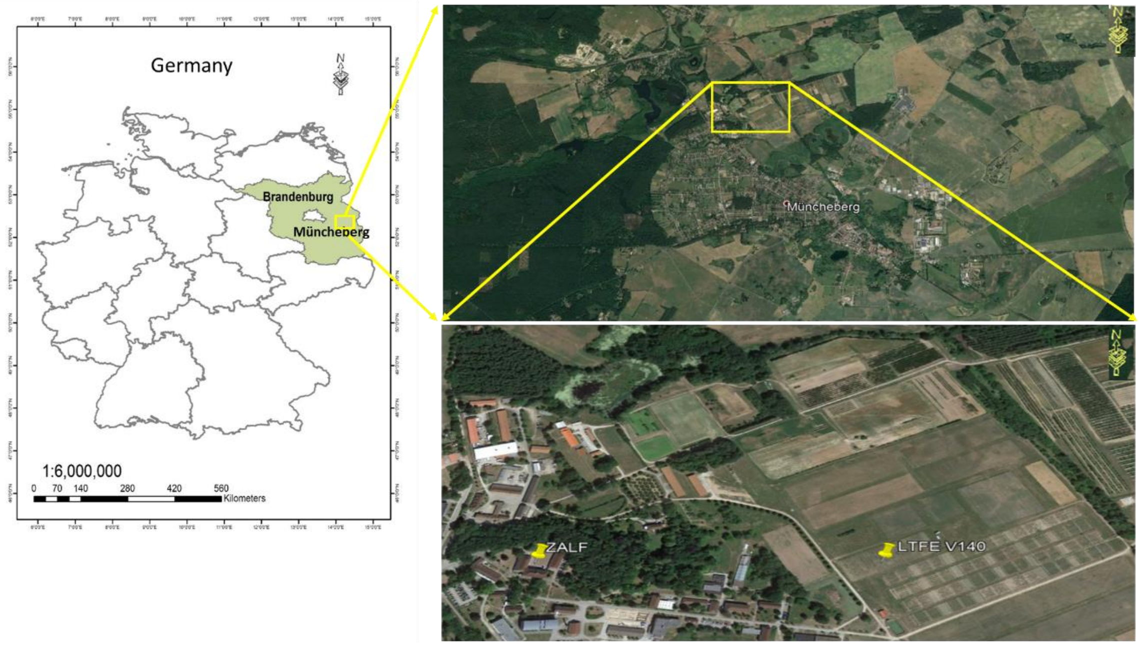



**Figure 1.2.** Map locations of the LTFE, "V140" in Müncheberg, Germany (Thai, 2019)


Source: Google earth

**Figure 1.3.** Experimental design

The soil of the study site is characterised by a high sand content of 740 g kg$^{-1}$ (50 g kg$^{-1}$ clay, 210 g kg$^{-1}$ silt), low total carbon contents (4.3-5.2 g kg$^{-1}$), a CEC of 31.5-35.6 mmolc kg$^{-1}$ and a $pH_{KCl}$ of 5.4-5.9 in the plough layer (0-25 cm) (Ellerbrock et al. 1999). According to the German Guidelines for Soil Assessment (Bodenschätzung) the dominating soil types are slightly loamy sand and sand (Sl4D and S4D). The most common soil sub type is Haplic Luvisol accounting to the FAO guideline (Barkusky, 2018; FAO, 1990). During long-term experimental from 1963-2016, the mean annual precipitation and average annual air temperature were 540 mm and 8.8°C, respectively (Thai et al, 2019).

V140 represents one of few active LTFEs on sandy soil in Germany. The experiment was set up on a flat land measuring 5712 $m^2$ involving 168 individual plots (Figure 1.3). The individual plots measured 6.0 m × 5.0 m, and a buffer zone of 1 m was allowed between the blocks. The experiment was arranged in a randomized complete block design (RCBD) comprising 21 treatments with eight blocks. The treatments includes five levels of mineral fertilization, each in combination with four levels of organic fertilization and one control. Before establishment of V140 the site was cultivated uniformly. During the running period just smaller management measures changes have been done.

Historically, V140 experiment was established to answer original questions regarding yield increase measures on sandy soil. Later, it was used to reveal the effects of organic and mineral fertilization on essential aspects of the soil fertility (Barkusky, 2018). Thus, the researchers' studies have so far considered more on soil characteristics, especially for soil carbon. For instance, the studies of (Ellerbrock and Gerke, 2016; Ellerbrock et al., 1999) had the main focus on soil organic carbon matter (SOM) or composition of SOM. The studies indicated that the total amount of SOM did not discriminate between different manurial practices over 34 years (from 1963 to 1997). The chemical composition of SOM was demonstrated effectively by the type of fertilization (Ellerbrock et al., 1999) while the composition of SOM was influenced by different long-term fertilization (Ellerbrock and Gerke, 2016). Likewise, study of Rogasik et al. (2004) focused on carbon factor in the soil under long term of agro-technical condition. They found that amount of carbon stored in agricultural soils depends on site specific climatic conditions and land use type as well as on management decisions. LTFEs were a useful tool for calculating the carbon sink or source potential of arable soils. Furthermore, the valuable data from the LTFE was also delivered for validation of models used to evaluate soil functions (Ellerbrock et al., 2005)

However, the researches related to inter-annual variability of crop yields in this experiment (i.e., yield stability) that can be associated with climate variability or changes of other factors have not been done. For instance, the study on the effects of different fertilization regimes on grain yield of different cereals over a long-term scale has not been evaluated. Therefore, this study aimed to assess the effects of different fertilization regimes on grain yield of one spring cereal (spring barley) and two winter cereals (winter rye and winter wheat) under varying annual weather condition in the LTFE.

### 1.3. Cereal crop

Cereal crops which are usually called grain crops are the top of the list for human nutritional needs. All cereals belong to the Gramineae family which has nine species: wheat (*Triticum*), rye (*Secale*), baley (*Hordeum*), oat (*Avena*), rice (*Oryza*), millet (*Pennisetum*), corn (*Zea*), sorghum (*Sorghum*), and triticale, which is a hybrid of wheat and rye.

Cereals provide essential nutrients and energy in the daily human diet through direct human consumption as food and indirect through using products such as meat which comprise a major livestock feed. Yielding grains are used for food, feed, and industrial purposes such as production of alcohol. Cereals are grown in greater quantities worldwide than any other type of crop and provide more food energy to human race than any other crop (Galanakis, 2018). According to the Food and Agriculture Organization (FAO) in 2020, cereal crops have a global cropping area of almost 740 million ha, with total cereals production of around 2982 million tons they supplied approximately 50% of the world's caloric intake (FAO, 2020).

Among cereals, wheat, rye, and barley are important cool-season cereals. In 2020, total production of wheat, rye, and barley in the world accounted for over 31% of the total global cereal production (FAO, 2020). In Europe scale, these cereals are the top three cool-season cereals, counting around 70 % of the total cereal production in 2020. In addition, the production of wheat, rye, and barley covered for more than 84% of the total cereal production in Germany. Among them, wheat leads in production, followed by barley while rye is the third most important cool temperate cereal. On the other hand, wheat and rye use is almost equal for animal feed and human consumption, while barley is predominantly used for animal feed.

Wheat, rye, and barley are the temperate species so they grow well in moderate weather and cease to grow in hot weather (approximately 30°C but this varies by species and variety). Therefore, they are grown mostly in the temperate regions or in some tropical regions during their cool season. Contrasted to tropical species, the temperate species can be grouped into spring and winter types. Therefore, most varieties of a particular species are either winter or spring types.

In this study, we considered yield variation of one spring cereal (spring barley) and two winter cereals (winter rye and winter wheat) which were cultivated in the LTFE. The aim of this investigation was to assess the effect of fertilization on grain yield of spring barley, winter rye and winter wheat over long period.

### 1.3.1. Spring cereals

Spring cereals are planted in early springtime and mature later during that same summer, without vernalization (exposure to low temperature for a genetically determined length of time). Spring cereals typically require more irrigation and yield less than winter cereals. The spring types will head quite normally when planted in the spring in the more northerly latitudes, or when grown during winter in the tropical regions. Spring planting of spring varieties that are adapted to a short season (90 days or less) can result in an adequate crop given the long summer day length of the more northerly regions. Spring barley, oats and wheat are the three main spring cereal crops in the world of which spring barley is a leader.

Spring barley (*Hordeum vulgare* L.) (SB) is the most important spring cereal in Germany. Reuters reported in 2017 that there is approximately 524,300 ha of spring grains including spring wheat, SB and oats sown in Germany. The total area for SB cultivation was approximately 400,000 ha (Hogan, 2017). Moreover, in 2020 the total grown area for barley was approximately 1.67 million ha (FAO, 2020), of which the total area for SB cultivation was approximately 346,000 ha (German Report, 2020).

Small area is cultivated for SB than for winter barley in Germany, however, SB is an important crop in the crop rotation and is used not only as animal feed but also for malting and roasting. In northeastern Germany, the location of the experimental site, SB has a short growth duration of approximately four months, normally from the end of March or the beginning of April to the end of July or early August. Compared to winter cereal, a relatively short period of weather influences the formation and differentiation of the yield characteristics (Chmielewski and Kohn, 1999).

### 1.3.2. Winter cereals

Winter cereals or winter grains are biennial cereal crops that are sown in the autumn, germinate and grow vegetatively before the freezing temperatures of winter set in, then become dormant during winter (winter rest). They resume growing in the springtime and mature in late spring or early summer.

Different with spring varieties, winter varieties requite vernalization, when crop exposure to low temperature for a genetically determined length of time. Therefore, winter cereals do not flower until springtime. Also, once these varieties are cultivated in the spring season, they do not normally head, flower and produce seed in the same season. The over-wintering period (this phase is called vernalization in scientific jargon) is thus necessary for the successful completion of their life cycle.

The general advantage of winter varieties is that the phenological stages such as tillering, shooting, and flowering appear earlier in the year during moderate temperatures, and plant development in general runs slower. However, they are harvested earlier than grains of the same type sown in springtime (Taylor and Cormack, 2002). Winter cereals generally have a much higher yield than their spring cousins and more stable grain yield (Olesen et al., 2000). This is partly a consequence of becoming established in the soil during the fall so that they are ready to begin growth as soon as the temperatures start to rise in early spring, while partly they can use winter moisture for growth. At that time the soil is generally very wet, making it impossible for the cultivation and seeding of spring types. Associated with the early spring growth is early maturity so that winter types not only escape the damaging effects of drought in late summer, but are usually harvested before disease has built up to severely damaging proportions.

In the Europe in general and in Germany in particular wheat, rye, barley and triticale are typically winter cereals. As reported by German statistics agency Destatis in 2021, there is approximately 5.1 million ha of winter grain. Whereas, the area of winter wheat is a big part with approximately 2.9 million ha, roughly 1.2 million hectares to grow winter barley, the area sown for rye and mixed winter cereals with approximately 593,300 ha, and triticale approximately 323,900 ha (Destatis, 2021). Thus, wheat, barley and rye are common winter varieties in Germany. In the LTFE, "V140", rye and wheat were cultivated for winter varieties, while barley was cultivated for spring variety.

Rye (*Secale cereale* L.) is an important temperate cereal in many countries. Rye production area was approximately 4.4 million ha, about 15 million tons production of grain and an average yield of 3.3 t $ha^{-1}$ worldwide in 2020 (FAO, 2020). Rye is almost exclusively cultivated as a winter crop. The cold tolerance and winter hardness contribute to its wide distribution in Central and Eastern Europe, where winter rye (WR) is cultivated on 3.6 million ha, which produce more than 75% of global rye production in 2020 (FAO, 2020). Germany is the leading producers of rye in Europe with a global production share of 23.4 % (FAO, 2020). In 2020, Germany recorded the highest rye production with approximately 3.5 million tons and its average grain yield at 5.5 t $ha^{-1}$.

Rye is primarily used for the production of bread flour, but also can be used for animal feed, alcohol, and biogas production. Rye is recognized to be the most drought tolerant cereal crop because of its extensive and well branched root system, which takes up water very efficiently (Starzycki and Bushuk, 1976). The root dry weight of rye exceeds that of wheat and triticale (Sheng and Hunt, 1991).

Despite WR is the most winter hardy and relatively good drought-resistant crop than all small grains and more productive than other cereals when grown on the same soil, which is on sandy, infertile, acid soil, poorly prepared land as well as on light soil with low water-holding capacity (Schittenhelm et al., 2014; Starzycki and Bushuk, 1976). Particularly, it is the best-adapted cereal crop on sandy soil. Therefore, WR is predominantly cultivated in marginal locations with low fertility, in which other cereals can hardly be grown (Miedaner et al., 2012). This makes WR especially vulnerable to drought events (Schlegel, 2013). Similar to other crops, grain yield of WR is affected not only by management practice but also by extremely weather such as drought, frost or unfavorable weather condition (Wittchen and Chmielewski, 2005). However, limited studies have been conducted to understand the effects of fertilizer management under such extreme weather conditions on its grain yield over long scale cultivation.

*** Winter wheat (WW)**

On the other hand, wheat (*Triticum aestivum* L.) is one of the oldest of all temperate cultivated plants and nutritionally important cereal worldwide. Currently, it has more than 50,000 varieties which are grown in a relatively wide range of climatic conditions. However, wheat is growing best in temperate climates. Therefore, in the Europe, wheat holds a unique place, which mean is the most popular cereal, covering almost half of the Europe's arable land. Winter wheat (WW) varieties are more commonly grown than spring wheat varieties.

In Germany, WW is sown on over 90% of the wheat cultivation areas, covering around 3.2 million hectares which accounts for around one-third of the total arable land area. The average total WW production from 2014 to 2018 was 24.7 million tons, with an average yield of 7.7 t $ha^{-1}$ (Destatis, 2020). The grain yield of WW in Germany has increased in recent decades from an average of less than 3 t $ha^{-1}$ in the 1960s to around 8 t $ha^{-1}$ in the 2000s (FAO, 2020). However, the grain yield of WW has fluctuated in recent years. Apart from crop breeding improvement, which has contributed dramatically to the wheat yield increase throughout the 20th century in Germany (Ahrends et al., 2018; Laidig et al., 2017), several other factors, such as enhanced agronomic management, favorable weather conditions, fertilization and soil improvement, also played an important role in yield development and yield stability (Macholdt and Honermeier, 2018; Macholdt et al., 2019). Thus, similar to other crops, yield variation in WW is the result of interdependencies and complex interactions among different factors. In this regard, identifying the major factors and their relationships that account for grain yield variation of WW is crucial to understanding how to maximize yields and minimize annual yield fluctuations each year. To achieve such investigations clearly it requires figuring out suitable analytical technicalities.

### 1.4. Thesis aim and objectives

The thesis aim to assess the cereal grain yield responses to fertilizer management in sandy soil in a long-term (1971 to 2016) fertilizer experiment in Northeast Germany.

To achieve this aim, three specific objectives are set to answer research questions in section 1.5 as follows:

(i) Analyze cereal grain yield responses to fertilizer management in the LTFE.

(ii) Analyze sensitive timing of weather events for cereal grain yield in the LTFE.

(iii) Compare different analysis models relevant for analyzing the grain yield in the LTFE.

### 1.5. Research questions and hypothesis

I pose and answer three specific questions as an attempt to address the main aim of the thesis:

(i) Does the cereal grain yield responses to fertilizer management in long-term field experiment in comprises relationships among climatic dependence and other factors?

(ii) Are there sensitive timing of weather events that determine/influence cereal grain yield in the LTFE?

(iii) What are the benefits of using different analysis models to analyze cereal grain yield in the LTFE?

I hypothesize that grain yield response of different cereals to fertilizer management in the long term involves complex relationships among climatic dependence, crop rotation, and soil characteristics. We postulate more challenges in assessing the yield of winter cereal than that of spring cereal

### 1.6. Conceptual framework

Figure 1.4 shows the conceptual framework of the study for cereals' grain yield response to fertilizer management in sandy soil in a long-term fertilizer field experiment in Northeast Germany. The main exploration of this study is based on analyzing input data collected as a part of data from LTFE, "V140" in Müncheberg, Germany. The input data focused on different factors i. e. fertilizer regimes (treatments), annual weather conditions, preceding crop, and soil, which may have temporal variability and affect cereal plant growth and the grain yield. The experimental factor is fertilizer including 21 treatments: five levels of mineral nitrogen fertilization, each combined with four levels of organic fertilization and one control. Non-experimental factors considered were weather, preceding crop, and soil data. For weather data, the work has focused on two important climatic variables: temperature and precipitation. For the preceding crop factor, the work has considered the preceding crop type, preceding crop yield.

In order to assess the cereal grain yield response to fertilizer in the long-term, the study addressed the specific objectives in the conceptual framework. Input data of three cereal crops namely spring barley, winter rye, and winter wheat were employed. The study used statistical analysis and machine learning (ML) methods. The statistical comprise Analysis of variance (ANOVA), General Linear Model (GLM), Multiple linear regression model (MRM), and Linear Mixed Model (LMM). The ML model employed was the the M5P model. Statistical analysis methods such as the ANOVA/GLM, and MRM (BMA) were used in the paper 1 (chapter 2) while the ANOVA/GLM, and M5P machine learning algorithm were used in the paper 2 (chapter 3). In paper 3 (chapter 4), the study used ANOVA, LMM, and M5P models. Therefore, this thesis contributes three scientific articles covering three specific objectives (in section 1.4) addressing the research aim.

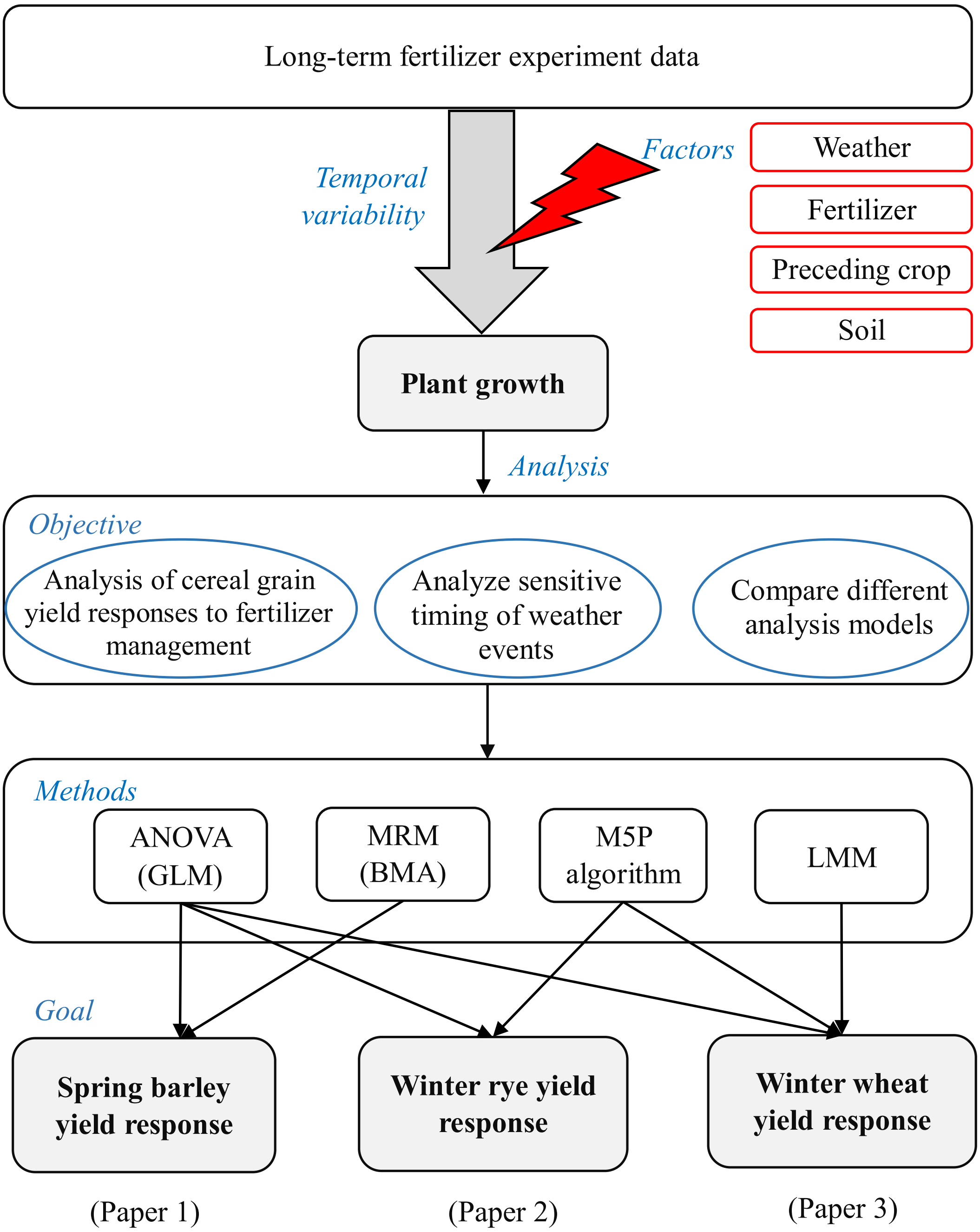



**Figure 1.4.** Conceptual framework (own design)

In Figure 1.4, the abbreviations are explained as following:

ANOVA: Analysis of variance

MRM: Multiple linear regression model

BMA: Bayesian model averaging

M5P: M5P machine learning algorithm/ M5P model

LMM: Linear mixed-effects model

GLM: General linear model

## 1.7. Research data and methods

### 1.7.1. Research data

This study used data collected as part of the V140 experiment in Müncheberg, Germany. The data in the LTFE used for analysis consisted of the experiment period from 1971-2016 where experimental design was stability and cereals were cultivated. The crop sequence in the experiment was not fixed and consisted of winter wheat, winter rye, spring barley, potatoes, sugar beets, maize, flax, and peas. Hence, sugar beet or potato were preceding crops for spring barley depend on growing season. Similar, maize or flax were preceding crops for winter rye, while potato, sugar beet or pea were preceding crops for winter wheat.

Dry mass (DM) grain yield data (Mg DM $ha^{-1}$) of SB, WR, WW obtained from every plot during the experiment period were used during the analysis. DM yield of the preceding crops for spring barley, winter rye, winter wheat in every treatment and every replication was obtained to estimate the effect of the preceding crop on the cereal yields.

Selected chemical soil parameters with eight replications in each treatment in available years were selected for input of model to estimate their effect on yield of winter cereal.

Meteorological data used in the analysis were obtained from an adjacent climate station of the German Meteorological Service (DWD), station number 03376 via the link: https://opendata.dwd.de/climate_environment/CDC/observations_germany/climate/daily/kl/historical). The daily mean air temperature, maximum temperature, minimum temperature and precipitation during the growing period of selected cereals were used to calculate the input weather variables for the study.

For spring cereal, the monthly mean temperature, cumulative precipitation during SB growing season, and average temperature, total precipitation during the whole SB growing season (April-July) and total precipitation from the prior winter (October-February) was used in statistical analyses.

For winter cereals: Monthly mean temperature, cumulative precipitation, cumulative number of days recorded having mean temperatures above 30°C in every month (days Tmax > 30°C), the cumulative number of days recorded with mean temperatures below 0°C or 32°F (freezing days in a month), and cumulative growing-degree days during the growing seasons were calculated. The maximum and minimum temperatures were used to calculate the growing degree days (GDD).

### 1.7.2. Methods

The present study has used a range of classical analytical and data mining techniques to unlock the complexities of factors and interdependencies influencing the yields of different cereals in a drought prone sandy soil with low nutrient inventory. Details of the methods are presented in each paper separately. However, they were summed as follows:

**In chapter 2 (paper 1),** statistical analysis methods were applied to estimate the main and interactive effects of treatments or year (annual weather conditions) on yield variation in SB using ANOVA and GLMs. Then, MRM was employed to evaluate the SB yield data as a function of weather parameters. By avoiding the collinearity effect, correlation analysis of weather variables was conducted to select the appropriate variables for the MRM. Bayesian method was used by Bayesian model averaging (BMA) for the MRM

**In chapter 3 (paper 2),** classical analytical methods were tested, then a data mining technique applied to analyze the WR grain yield. I used GLM to figure out the main factors influencing the WR yield. The yield variation due to the effects of treatment (fertilizer), yearly effects (weather conditions), and the size of their effects were estimated. For data mining techniques, I used the M5P machine learning algorithm to show up by decision tree model for the WR yield response.

**In chapter 4 (paper 3),** I compared the statistical model to the machine learning model in analyzing WW yield response. There are two main steps in the analysis. I first explored the WW grain yield and yield variability using descriptive analysis and ANOVA within the GLM and afterward applied nonparametric methods involving the LMM and M5P model to analyze the grain yield response.

## 1.8. Thesis structure and contribution of articles

This is a cumulative thesis comprising three scientific papers that together deal with three specific objectives (in section 1.4) and answer three research questions.

Chapter 2 is paper 1, investigates the grain yield response of SB to fertilizer management in a LTFE. The study provided evidence that different fertilizer regimes, weather, and their interaction have effects on the grain yield of SB. In addition, the study revealed that non-experimental factor such as the preceding crop was an important variable that could influence the yield variation of SB. Also, soil parameters should be considered in further research on the grain yield of SB. Chapter 3 is paper 2 and chapter 4 is paper 3 provide a detailed analysis of grain yield response of winter cereals to fertilizer, weather, preceding crop, and soil factors using long term data. Chapter 3 focuses on analyzing and evaluating the responses of grain yield of WR while in chapter 4, WW yield response was evaluated. Chapter 5, sum up the main outcomes, general discussion, and overall conclusions.

## References


Ahmad, S., Abbas, G., Ahmed, M., Fatima, Z., Anjum, M. A., Rasul, G., Khan, M. A., and Hoogenboom, G., 2019. Climate warming and management impact on the change of phenology of the rice-wheat cropping system in Punjab, Pakistan. Field Crops Research. 230, 46-61.

Ahrends, H. E., Eugster, W., Gaiser, T., Rueda-Ayala, V., Hüging, H., Ewert, F., and Siebert, S., 2018. Genetic yield gains of winter wheat in Germany over more than 100 years (1895-2007) under contrasting fertilizer applications. Environmental Research Letters. 13, 104003.

Barkusky, D., 2018. Long-term field experiment V140 Muencheberg from 1963 to 2009 - Plots (Version 1.0). BonaRes Repository-BonaRes Data Centre (Leibniz Centre for Agricultural Landscape Research (ZALF)).

Blanchet, G., Gavazov, K., Bragazza, L., and Sinaj, S., 2016. Responses of soil properties and crop yields to different inorganic and organic amendments in a Swiss conventional farming system. Agriculture, Ecosystems & Environment. 230, 116-126.

Brentrup, F., Küsters, J., Lammel, J., Barraclough, P., and Kuhlmann, H., 2004. Environmental impact assessment of agricultural production systems using the life cycle assessment (LCA) methodology II. The application to N fertilizer use in winter wheat production systems. European Journal of Agronomy. 20, 265-279.

Challinor, A. J., Watson, J., Lobell, D. B., Howden, S., Smith, D., and Chhetri, N., 2014. A meta-analysis of crop yield under climate change and adaptation. Nature climate change. 4, 287-291.

Chmielewski, F. M., and Kohn, W., 1999. Impact of weather on yield components of spring cereals over 30 years. Agricultural and Forest Meteorology. 96, 49-58.

Dawe, D., Dobermann, A., Moya, P., Abdulrachman, S., Singh, B., Lal, P., Li, S. Y., Lin, B., Panaullah, G., Sariam, O., Singh, Y., Swarup, A., Tan, P. S., and Zhen, Q. X., 2000. How widespread are yield declines in long-term rice experiments in Asia? Field Crops Research. 66, 175-193.

Debreczeni, K., and Körschens, M., 2003. Long-term field experiments of the world. Archives of Agronomy and Soil Science. 49, 465-483.

Destatis. Available online: https://www.destatis.de/DE/ZahlenFakten/Wirtschaftsbereiche/LandForstwirtschaftFischerei/LandwirtschaftlicheBetriebe/Tabellen/LandwirtschaftlicheBetriebeFlaechenHauptnutzungsarten.html (accessed on 20 April 2020).

Destatis. Available online: https://www.destatis.de/EN/Themes/Economic-Sectors-Enterprises/Agriculture-Forestry-Fisheries/Field-Crops-Grassland/Tables/first-estimate-of-cereal-rape-yields.html (accessed on 30 March 2022).

DWD, S. Der KLIMA-Report 2007. Available at http://www.dwd.de (accessed on 8 October 2020)

Ellerbrock, R. H., and Gerke, H. H., 2016. Analyzing Management-Induced Dynamics of Soluble Organic Matter Composition in Soils from Long-Term Field Experiments. Vadose Zone Journal. 4, 15.

Ellerbrock, R. H., Hohn, A., and Rogasik, J., 1999. Functional analysis of soil organic matter as affected by long-term manurial treatment. European Journal of Soil Science. 50, 65-71.

Ellerbrock, R. H., Kersebaum, K. C., and Kaiser, M., 2005. Isolation and characterization of soil organic matter fractions different in solubility as a possibility to evaluate and to improve C-pools in C-turnover models. Archives of Agronomy and Soil Science. 51, 209-219.

FAO, 1990. Guidelines for Soil Profile Description, 3rd edn. Food and Agriculture Organization of the United Nations, Rome.

FAO, 2022. FAOSTAT online database, available at link http://faostat.fao.org/. Accessed in March 2022.

Galanakis, C. M., 2018. Sustainable recovery and reutilization of cereal processing by-products. Woodhead Publishing.

Gourdji, S. M., Sibley, A. M., and Lobell, D. B., 2013. Global crop exposure to critical high temperatures in the reproductive period: historical trends and future projections. 8, 024041.

Grosse, M., Ahlborn, M. C., and Hierold, W., 2021. Metadata of agricultural long-term experiments in Europe exclusive of Germany. 38, 107322.

Grosse, M., Heinrich, U., and Hierold, W., 2019. Fact Sheet for the Description of Long-Term Field Experiments. BonaRes Series 2019/2. Editor: BonaRes Centre for Soil Research.

Grosse, M., and Hierold, W., 2019. Online overview and data-Long-term Field Experiments in Germany. BonaRes Data Centre ZALF.

Grosse, M., Hierold, W., Ahlborn, M. C., and Helming, K., 2020. Long-term Field Experiments in Germany: Classification and spatial Representation. SOIL Discussions, 1-32.

Hartmann, J., West, A. J., Renforth, P., Köhler, P., De La Rocha, C. L., Wolf-Gladrow, D. A., Dürr, H. H., and Scheffran, J., 2013. Enhanced chemical weathering as a geoengineering strategy to reduce atmospheric carbon dioxide, supply nutrients, and mitigate ocean acidification. Reviews of Geophysics. 51, 113-149.

Hogan, M., 2017. German farmers plant more winter wheat, less rapeseed. Reuters report.

Johnston, A. E, and Poulton, P. R., 2018. The importance of long-term experiments in agriculture: their management to ensure continued crop production and soil fertility; the Rothamsted experience. European journal of soil science. 69, 113-125.

Kirchmann, H., 2007. Success Stories of Agricultural Long-term Experiments. Report from a conference at the Royal Swedish Academy of Agriculture and Forestry, 28-29 May 2007.

Körschens, M., 2006. The importance of long-term field experiments for soil science and environmental research - A review. Plant, Soil and Environment. 52, 1-8.

Körschens, M., Albert, E., Armbruster, M., Barkusky, D., Baumecker, M., Behle-Schalk, L., Bischoff, R., Čergan, Z., Ellmer, F., Herbst, F., Hoffmann, S., Hofmann, B., Kismanyoky, T., Kubat, J., Kunzova, E., Lopez-Fando, C., Merbach, I., Merbach, W., Pardor, M. T., Rogasik, J., Rühlmann, J., Spiegel, H., Schulz, E., Tajnsek, A., Toth, Z., Wegener, H., and Zorn, W., 2013. Effect of mineral and organic fertilization on crop yield, nitrogen uptake, carbon and nitrogen balances, as well as soil organic carbon content and dynamics: results from 20 European long-term field experiments of the twenty-first century. Archives of Agronomy and Soil Science. 59, 1017-1040.

Laidig, F., Piepho, H.-P., Rentel, D., Drobek, T., Meyer, U., and Huesken, A., 2017. Breeding progress, variation, and correlation of grain and quality traits in winter rye hybrid and population varieties and national on-farm progress in Germany over 26 years. Theoretical and applied genetics. 130, 981-998.

Macholdt, J., and Honermeier, B., 2018. Stability analysis for grain yield of winter wheat in a long-term field experiment. Archives of Agronomy and Soil Science. 65.

Macholdt, J., Piepho, H.-P., and Honermeier, B., 2019. Mineral NPK and manure fertilization affecting the yield stability of winter wheat: Results from a long-term field experiment. European Journal of Agronomy. 102, 14-22.

Merbach, W., and Deubel, A., 2007. "The Long-Term Fertilization Trials in Halle (Saale): A Tool for Sustainable and Environmentally Compatible Land Management-Contributions of Plant Nutrition Science in Halle 13," Springer Science & Business Media.

Merbach, W., and Deubel, A., 2008. Long-term field experiments - museum relics or scientific challenge? Plant Soil and Environment. 54, 219-226.

Merbach, W., Herbst, F., Eißner, H., Schmidt, L., and Deubel, A., 2013. Influence of different long-term mineral-organic fertilization on yield, nutrient balance and soil C and N contents of a sandy loess (Haplic Phaeozem) in middle Germany. Archives of Agronomy and Soil Science. 59, 1059-1071.

Miedaner, T., Koch, S., Seggl, A., Schmiedchen, B., and Wilde, P., 2012. Quantitative genetic parameters for selection of biomass yield in hybrid rye. Plant breeding. 131, 100-103.

Olesen, J. E., Jensen, T., and Petersen, J., 2000. Sensitivity of field-scale winter wheat production in Denmark to climate variability and climate change. Climate Research. 15, 221-238.

Rasmussen, P. E., Goulding, K. W., Brown, J. R., Grace, P. R., Janzen, H. H., and Körschens, M., 1998. Long-term agroecosystem experiments: assessing agricultural sustainability and global change. Science. 282, 893-896.

Report, G., 2020. Crop Report N1: German spring barley. Available at link: https://www.braugerstengemeinschaft.de/en/press-report/crop-report-1-german-spring-barley-april-23-2021/

Richter, D. D., Hofmockel, M., Callaham, M. A., Powlson, D. S., and Smith, P., 2007. Long-term soil experiments: Keys to managing Earth's rapidly changing ecosystems. Soil Science Society of America Journal. 71, 266-279.

Rogasik, J., Schroetter, S., Funder, U., Schnug, E., and Kurtinecz, P., 2004. Long-term fertilizer experiments as a data base for calculating the carbon sink potential of arable soils. Archives of Agronomy and Soil Science. 50, 11-19.

Rosenzweig, C., Elliott, J., Deryng, D., Ruane, A. C., Müller, C., Arneth, A., Boote, K. J., Folberth, C., Glotter, M., and Khabarov, N., 2014. Assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison. Proceedings of the national academy of sciences. 111, 3268-3273.

Schittenhelm, S., Kraft, M., and Wittich, K.-P., 2014. Performance of winter cereals grown on field-stored soil moisture only. European journal of agronomy. 52, 247-258.

Schlegel, R. H., 2013. "Rye: genetics, breeding, and cultivation," Crc Press.

Schmidhuber, J., and Tubiello, F., 2007. Global food security under climate change. Proceedings of the National Academy of Sciences. 104, 19703-19708.

Sheng, Q., and Hunt, L., 1991. Shoot and root dry weight and soil water in wheat, triticale and rye. Canadian Journal of Plant Science. 71, 41-49.

Starzycki, S., and Bushuk, W., 1976. Diseases, pests, and physiology of rye. Rye: Production, Chemistry, and Technology. St. Paul, MN: American Association of Cereal Chemists, 27-61.

Statistisches, B., 2019. GENESIS-Online Datenbank. Available at: https://wwwgenesis.destatis.de/genesis/online [last accessed 09.8.19].

Sultana, H., Ali, N., Iqbal, M. M., and Khan, A., 2009. Vulnerability and adaptability of wheat production in different climatic zones of Pakistan under climate change scenarios. Climatic Change. 94, 123-142.

Taylor, B., and Cormack, W., 2002. Choice of cereal and pulse species and varities. In "Organic cereals and pulses. Papers presented at conferences held at the Heriot-Watt University, Edinburgh, and at Cranfield University Silsoe Campus, Bedfordshire, 6 and 9 November 2001", pp. 9-28. Chalcombe Publications.

Thai, T. H., 2019. Map of V140 long-term experimental locations in Müncheberg, Germany. Own study.

Thai, T. H., Bellingrath-Kimura, S. D., Hoffmann, C., and Barkusky, D., 2019. Effect of long-term fertiliser regimes and weather on spring barley yields in sandy soil in North-East Germany. Archives of Agronomy and Soil Science. 66, 1812-1826.

Vitousek, P. M., Aber, J. D., Howarth, R. W., Likens, G. E., Matson, P. A., Schindler, D. W., Schlesinger, W. H., and Tilman, D., 1997. Human alteration of the global nitrogen cycle: sources and consequences. Ecological applications. 7, 737-750.

Willocquet, L., Aubertot, J., Lebard, S., Robert, C., Lannou, C., and Savary, S., 2008. Simulating multiple pest damage in varying winter wheat production situations. Field Crops Research. 107, 12-28.

Wittchen, U., and Chmielewski, F. M., 2005. Phytoclimate of winter rye stands. Meteorologische Zeitschrift. 14, 183-189.

Ziervogel, G., Nyong, A., Osman, B., Conde, C., Cortés, S., Downing, T., 2006. Climate variability and change: Implications for household food security. AIACC Working Paper No. 20.

Zoellick, R. B., 2009. Securing development.

# Chapter 2

## Effect of long-term fertilizer regimes and weather on spring barley yields in sandy soil in North-East Germany

Thi Huyen Thai, Sonoko Dorothea Bellingrath-Kimura, Carsten Hoffmann, and Dietmar Barkusky

*Archives of Agronomy and Soil Science (published. DOI: 10.1080/03650340.2019.1697436)*

# Effect of long-term fertilizer regimes and weather on spring barley yields in sandy soil in North-East Germany

Thi Huyen Thai[a,b,c], Sonoko Dorothea Bellingrath-Kimura[a,b*], Carsten Hoffmann[a], and Dietmar Barkusky[a]

[a]Leibniz Centre for Agricultural Landscape Research (ZALF), Eberswalder Str. 84, 15374 Müncheberg, Germany

[b]Institute of Agriculture and Horticulture, Faculty of Life Science, Humboldt University of Berlin, Albrecht-Thaer-Weg 5, 14195 Berlin, Germany

[c]Faculty of Agronomy, University of Agriculture and Forestry (HUAF), Hue University, 102 Phung Hung street, Hue, Vietnam

*Correspondence: belks@zalf.de; Tel.: +49(0) 33432 82-207

## Abstract

The interaction effects of different fertilization regimes and weather variability on crop yield is a challenge that requires long-term investigation. Therefore, yield data for spring barley (SB) in an agricultural long-term field experiment, established in 1963 in Müncheberg, northeast Germany, were analyzed to reveal the effects of 21 fertilizer regimes and different weather conditions on SB yields. SB yields were significantly affected by fertilization regimes (11%), annual weather conditions (55%) and their interaction effect (8%). Mineral N fertilization decreased overall yield variability across seasons as compared to no fertilization and organic fertilization regimes showed higher yield variability. A suitable combined application of mineral nitrogen and organic fertilizer was found to be an effective way to produce higher SB yields than the application of either mineral nitrogen or organic fertilizer alone. A Bayesian linear regression model showed total precipitation during the growing season (April-July) positively affected on SB yields when high mineral N was supplied. At the early growth stage, a precipitation rate (March) and temperature (April or sowing day) negatively affected on SB yield.

**Keywords:** barley, mineral N fertilizer, organic fertilizer, precipitation, temperature.

## 2.1. Introduction

Barley (*Hordeum vulgare* L.) is one of the most important cereals after wheat, maize and rice, grown in more than 100 countries in the world (FAO 2019). Based on statistical evaluations of the Food and Agricultural Organization (2019), the total production area was approximately 47 million hectares (ha) with a production of approximately 149 million tons (t) of grain at an average yield of 3.2 t ha$^{-1}$ in 2017. In Europe, barley is the second most important cool temperature cereal after wheat. In the last decade, Europe has produced about 60% of global production despite the decline in the production area and grain production (FAO 2019). The major European barley-producing countries are France, Germany, Russia, Ukraine, and Spain. In Germany, the total growing area for barley was approximately 1.6 million ha in 2016, of which the total area for spring barley (SB) cultivation was approximately 340,000 ha (Walter 2017; FAO 2019). Although less area is cultivated for SB than for winter barley, SB is an important crop in the crop rotation and is used not only as animal feed but also for malt production (Friedt and Ordon 2013). In northeastern Germany, SB has a short growth duration of approximately four months and is usually sown at the end of March or in early April and harvested at the end of July.

In general, food security under a globally changing climate requires a comprehensive understanding of fertilization practices to achieve optimal crop yields while minimizing the harmful effects on agroecosystems (Timsina 2018). Fertilizer management is considered an important factor in agricultural production for sustaining crop productivity in agroecosystems (Blanchet et al. 2016). Integrating mineral nitrogen fertilizer with organic fertilizer is a promising management strategy for sustainable agricultural production systems, especially for fields with low soil organic matter contents and dry condition (Yang et al. 2015; Wei et al. 2016; Muller et al. 2017). The effect of integrated organic-mineral fertilization on SB yield was observed in several studies; the trails, however, show that the strength of the effect may differ depending on soil condition, weather condition in a year, and agronomical factors (Příkopa et al. 2005; Váňová et al. 2006; Černý et al. 2010).

Temperature and precipitation are two major climatic factors that determine crop yields (Peltonen-Sainio et al. 2011). They are important predictors of yield at sensitive crop phenological stages (Peltonen-Sainio et al. 2010; Trnka et al. 2011). Studies by Lobell and Field (2007) found that seasonal temperature and precipitation explained 30% or more of the year-to-year variation in global average yield for the six most commonly grown crops. Other studies in the UK found that 33% of the variation in grain yield and 50% of the variation in straw yield of winter wheat could be explained by precipitation and temperature variation (Chmielewski and Potts 1995). Chmielewski and Köhn (1999) reported that it is possible to explain nearly 60% of grain yield variability in SB and oats with meteorological variables. Thus, detailed observations of weather variables can help to explain yield variabilities for each crop. Furthermore, Freckleton et al. (1999) and Fisher (1925) indicated that there may be significant effects of weather on crop yield and yield variability. However, both studies showed that the directions of these effects were not necessarily consistent, and these effects may interact with nutrient input and the environmental conditions of the study site.

We postulate that using long-term datasets, it is a better way to understand effects of the two covariates. To obtain effects of fertilizer regimes and weather conditions on SB yields, the data from the "V140" agricultural long-term field experiment in Müncheberg, northeastern Germany for nine growing seasons over the period from 1976-2016 were analyzed. SB was chosen as the sample crop since which shows the highest yield variability and is highly sensitive to weather in spring period (Trnka et al. 2007). We expect that these effects appear more significant at sandy soils with low nutrient inventory and low annual precipitation rates.

Therefore, the objective of this study is to assess the effects of different long-term fertilization regimes on SB yield under varying annual weather conditions and their interactions. The study addressed the following three research questions: 1) Which weather variables determine SB yield variation?, 2 ) How fertilizer regimes and weather affect SB yield variation?, and 3) What different fertilizer management strategies affect SB yield in the long term?

## 2.2. Materials and methods

### 2.2.1. Site description

The data were collected from the agricultural long-term field experiment (LTFE) "V140", which was established in 1963 by the German Academy of Agriculture Institute of Agriculture and Crop Production in Müncheberg, Germany (latitude: 52° 30' N; longitude: 14° 8' E; altitude: 62 m a.s.l.). The field trial is located approximately 50 km east of Berlin. The soil of the study site is characterised by a high sand content of 740 g $kg^{-1}$ (50 g $kg^{-1}$ clay, 210 g $kg^{-1}$ silt), low total carbon contents (4.3-5.2 g $kg^{-1}$), a CEC of 31.5-35.6 mmolc $kg^{-1}$ and a pHKCl of 5.4-5.9 in the plough layer (0-25 cm) (Ellerbrock et al. 1999). The soil type was classified as a Podzoluvisol to Arenosol (FAO 2006). During the experimental from 1963-2016, the mean annual precipitation and average annual air temperature were 540 mm (range: 343-793 mm) and 8.8°C (range: 6.5-10.4°C), respectively.

### 2.2.2. Experimental design

The experiment followed a randomized complete block design with 21 treatments (including a control, Table 2.1), each with eight replicates. The plot size was 30 $m^2$ (6.0 m × 5.0 m). The cropping system was conventional tillage with ploughing in autumn or in the spring depending on date of harvesting of preceding crop and weather conditions in autumn. The seedbed was prepared immediately before sowing. Different crops (winter wheat, winter rye, spring barley, potato, sugar beet, pea, maize, and oil flax) were annually cultivated in a cropping system (supplementary Table S 2.1). The present study focused on dry mass grain yield data of nine SB growing seasons in "V140" from 1976-2016. Because of lack of the eight replication of the treatment in the first 9 years (1963-1971). Since 1972, SB started to cultivate in crop rotation since 1976.

Spring barley was sown from end of March to early or mid-April (mean daily air temperature mostly > 5.0°C) and was commonly harvested between the end of July to the beginning of August depending on weather conditions. The experimental period was separated into two distinct periods: (1) six SB seasons from 1976 to the period before 2000, where farmyard manure (fym) was applied every two years to sugar beet fields (preceding crop), and (2) three SB seasons from 2000 to 2016, where fym was amended every four years to potato fields (preceding crop). The separation is necessary due to changes in crop rotations and change time for applying manure in 1999. Sugar beet used in rotation of period 1 in 1975, 1977, 1979, 1981, 1985 and 1989. Potato served as the preceding crop in rotation of period 2 in 1999, 2007, 2015. Average nutrient contents of dry mass manure used in the experiment were 2.3% N, 0.9% $P_2O_5$, 2.3 % K2O, 1.6% Mg and 55.9% organic matter. Straw was applied every two years throughout both periods (using the straw from the harvested cereal). Average nutrient contents of dry mass straw used in the experiment were 0.6% N, 0.1% $P_2O_5$, 1.5% K2O and 0.08% Mg. The ploughing, cultivation, sowing, and liming methods and seeding rate were the same for all plots. The phosphorus and potassium fertilization rates (50 kg ha$^{-1}$ $P_2O_5$ a$^{-1}$, 150 kg ha$^{-1}$ $K_2O$ a$^{-1}$) were the same for all plots (In control plot phosphorus and potassium fertilizer were applied only in the years 1978 and 1980). Mineral nitrogen fertilizer was annually applied two twice during SB growth, after seeding (the end of March or early or mid-April) and between shooting to full bloom (the end of May or early June). Mean values of soil chemical analyses of each treatment through eight growing years of SB (accept 2000 lack of data) are shown in the supplementary Table S 2.2. Soil influence on SB yield could study in another paper. The used varieties of SB changed over time: "Trumpf" variety was used in two years (1976 and 1978), different varieties were cultivated between 1980 and 2000, and "Simba" variety was used in both years 2008 and 2016. Weeds were controlled with a postemergence herbicide. SB was harvested at the time of technological maturity by plot harvester.

### 2.2.3. Description of the treatments

Five different rates of mineral nitrogen fertilization (MN) were applied with four organic fertilizer (OR) regimes: i) no OR, ii) 1.2 t dry mass (DM) ha$^{-1}$ a$^{-1}$ farmyard manure (= fym1), iii) 3.2 t DM ha$^{-1}$ a$^{-1}$ farmyard manure (= fym2), and iv) 2.0 t DM ha$^{-1}$ a$^{-1}$ straw (Table 2.1). In addition to the control (no fertilization), the five MN rates were 25, 50, 75, 100, and 125 kg ha$^{-1}$ N, which are referred to as N0, N1, N2, N3, N4, and N5, respectively. In the regime with no OR, fym1 and straw, the N1, N2, N3, N4 and N5 rates were included, while in the fym2 regime, N0, N1, N2, N3, and N4 were included, respectively. Together with the control treatment, 21 treatments were included in this experiment. Treatments were grouped as shown in Table 2.1. The straw from each crop was removed from the experimental plots after harvest.

**Table 2.1.** Description of the experimental treatments

| Treatment Code | Group treatment | Mineral nitrogen fertilizer-NF (kg ha$^{-1}$) | Organic fertilizer-ORF | Fertilizer application |
|---|---|---|---|---|
| 0 | Control | 0 | 0 | 0 |
| 1.1 | NPK | 25 | 0 | MF |
| 1.2 | | 50 | | |
| 1.3 | | 75 | | |
| 1.4 | | 100 | | |
| 1.5 | | 125 | | |
| 2.1 | NPK+fym1 | 25 | 1.2 t ha$^{-1}$ year$^{-1}$ DM farmyard manure | fym1 |
| 2.2 | | 50 | | |
| 2.3 | | 75 | | |
| 2.4 | | 100 | | |
| 2.5 | | 125 | | |
| 3.1 | PK+fym2 | 0 | 3.2 t ha$^{-1}$ year$^{-1}$ DM farmyard manure | fym2 |
| 3.2 | NPK+fym2 | 25 | | |
| 3.3 | | 50 | | |
| 3.4 | | 75 | | |
| 3.5 | | 100 | | |
| 4.1 | NPK+Straw | 25 | 2.0 t ha$^{-1}$ year$^{-1}$ DM straw | straw |
| 4.2 | | 50 | | |
| 4.3 | | 75 | | |
| 4.4 | | 100 | | |
| 4.5 | | 125 | | |

Treatment codes (1.1-1.5; 2.1-2.5; 4.1-4.5): each rate of mineral nitrogen fertilizer-NF (five levels NF: 25, 50, 75, 100, 125 kg ha$^{-1}$, respectively) with organic fertilizer-ORF (three types: no ORF, 1.2 t dry mass (DM) ha$^{-1}$ farmyard manure (FYM) and 2.0 t DM ha$^{-1}$ straw. Treatment codes (3.1-3.5): each 3.2 t DM ha$^{-1}$ FYM with each level NF (five levels: 0, 25, 50, 75, 100 kg ha$^{-1}$, respectively). Treatment code "0" or control: no fertilizer inputs. Fertilizer application (MF application: sole mineral fertilizer applied at 25, 50, 75, 100 and 125 kg ha$^{-1}$ NF; fym1: FYM applied at 25, 50, 75, 100 and 125 kg ha$^{-1}$ NF; fym2: FYM applied at 0, 25, 50, 75 and 100 kg ha$^{-1}$ NF; straw: straw applied at 35, 70, 105, 140 and 175 kg ha$^{-1}$ NF.

### 2.2.4. Meteorological and crop data

Dry mass grain yield data of SB obtained from every plot in nine years of SB cultivation from "V140" from 1976 to 2016 was used for analysis in this study. Due to the irrigation conducted in 1976, 1978 and 1980, only four replicates without irrigation were evaluated for these years.

Meteorological data used in the analysis were obtained from an adjacent climate station of the German Meteorological Service (DWD station number 03376 via the link opendata.dwd.de/climate_environment/CDC/observations_germany/climate/daily/kl/histori cal). For every year of SB cultivation, averages daily air temperature and sum daily precipitation were used to calculate the average monthly temperature, monthly precipitation during the growing season, and average temperature and total precipitation during the whole growing season (April-July) to estimate weather effects on yield and yield variability. Additionally, total precipitation from the prior winter (October-February) was used in statistical analyses.

All corresponding agricultural data of the V140 experiment, including the yield, fertilizer, plant and soil laboratory data, are open access and can be downloaded from the BonaRes Data Portal (BonaRes 2019), excluding the data from the last ten years.

**2.2.5. Statistical analysis**

Analysis of variance (ANOVA) was used to estimate SB yield variation due to the effects of treatment or year (annual weather conditions) and interaction effects between year and treatment by using a general linear model. In the case of a significant ANOVA result, Tukey's HSD post hoc test was used to assess the differences in mean yields among treatments every year and over the years. The treatment effects were declared significant at $P < 0.05$. When the SB yield data were evaluated over the years, fertilizer applications were included as fixed factors and SB planting year were included as random factors in the model. The ANOVA took into consideration the randomized complete block design of the experiment.

A multiple linear regression model (MRM) was used to evaluate the SB yield data as a function of weather parameters. To avoid the effects of collinearity, a correlation analysis of weather variables was conducted to choose the appropriate variables to be included in the MRM. Temperature and precipitation were tested using linear regression analysis, with a Pearson correlation matrix as the starting point. The proportion of significant results obtained from the matrix indicated whether a particular variable should be included in the MRM. The tested factors were considered to be statistically significant at $P < 0.05$. We used the Bayesian method for the MRM by Bayesian model averaging (BMA), following (Raftery 1995; Raftery et al. 1997; Hoeting et al. 1999). The model for yield response (yi: dependent variable) to k weather variables ($x_{1i}$, $x_{2i}$... $x_{ki}$: independent variables) has the following form (Gomez KA and Gomez AA 1983):

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_k x_{ki} + \varepsilon_i \qquad (1)$$

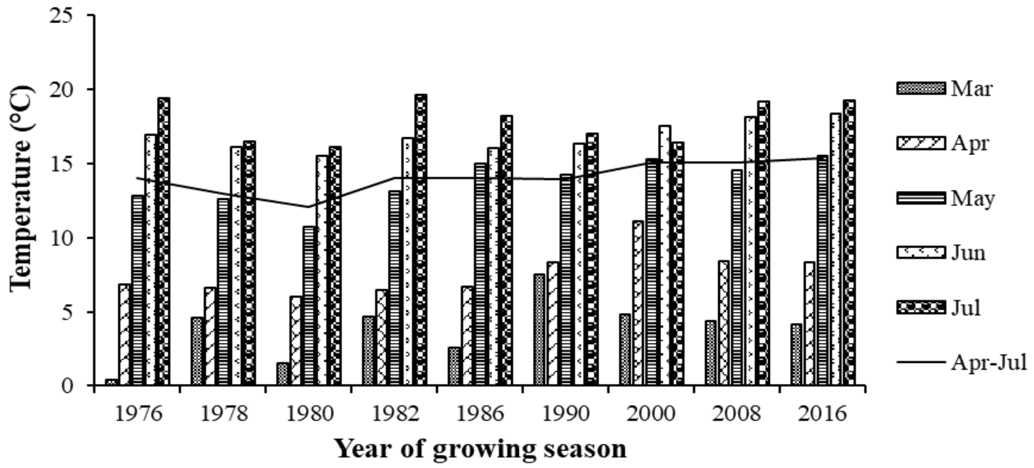
where i is the $i^{th}$ data point (i=1, …, n), yi is the yield of SB in particular treatments over nine years, $\beta_0$ is the intercept term, other weights, i.e., $\beta_1$, $\beta_2$, …, and $\beta_k$, are regression coefficients (k slope) of the k weather variables ($x_{1i}$, $x_{2i}$, …, $x_{ki}$), respectively, and εi is the error term, i.e., the residual of point i from the fitted surface. BMA usually displays the five best models found, but in this study, we report the first model since it is usually the best. The BMA is the model that includes all explanatory variables whose posterior probability (P! =0) is greater than 50%. "P! =0" is the posterior probability that the regression coefficient of each variable is non-zero (in %). The BIC (Bayesian information criterion) “is a criterion for model selection among a finite set of models; the model with the lowest BIC is preferred”. We performed all statistical analyses in SPSS version 22, R version 3.4.4 and Excel 2013.

### 2.3. Results

#### 2.3.1. Temperature and precipitation during the spring barley seasons

The lowest average monthly March temperature was 0.4°C, in 1976, while in 1990, it was highest, at 7.5°C (Figure 2.1a). The highest monthly average temperature was in July 1982, at 19.6°C. The average SB growing season temperature increased by approximately 1.4°C between 1976 and 2016. The total precipitation in each SB growing season was between 82 and 271 mm (Figure 2.1b). In six of the nine years (1976, 1978, 1982, 2000, 2008, and 2016), the largest amount of precipitation during the growing season was less than 170 mm. In most growing seasons, the month with the lowest precipitation was May. This was especially the case in 2008 when the amount of precipitation was 4.3 mm in May. The wettest month was June in most growing seasons. In June 1990, the highest amount of precipitation was recorded (165 mm)

a)




b)

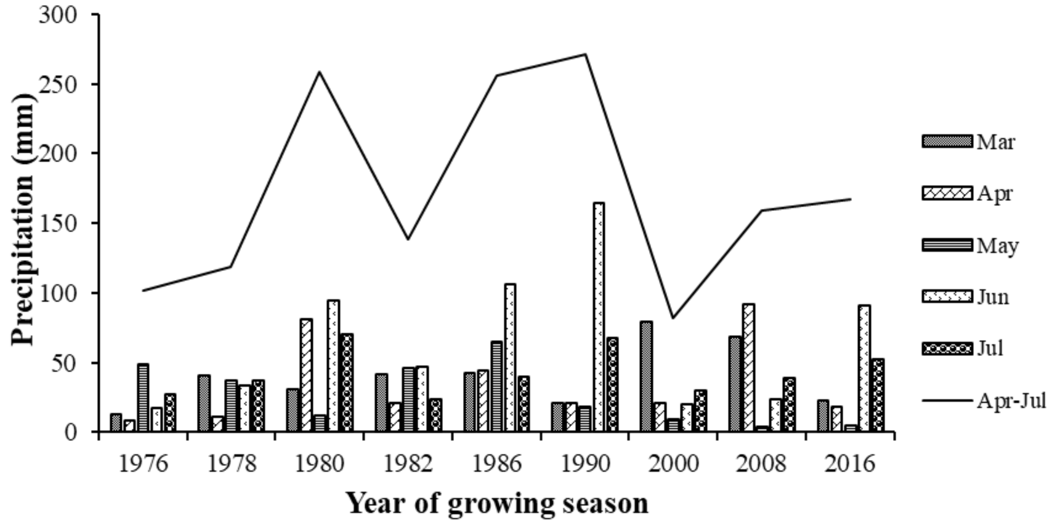



**Figure 2.1.** a) Average monthly temperature and b) total monthly precipitation during spring barley growing season in the long term experiment.

The bars show the average temperature and sum of precipitation for each month (March to July) just before sowing and during spring barley growing season; the solid line shows an average of temperature and the total amount of precipitation during the growing season (April-July).

### 2.3.2. Effect of temperature and precipitation on spring barley yield

SB yield was significantly negatively correlated with the average temperature in April in 16 of the 21 treatments; air temperature on the sowing day in 10 of the 21 treatments; temperature on the harvest day in three of the 21 treatments; average temperatures in May, July and the growing season (April-July) in one of the 21 treatments (Figure 2.2a); and the amount of precipitation in March in 15 of the 21 treatments (Figure 2.2b). The yield was significantly positively correlated with the amount of precipitation in June in one of the 21 treatments and total precipitation from April to July in seven of the 21 treatments.

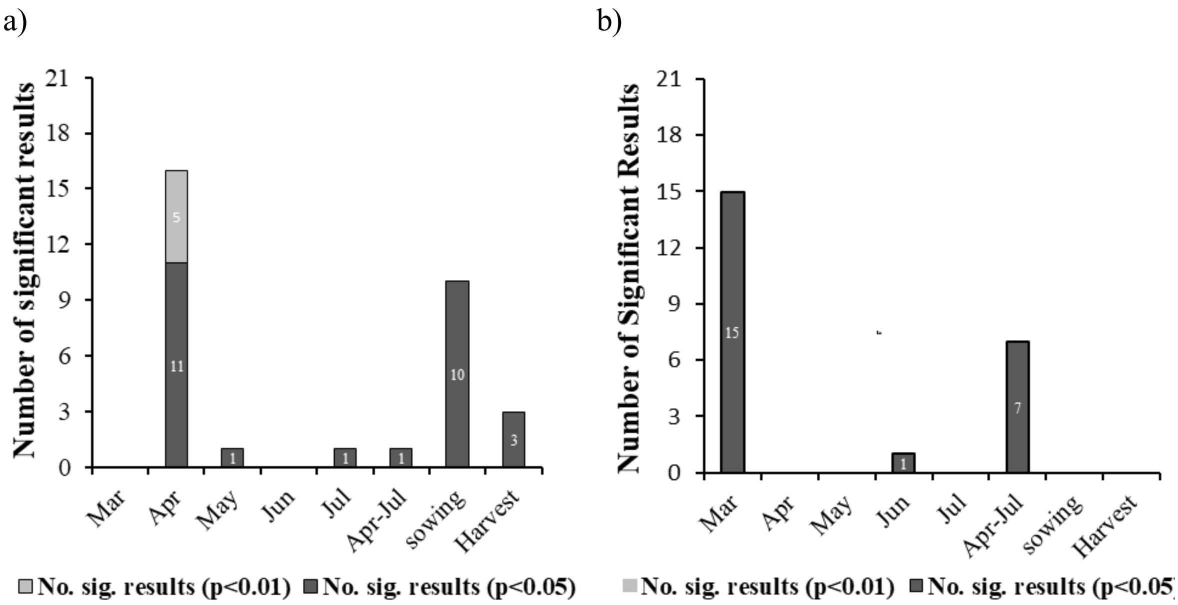



**Figure 2.2.** Number of significant results ($P< 0.05$, $P<0.01$) obtained from linear regressions of 21 mean yields from each treatment on a) average monthly temperature and b) total monthly precipitation.

The results from the multiple regression analysis by BMA presented in Table 2.2 revealed significant negative effects of the average temperature in April and sowing day temperature on SB yield. The effects were explained in 12 treatments, with $\beta_1$ values ranging from 0.101-5.639, for the temperature in April and in 16 treatments, with $\beta_2$ values ranging from 0.053-1.265, for sowing day temperature. Regarding precipitation, there were negative effects of the amount of precipitation in March on yield but positive effects of the total amount of precipitation from April-July on yield. The negative effects of the precipitation in March on yield was explained in 19 treatments, with $\beta_3$ values ranging from 0.010-1.806. The positive effects of the precipitation from April-July on yield were explained in 12 treatments except treatments low MN (treatment 1.1, 1.2, 2.1, 2.2, 3.1- 3.3, 4.1, 4.2), with $\beta_4$ values ranging from 0.004-7.194. Based on R-squared values, the weather variable effects explained 65-99% of the variation in yield among the treatments. The coefficients of variation ($Cv_s$) in SB yield among the different fertilizer treatments ranged from 0.35-0.50.

**Table 2.2.** Summary of means, variation coefficients in yields and the results of multiple regression models of yields on weather variables.

| Treatment | Yield ($Mg\ DM\ ha^{-1}$) | | Intercept ($\beta_0$) | $\beta_1$ | P!=0 (%) | $\beta_2$ | P!=0 (%) | $\beta_3$ | P!=0 (%) | $\beta_4$ | P!=0 (%) | R square |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | CV | | | | | | | | | | |
| 0 | 1.17 | 0.46 | 3.104 | - | - | -0.073 | 61.8 | -0.016 | 68.8 | -0.004 | 60.9 | 0.685 |
| 1.1 | 2.16 | 0.37 | 5.047 | -0.314 | 92.0 | - | - | - 0.012 | 57.8 | - | - | 0.752 |
| 1.2 | 2.56 | 0.37 | 5.928 | -0.264 | 90.4 | -0.089 | 80.8 | -0.015 | 71.7 | - | - | 0.871 |
| 1.3 | 2.80 | 0.40 | 5.293 | -0.260 | 80.8 | -0.076 | 62.8 | -0.016 | 64.7 | 0.004 | 67.0 | 0.866 |
| 1.4 | 2.80 | 0.38 | 4.116 | -0.149 | 76.7 | -0.099 | 100 | -0.013 | 91.7 | 0.007 | 100 | 0.957 |
| 1.5 | 2.80 | 0.40 | 3.782 | - | - | -0.116 | 100 | -0.026 | 100 | 0.006 | 96.3 | 0.923 |
| 2.1 | 2.29 | 0.36 | 3.918 | - | - | -0.103 | 67.5 | -0.019 | 58.9 | - | - | 0.647 |
| 2.2 | 2.91 | 0.36 | 6.611 | -0.291 | 90.3 | -0.086 | 73.7 | -0.019 | 78.1 | - | - | 0.872 |
| 2.3 | 2.86 | 0.38 | 4.314 | - | - | -0.123 | 98.1 | -0.027 | 98.1 | 0.004 | 63.5 | 0.886 |
| 2.4 | 2.92 | 0.39 | 3.798 | - | - | -0.145 | 100 | -0.020 | 100 | 0.006 | 100 | 0.946 |
| 2.5 | 3.29 | 0.38 | 5.600 | -0.383 | 81.6 | - | - | -0.018 | 65.1 | 0.006 | 57.6 | 0.800 |
| 3.1 | 2.26 | 0.50 | 5.778 | -0.500 | 58.2 | - | - | - | - | - | - | 0.409 |
| 3.2 | 3.01 | 0.43 | 7.126 | -0.393 | 67.6 | - | - | -0.028 | 80.3 | - | - | 0.712 |
| 3.3 | 3.13 | 0.38 | 5.683 | - | - | -0.161 | 90.7 | -0.030 | 89.0 | - | - | 0.795 |
| 3.4 | 3.09 | 0.38 | 4.434 | - | - | -0.135 | 97.6 | -0.026 | 97.4 | 0.005 | 71.6 | 0.877 |
| 3.5 | 3.07 | 0.37 | 4.148 | - | - | -0.122 | 93.8 | -0.024 | 95.5 | 0.005 | 81.5 | 0.876 |
| 4.1 | 2.62 | 0.39 | 6.927 | -5.639 | 100 | - | - | - | - | - | - | 0.765 |
| 4.2 | 2.57 | 0.38 | 5.802 | -0.242 | 75.3 | -0.096 | 67.2 | -0.015 | 56.5 | - | - | 0.782 |
| 4.3 | 3.11 | 0.35 | 5.448 | -0.330 | 100 | -0.053 | 56.6 | -0.010 | 51.9 | 0.006 | 100 | 0.933 |
| 4.4 | 3.00 | 0.37 | 4.738 | -0.101 | | -0.111 | 83.7 | -0.023 | 100 | 0.005 | 100 | 0.986 |
| 4.5 | 2.74 | 0.40 | 3.269 | - | - | -1.265 | 100 | -1.806 | 100 | 7.194 | 100 | 0.977 |

The mean (Mg dry mass $ha^{-1}$), the coefficient of variation in yields for each treatment over the years. The weights $\beta_1$, $\beta_2$, $\beta_3$, and $\beta_4$ are regression coefficients (slopes) of weather variables: April temperature ($x_1$), temperature of sowing day ($x_2$), March precipitation ($x_3$), precipitation from April-Jul ($x_4$), respectively. P!=0 is the posterior probability that each variable is non-zero (in percent), P!=0 in model select > 50%. $R^2$ values for the models.

### 2.3.3. Spring barley yield

#### *2.3.3.1. Spring barley yield under different fertilizer regimes and weather conditions*

The lowest yields were observed in the control treatments (P<0.05) (Figure 2.3) for all years except 2008, where the yield was not significantly different from the yield in fym2. The highest yields were observed in 1980 and 2016, and the lowest yields were observed in 2000. The yields in 1976, 1978, and 1986 were not significantly different. Additionally, the yields in 1982, 1990, and 2008 were not significantly different (P<0.05). There was no significant difference in SB yield among the different fertilizer treatments in most crop years. In 1980, 1990, and 2016 the yields in NPK+fym2 were significantly higher than those of all other fertilizer treatments.

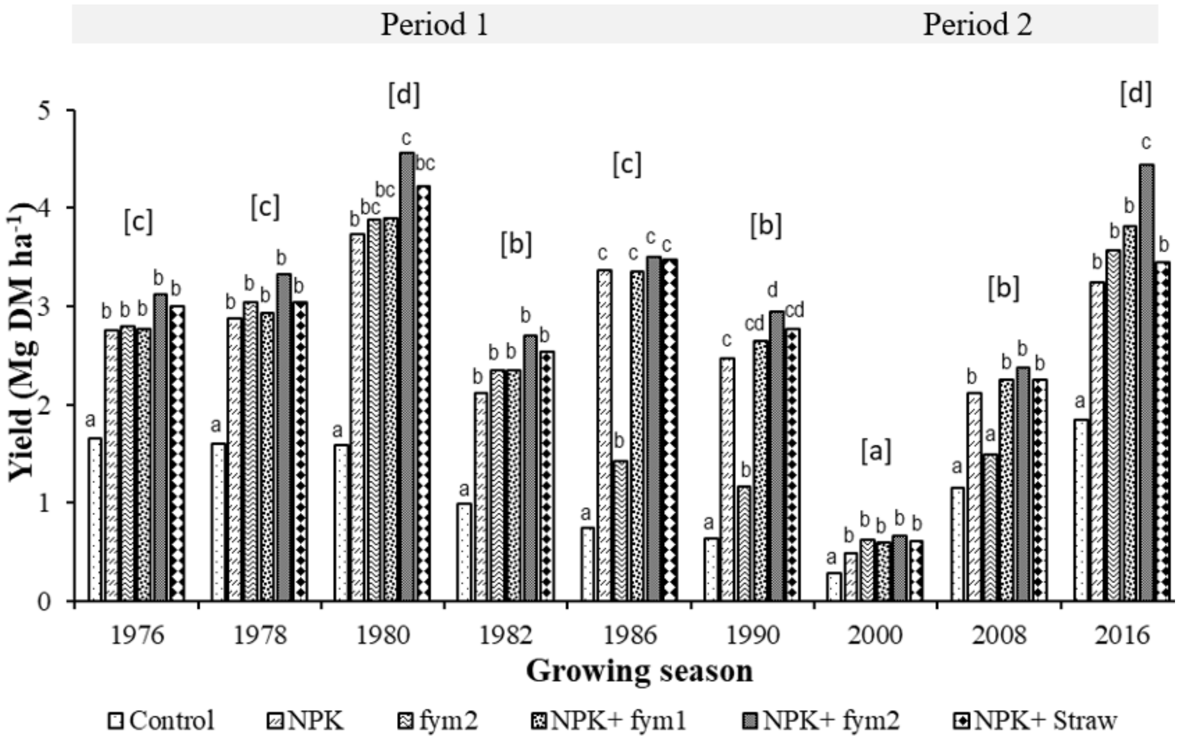



**Figure 2.3.** Effect of fertilizer applications (group treatments) on the spring barley yield (Mega gram dry mass ha$^{-1}$) every year.

Significant difference mean spring barley yield by group treatments or by average all treatments (include control) in a certain year by Tukey Test. Means sharing the same letters are not significantly different (P<0.05). Letters in square brackets at the top of bars compare mean SB yield of all treatments between different years. Letters at the top (without square brackets) of bars compare mean SB yield of group treatments within a year. Treatment groups are given in Table 2.1.

#### *2.3.3.2. The effects of fertilizer management on SB yield in the long term*

The average SB yield in the most productive treatment (NPK+fym2) was approximately 3.1 t ha$^{-1}$ a$^{-1}$ (Table 2.3). The average yield in the control treatment was 1.2

t ha$^{-1}$ a$^{-1}$, which was approximately 61% lower than that in the most productive treatment. The effect of the combined application of MN and OR (NPK+fym1, NPK+fym2 or NPK+straw) on SB yield was significantly greater than that of OR only (fym2), but only the combination of NPK+fym2 had a significantly different effect on SB yield than the treatment with MN only (NPK). The size of the positive effect of the combined application ranged between 134 and 163% compared to the control treatment, while fym2 and NPK increased SB yield by 93% and 121% compared to the control treatment, respectively. The combined application increased the SB yield by 21-36% compared to the fym2 treatment and by 6-19% compared to the NPK treatment. The coefficients of variation (Cvs) in SB yield among the different combinations of fertilizer application and NPK application were a lower value than those in the control and fym2.

**Table 2.3.** Averaged yield and variation of spring barley yields between group treatments based on the level of mineral nitrogen and organic fertilizer application in the long-term experiment through the seasons.

| Group treatment | Yield (Mg DM ha$^{-1}$) | ±Se | CV | Yield increase compare to control (%) | Yield increase compare to NPK (%) | Yield increase compare to fym2 (%) |
|---|---|---|---|---|---|---|
| Control | 1.17$^{a}$ | 0.18 | 0.46 | - | - 55 | - 48 |
| NPK | 2.58$^{c}$ | 0.32 | 0.37 | 121 | - | 14 |
| fym2 | 2.26$^{b}$ | 0.38 | 0.50 | 93 | - 12 | - |
| NPK+ fym1 | 2.74$^{c}$ | 0.33 | 0.36 | 134 | 6 | 21 |
| NPK + fym2 | 3.08$^{d}$ | 0.39 | 0.38 | 163 | 19 | 36 |
| NPK + straw | 2.83$^{c}$ | 0.34 | 0.36 | 142 | 10 | 25 |

Group treatments are given in Table 2.1. Mg DM: Mega gram dry mass; Se: standard error; CV: coefficient of variation. Different letters in the same column present that the difference was significant at P< 0.05.

Under all four fertilization applications (NPK, fym1, fym2, and straw), SB yield was significantly higher than that in the control treatment (Figure 2.4). The application of fym2 had a significant effect on SB yield at the low MN supply rates of 25-50 kg N ha$^{-1}$. While the application of fym1 and application of straw did not have a significant effect on yield at low MN supply rates of 25-50 kg N ha$^{-1}$. The highest SB yields of this study were obtained when 50 kg ha$^{-1}$ MN was applied with NPK, fym1 or 75 kg ha$^{-1}$ N was applied with straw or 25 kg ha$^{-1}$ N was applied with fym2. The result of ANOVA (sum of squares, type III) showed that barley yield variability was significantly (P<0.05) affected by year, representing annual weather conditions (55%), followed by treatment, representing fertilizer application (11%), and the year × treatment interaction (8%); 26% of the variation was due to error (other factors), and the adjusted R-squared was 0.703 (supplementary Table S 2.3).

The SB yields in the fertilizer treatment fluctuated in the first seasons (1976, 1978, 1980, 1982, 1986), steady decline in 1990 and dropped sharply in the year 2000 (in period (1)), when the yield increased rapidly for two consecutive seasons. The variability was similar in the control at a lower level, but with a much smoother gradual went down and reached the bottom in 2000, later increases quickly in SB yield (Figure 2.3, supplementary Figure S 2.1). Among other factors, the preceding crop could also have an effect on SB yields. In this experiment, the preceding crop used in rotation for the seasons before 2000 was sugar beet, while from 2000 to later seasons, potato was used as the preceding crop.

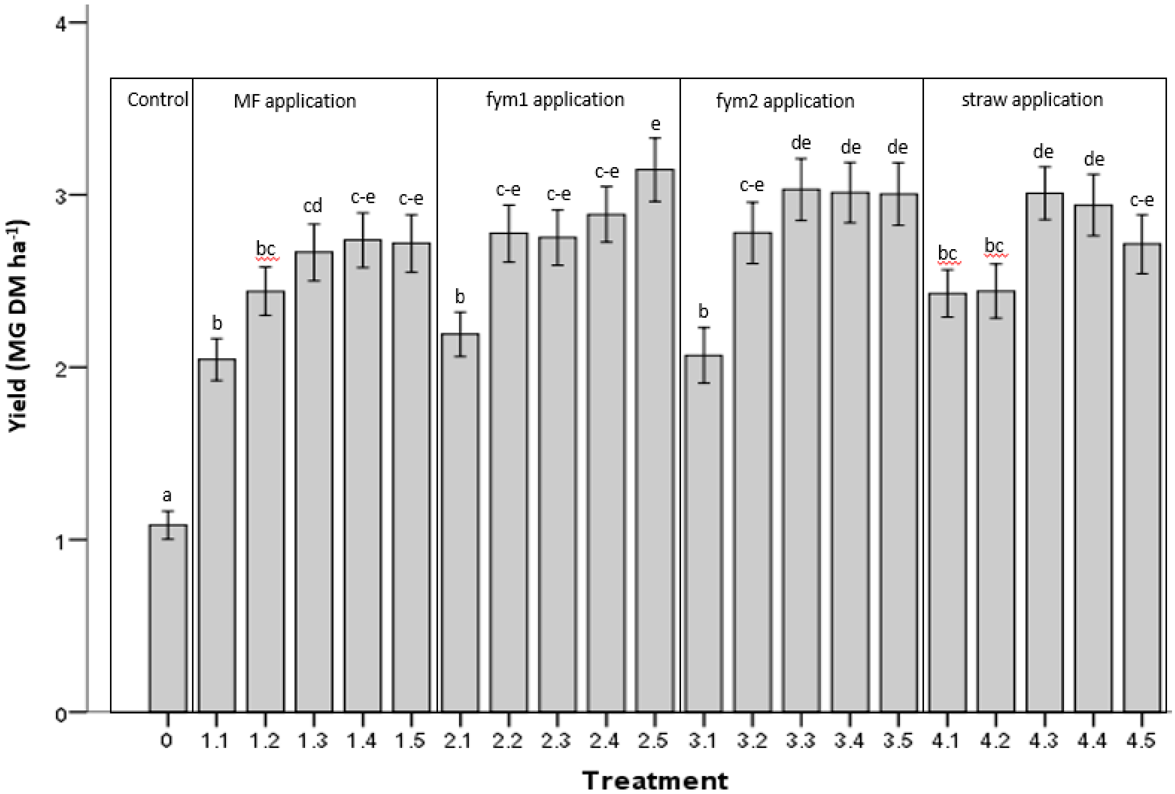



**Figure 2.4.** Effect of fertilizers on spring barley yields through nine growing seasons in every treatment.

Vertical lines indicate standard error (SE). Significant difference in spring barley by individual treatments over the year by Tukey Test. Treatments sharing the same letter are not significantly different ($P<0.05$). Treatment numbers are given in Table 2.1.

## 2.4. Discussion

### 2.4.1. Yield response to weather at the early growth stage

The weather conditions during the early growing stages are key determinants of the germination and emergence of a crop, which determine the crop yield (Zhou et al. 2007; Peltonen-Sainio et al. 2010; Hakala et al. 2012). This relationship was confirmed by the findings from this study: the SB yields were negatively affected by the average temperature in April, the temperature on the sowing day and the precipitation rates in March (Figure 2.2a, Figure 2.2b, and Table 2.2). This finding reflects an old Finnish saying referenced in the study of Hakala et al. (2012): "shivering sets the seed," implying that cold weather at the beginning of plant growth assures better yields in a temperate climate. Our result indicates that low temperatures at the early stage of plant growth (April) have beneficial effects on yield, as slow growth of the aboveground plant parts makes the plant tolerant to cold weather. Therefore, stem and leaf development is delayed, and roots may reach deeper into the soil, which helps the plant acquire more nutrients and water. In contrast, high temperatures may hasten growth, shorten developmental stages, especially the grain-filling period, and reduce yield (Evans 1976; Peltonen-Sainio et al. 2011). Moreover, higher temperatures also lead to higher evapotranspiration and subsequently increased soil water losses, resulting in drought, which can lower the yield capacity and result in a lower yield (Rajala et al. 2011; Peltonen-Sainio et al. 2015). Similar findings were reported by Peltonen-Sainio et al. (2011), who demonstrated negative yield responses to high temperature during early and mid-developmental stages. The results of this study are in accordance with results from Chmielewski and Köhn (1999), which showed that the yields of barley and oats in Germany decreased when the temperatures during the early growth stage were higher than the average temperature.

Regarding precipitation, it showed that high precipitation in March (before the growing season) led to a decrease in yield. Leaching may reduce nutrient availability and restrain seedling emergence and causing lower suboptimal density, which leads to lower yields. This is, especially then the case, when the period before sowing (March) is unusually wet or even affected by heavy rainfall and the period after sowing (April) is dry and warm (Zhou et al. 2007). Additionally, a large amount of pre-sowing precipitation can delay spring sowing due to soil saturation (Trnka et al. 2011), leading to a decrease in yield. Peltonen-Sainio et al. 2015) also found that delayed sowing is a cause of reduced yield when the conditions after sowing are too hot and dry for optimal yield formation. This explanation is shown explicitly in 2000, when high precipitation in March (80 mm) was followed by very late sowing (April 19th), a dry period after sowing which together and end up in low yields in summer. In contrast, in years of early sowing days (April 4th, 1980, March 24th, 2016) SB yields were above average (Figure 2.3). As high precipitation in March and high temperature during the early growing season have a strong effect on later growth.

### 2.4.2. Yield response to nutrients and weather

The interaction effect of annual weather conditions and fertilizer application on the variation in SB yield from year to year was pronounced. The lowest yield of SB was in the year 2000, which had high drought stress in the growing season after wet conditions in March (Figure 2.3). The average temperature and total precipitation in the growing season (April-July) in 2000 were 15.1°C and 82 mm, respectively. Consistent with other findings (Fernandez-Getino et al. 2015), our data showed that the average SB yield in the best fertilizer treatment (NPK+fym2) was significantly different from that in some other treatments in the years (1980 and 1990) with greater precipitation during the growing season.

These findings can be attributed to the effect of fertilization on the crops under sufficient water and drought conditions (Freckleton et al. 1999). The findings are further confirmed by the results of the MRM of SB yield and weather variables presented in Table 2.2, which showed that the yield in the NPK+fym2 treatment varied due to the sowing day temperature and March precipitation, which are important water availability factors for crops in the initial stages. The results of the model indicated the importance of favorable weather conditions for the growth of SB, as they later led to an increase in yield. Considerably lower yields were obtained in dry years (1982, 2000, and 2008, with total growing season precipitation < 160 mm) as a consequence of hot temperatures and dry conditions. The SB in the fertilizer treatments did not perform well since the crop is very sensitive to heat and water deficits, especially during tillering (Svobodová and Misa 2004; Pohanková et al. 2018).

In period (1), the barley yields obtained in 1976 and 1978 were not significantly different when the same variety ("Trumpf") was used (Figure 2.3). The average temperature and total precipitation in the growth period in 1976 and 1978 were similar, with averages of 13.9°C and 12.9°C and 102 mm and 119 mm, respectively (Figure 2.1a, Figure 2.1b). However, in period (2), the yields obtained in 2008 and 2016 were significantly different when using the same variety ("Simba"). The average temperature and total precipitation in the growth period 2008 and 2016 were similar, with averages of 15.1°C and 15.4°C and 159 mm and 167 mm, respectively. However, precipitation rates were highest in March and April in 2008 but decreased in the later stages (May-July). In 2016, in contrast, there was less precipitation in the early growth stage from March to May and higher precipitation from June to July, particularly in June. This may have caused the lower SB yield in 2008 than in 2016.

The coefficients of variation indicated that increasing the level of nutrients applied to SB decreased the degree to which the yield responded to the climate in general, with the exception of fym2. This is consistent with findings from Macholdt et al. (2019) and Ellmer et al. (2001), who noted that a stable supply of nutrients to crops could improve not only their grain yield but also their yield stability. The chemical fertilization treatment improved the nutrient availability more than the control and organic material application alone. The fym2 treatment had a lower stable yield than the control treatment. The cause for this result might be competition for nutrients between the plants and microorganisms that break down organic matter (Kaye and Hart 1997; Hodge et al. 2000a, 2000b). This microbial process depends on soil and weather conditions, such as the amount of soil water, soil temperature, air temperature, and precipitation (Bardgett et al. 2003; Davidson and Janssens 2006; Kuzyakov and Xu 2013; Ihara H et al. 2014). Sandy soil and low precipitation do not provide favourable conditions for the activities of microbes that break down organic matter (Mengel and Kirkby 2001; Koorem et al. 2014; Fujii et al. 2018). In summary, the amounts of nutrients released and available from organic material in the study site were unfavorable for SB growth. In addition, the weather variables changed every year, which influenced the process of organic matter breakdown, leading to enhanced yield variability in the fym2 treatment compared with that in the control treatment.

### 2.4.3. Long-term effects of fertilization regimes on SB yield

Similar to the observations of other authors (Tajnšek et al. 2013; Yang et al. 2015; Wei et al. 2016), we also found that yield was influenced by both the fertilization regime (NPK, fym2, NPK+fym1, NPK+fym2, and NPK+straw) and the MN application rate (25, 50, 75, 100 and 125 kg) during the experimental period. The average yield over the years increased with increasing MN. However, the rate of increase in yield differed according to the OR application. The highest combined effect of MN and OR was found for the fym2 application, while there were no clearly different effects among the other OR applications, such as between fym1 and straw (Table 2.3 and Figure 2.4). The results indicated that the average SB yield increased in the nutrient application treatments in the following order: NPK+fym2 > NPK+fym1, NPK+straw or NPK > fym2 > control. However, the highest yields under NPK or fym1 application was obtained at a 50 kg $ha^{-1}$ N supply; under the fym2 application, at a 25 kg $ha^{-1}$ N supply; and under straw application, at a 75 kg $ha^{-1}$ N supply. The fym2 treatment resulted in slightly lower yields than the NPK treatment, which was caused by the slow release and low utilization efficiency of organic N. When NPK, NPK+fym1 and NPK+straw were compared, fym1 and straw did not have any advantage over the regime with NPK. The higher yield in the NPK+fym2 treatment than in the other treatments may be due not only to a greater benefit from the organic N in fym and a higher rate of additional fym supply but also to the improvement of other properties, e.g., soil chemical, physical and biological effects (Kismányoky and Tóth 2013). The positive effect of fym2 application was notable at lower MN supply rates (N1 and N2) (Figure 2.4). Hence, supplying MN to SB at low rates promoted the processes of breaking down organic matter and releasing available N for the growth of SB, resulted in an increased yield.

Due to fluctuations in precipitation among growing seasons in period (1), SB yields in the fertilizer treatment also fluctuated and drastically went down and reach the bottom in 2000, the year that was also the driest. The yield in the control gradually declined to trough in 2000. However, the SB yield for all treatments (including control) increased after 2000 (period 2). One important factor that we do not statistically evaluate but could have an important effect on SB yields and its variability is the preceding year crop. In this experiment period, all the treatments in period (1) were preceded by sugar beet and by potato in period (2). Generally, sugar beet produced higher yields (21 t $ha^{-1}$ $a^{-1}$ in unfertilized/ control and approximately 46 t $ha^{-1}$ $a^{-1}$ in fertilizer treatment) than potato (approx. 13 t $ha^{-1}$ $a^{-1}$ in unfertilized/ control and 29 t $ha^{-1}$ $a^{-1}$ on fertilizer treatments) (see Figure 2.5). Kunzová and Hejcman (2009, 2010) and Hejcman Kunzová (2010) reported that the yield level of the preceding crop is an important determinant of the successive crop yield. Thus, the different biomass yields of sugar beet and potato could be implicated in the SB yield variability over time. Because preceding crop type and preceding crop yield resulted in uptake of nutrients and moisture in the soil, which related the growth of SB as a succeeding crop.

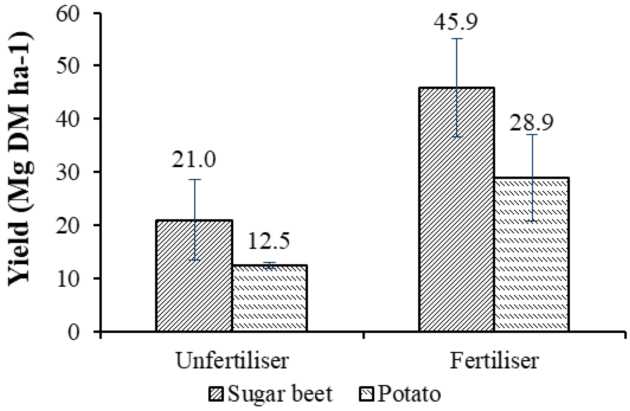



**Figure 2.5.** Yield of preceding crop (sugar beet and potato) over time. The yield in unfertilized (control), and fertilizer treatment are given.

Vertical lines indicate standard deviation (SD).

## 2.5. Conclusions

The SB yields were affected by fertilization regimes, annual weather conditions, and their interactions. Mineral N fertilization decreased overall yield variability across seasons as compared to no fertilization and organic fertilization regime show higher yield variability. The combined application of MN and OR produced higher SB yields than the application of either MN or OR. At the highest SB yields were found in NPK+fym2. Greater total precipitation during the growing season (April-July) increased SB yields when supplied high MN (N >75 kg/ha$^{-1}$), while at the early growth stage, a higher precipitation rate (March) and higher temperature (April or sowing day) negatively affected SB yield. One important factor that could have also influenced SB yields and yield stability is the preceding crop which could statistically evaluate in further study. The results of this analysis contribute to comprehensive crop production sustainability with regard to climate change. Further analysis of the effect of long-term fertilizer treatments on soil elements will be important for explaining the dynamics of nutrient depletion in the soil over time.

## Acknowledgements

This work was financially supported by the Vietnamese Government through the 911 project. We would like to thank the Müncheberg Experimental Station, ZALF, for providing the data that were used in this analysis. Special thanks go to the current manager of the Müncheberg Experimental Station and several generations of technical staff for long-term maintenance of the experiment.

## References

Bardgett RD, Streeter TC, Bol R. 2003. Soil microbes compete effectively with plants for organic-nitrogen inputs to temperate grasslands. Ecology. 84:1277-1287.

Blanchet G, Gavazov K, Bragazza L, Sinaj S. 2016. Responses of soil properties and crop yields to different inorganic and organic amendments in a Swiss conventional farming system. Agric Ecosyst Environ. 230:116-126.

BonaRes Data Portal. 2019. BonaRes: Soil as a sustainable resource for the bioeconomy; [accessed 2019 Feb 25]. https://maps.bonares.de/mapapps/resources/apps/bonares/index.html?lang=en. Registration required.

Černý J, Balík J, Kulhánek M, Čásová K, Nedvěd V. 2010. Mineral and organic fertilization efficiency in long-term stationary experiments. Plant Soil Environ. 56:28-36.

Chmielewski FM, Potts J. 1995. The relationship between crop yields from an experiment in southern England and long-term climate variations. Agric Forest Meteorol. 73:43-66.

Chmielewski FM, Köhn W. 1999. Impact of weather on yield components of spring cereals over 30 years. Agric For Meteorol. 96:49-58.

Davidson EA, Janssens IA. 2006. Temperature sensitivity of soil carbon decomposition and feedbacks to climate change. Nature. 440:165.

Ellerbrock RH, Höhn A, Rogasik J. 1999. Functional analysis of soil organic matter as affected by long-term manurial treatment. Eur J Soil Sci. 50:65-71.

Ellmer F, Erekul O, Kohn W. 2001. Influence of long-term different organic-mineral fertilisation on yield, yield structure and bread-making quality of winter wheat. Arch Agron Soil Sci. 47:423-444.

Evans LT. 1976. Aspects of the comparative physiology of grain yield in cereals. Adv Agron (USA). 28:301-359.

[FAO] Food and Agriculture Organization of the United Nations. 2006. Guidelines for Soil Profile Description. 4th ed. Rome.

[FAO] The Food and Agriculture Organization Corporate. 2019. Statistical Online Database; [accessed 2019 Jan 5]. http://faostat.fao.org/.

Fernandez-Getino AP, Santin-Montanya MI, Zambrana E, de Andres EF, Tenorio JL. 2015. The response of barley to rainfall and temperature in different tillage and crop rotation systems in semi-arid conditions. Ann Appl Biol. 166:143-153.

Fisher RA. 1925. III. The influence of rainfall on the yield of wheat at Rothamsted. Philos Trans Royal Soc B. 213:89-142.

Freckleton RP, Watkinson AR, Webb DJ, Thomas TH. 1999. Yield of sugar beet in relation to weather and nutrients. Agric For Meteorol. 93:39-51.

Friedt W, Ordon F. 2013. Advance in Barley Sciences: Barley production and breeding in Europe: modern cultivars combine disease resistance, malting quality and high yield. p. 389-400. Springer.

Fujii K, Shibata M, Kitajima K, Ichi T, Kitayama K, Turner BL. 2018. Plant-soil interactions maintain biodiversity and functions of tropical forest ecosystems. Ecol Res. 33:149-160.

Hakala K, Jauhiainen L, Himanen SJ, RÖTter R, Salo T, Kahiluoto H. 2012. Sensitivity of barley varieties to weather in Finland. J Agric Sci. 150:145-160.

Hejcman M, Kunzova E. 2010. Sustainability of winter wheat production on sandy-loamy Cambisol in the Czech Republic: Results from a long-term fertilizer and crop rotation experiment. Field Crops Res. 115(2):191-199.

Hodge A, Robinson D, Fitter A. 2000a. Are microorganisms more effective than plants at competing for nitrogen?. Trends Plant Sci. 5:304-308.

Hodge A, Stewart J, Robinson D, Griffiths B, Fitter A. 2000b. Competition between roots and soil micro-organisms for nutrients from nitrogen-rich patches of varying complexity. Ecology. 88:150-164.

Hoeting JA, Madigan D, Raftery AE, Volinsky CT. 1999. Bayesian model averaging: a tutorial. Stat Sci. 382-401.

Ihara H, Kato N, Takahashi S, Nagaoka K. 2014. Effect of soil solarisation on subsequent nitrification activity at elevated temperatures. Soil Sci Plant Nutr. 60(6):824-831.

Kaye JP, Hart SC. 1997. Competition for nitrogen between plants and soil microorganisms. Trends Ecol Evol. 12:139-143.

Kismányoky T, Tóth Z. 2013. Effect of mineral and organic fertilization on soil organic carbon content as well as on grain production of cereals in the IOSDV (ILTE) long-term field experiment, Keszthely, Hungary. Arch Agron Soil Sci. 59:1121-1131.

Koorem K, Gazol A, Öpik M, Moora M, Saks Ü, Uibopuu A, Sõber V, Zobel M. 2014. Soil nutrient content influences the abundance of soil microbes but not plant biomass at the small-scale. PLoS One. 9(3):e91998.

Kunzová E, Hejcman M. 2009. Yield development of winter wheat over 50 years of FYM, N, P and K fertilizer application on black earth soil in the Czech Republic. Field Crops Res. 111(3):226-234.

Kunzová E, Hejcman M. 2010. Yield development of winter wheat over 50 years of nitrogen, phosphorus and potassium application on greyic Phaeozem in the Czech Republic. Eur J Agron. 33(3):166-174.

Kuzyakov Y, Xu X. 2013. Competition between roots and microorganisms for nitrogen: mechanisms and ecological relevance. New Phytol. 198:656-669.

Gomez KA, Gome AA. 1983. Statistical procedures for agricultural research. International Rice Research Institute, Philippines, 1976: A Wiley. p. 382-388.

Lobell DB, Field CB. 2007. Global scale climate–crop yield relationships and the impacts of recent warming. Environ Res Lett. 2:014002.

Macholdt J, Piepho HP, Honermeier B. 2019. Mineral NPK and manure fertilization affecting the yield stability of winter wheat: Results from a long-term field experiment. Eur J Agron. 102:14-22.

Mengel K, Kirkby E. 2001. Principles of plant nutrition. 5th ed. Dordrecht, the Netherlands: Kluwer Academic.

Muller A, Schader C, Scialabba NE-H, Brüggemann J, Isensee A, Erb KH, Smith P, Klocke P, Leiber F, Stolze M. 2017. Strategies for feeding the world more sustainably with organic agriculture. Nat Commun. 8:1290.

Peltonen-Sainio P, Jauhiainen L, Trnka M, Olesen JE, Calanca P, Eckersten H, Eitzinger J, Gobin A, Kersebaum K C, Kozyra J. 2010. Coincidence of variation in yield and climate in Europe. Agric Ecosyst Environ. 139:483-489.

Peltonen-Sainio P, Jauhiainen L, Hakala K. 2011. Crop responses to temperature and precipitation according to long-term multi-location trials at high-latitude conditions. J Agric Sci. 149:49-62.

Peltonen-Sainio P, Rajala A, Känkänen H, Hakala K. 2015. Crop Physiology: Improving farming systems in northern Europe. 2nd ed. p. 65-91. Elsevier.

Pohanková E, Hlavinka P, Orság M, Taká?, J., Kersebaum, K., Gobin, A., and Trnka, M. (2018). Estimating the water use efficiency of spring barley using crop models. J Agric Sci. 156:628-644.

Příkopa M, Richter R, Zimolka J, Cerkal R. 2005. The influence of the year, fore-crops and fertilisation on yield and content of crude protein in spring barley. Plant Soil Enviro. 51:144-150.

Raftery AE. 1995. Bayesian model selection in social research. Sociol Methodol. 25:111-164.

Raftery AE, Madigan D, Hoeting JA, 1997. Bayesian model averaging for linear regression models. J Amer Statist Assoc. 92:179-191.

Rajala A, Hakala K, Mäkelä P, Peltonen-Sainio P. 2011. Drought effect on grain number and grain weight at spike and spikelet level in six-row spring barley. J Agron Crop Sci. 197:103-112.

Svobodová I, Misa P. 2004. Effect of drought stress on the formation of yield elements in spring barley and the potential of stress expression reduction by foliar application of fertilizers and growth stimulator. Plant Soil Environ. 50:439-446.

Tajnšek A, Čergan Z, Čeh B. 2013. Results of the long-term field experiment IOSDV Jable at the beginning of the 21st century. Arch Agron Soil Sci. 59:1099-1108.

Timsina J. 2018. Can organic sources of nutrients increase crop yields to meet global food demand? Agronomy. 8:214.

Trnka M, Hlavinka P, Semerádová D, Dubrovský M, Žalud Z, Možný M. 2007. Agricultural drought and spring barley yields in the Czech Republic. Plant Soil Environ. 53:306–316

Trnka M, Olesen JE, Kersebaum KC, Skjelvåg AO, Eitzinger J, Seguin B, Peltonen-sainio P, Rötter R, Iglesias A, Orlandini S. 2011. Agroclimatic conditions in Europe under climate change. Glob Change Biol. 17:2298-2318.

Váňová M, Palík S, Hajšlová J, Burešová I. 2006. Grain quality and yield of spring barley in field trials under variable growing conditions. Plant Soil Environ. 52:211-219.

Walter K. 2017. Crop Report No.1 German 2017 Spring Barley. German: Braugersten-Gemeinschaft e.V.; [accessed 2018 Oct 16]. https://bestmalz.de/en/crop-report-1-on-german-2017-spring-barley/.

Wei W, Yan Y, Cao J, Christie P, Zhang F, Fan M. 2016. Effects of combined application of organic amendments and fertilizers on crop yield and soil organic matter: An integrated analysis of long-term experiments. Agric Ecosyst Environ. 225:86-92.

Yang J, Gao W, Ren S. 2015. Long-term effects of combined application of chemical nitrogen with organic materials on crop yields, soil organic carbon and total nitrogen in fluvo-aquic soil. Soil Till Res. 151:67-74.

Zhou M, Li H, Mendham N. 2007. Combining ability of waterlogging tolerance in barley. Crop Sci. 47:278-28

**Supplement**

**Table S 2.1**. Crop rotation of the long-term experiment "V140". In bold: Spring barley, in grey: preceding crops.

| Year | Crop | Year | Crop | Year | Crop |
|---|---|---|---|---|---|
| 1963 | Maize | 1981 | Sugar beet | 1999 | Potato |
| 1964 | Winter rye | 1982 | **Spring barley** | 2000 | **Spring barley** |
| 1965 | Potato | 1983 | Potato | 2001 | Pea |
| 1966 | Winter rye | 1984 | Winter wheat | 2002 | Winter wheat |
| 1967 | Potato | 1985 | Sugar beet | 2003 | Maize |
| 1968 | Summer wheat | 1986 | **Spring barley** | 2004 | Winter rye |
| 1969 | Sugar beet | 1987 | Potato | 2005 | Oil flax |
| 1970 | **Spring barley** | 1988 | Winter wheat | 2006 | Winter rye |
| 1971 | Maize | 1989 | Sugar beet | 2007 | Potato |
| 1972 | Winter rye | 1990 | **Spring barley** | 2008 | **Spring barley** |
| 1973 | Potato | 1991 | Potato | 2009 | Pea |
| 1974 | Winter wheat | 1992 | Winter wheat | 2010 | Winter wheat |
| 1975 | Sugar beet | 1993 | Sugar beet | 2011 | Maize |
| 1976 | **Spring barley** | 1994 | Winter wheat | 2012 | Winter rye |
| 1977 | Sugar beet | 1995 | Maize | 2013 | Oil flax |
| 1978 | **Spring barley** | 1996 | Winter rye | 2014 | Winter rye |
| 1979 | Sugar beet | 1997 | Oil flax | 2015 | Potato |
| 1980 | **Spring barley** | 1998 | Winter rye | 2016 | **Spring barley** |

The position of spring barley within the crop rotation was changed during the long term experiment i.e. 2 years from 1975-1982: 4 rounds of sugar beet-spring barley; 4 years from 1983-1990: 4 rounds of potato-winter wheat-sugar beet-spring barley; 7 years from 1999-216: 3 rounds of potato-spring barley- pea-winter wheat-maize-winter rye-oil flax- winter rye. Between 1991- 1998 no spring barley was seeded.

**Table S 2.2**. Soil chemical properties (0-25 cm) in each treatment through eight spring barley seasons (except 2000).

| Treatment | pH (KCl) | Nt (mg/100g soil) | Ct (mg/100g soil) | available P (mg/100g soil) | available K (mg/100g soil) | Mg (CaCl2) (mg/100g soil) |
|---|---|---|---|---|---|---|
| 0 | 5.8 | 44.7 | 469.2 | 6.6 | 11.2 | 4.0 |
| 1.1 | 5.8 | 46.9 | 492.2 | 8.0 | 12.5 | 3.9 |
| 1.2 | 5.8 | 48.4 | 507.7 | 7.8 | 11.9 | 3.9 |
| 1.3 | 5.7 | 47.4 | 495.8 | 7.4 | 11.3 | 3.9 |
| 1.4 | 5.7 | 49.0 | 505.9 | 7.7 | 11.5 | 4.0 |
| 1.5 | 5.5 | 49.0 | 513.6 | 8.0 | 11.0 | 4.0 |
| 2.1 | 5.9 | 49.7 | 526.4 | 9.5 | 13.1 | 4.3 |
| 2.2 | 5.9 | 51.8 | 538.0 | 8.8 | 12.6 | 4.3 |
| 2.3 | 5.8 | 53.9 | 552.6 | 9.3 | 12.2 | 4.4 |
| 2.4 | 5.8 | 52.7 | 553.7 | 9.3 | 11.6 | 4.4 |
| 2.5 | 5.7 | 53.8 | 571.2 | 9.3 | 11.9 | 4.4 |
| 3.1 | 6.0 | 54.5 | 579.1 | 11.5 | 14.2 | 4.6 |
| 3.2 | 5.9 | 59.6 | 619.1 | 11.5 | 14.4 | 5.0 |
| 3.3 | 5.9 | 58.2 | 598.3 | 11.1 | 13.4 | 5.0 |
| 3.4 | 5.9 | 59.6 | 604.4 | 11.4 | 13.3 | 4.8 |
| 3.5 | 5.8 | 58.2 | 603.4 | 11.1 | 12.9 | 4.6 |
| 4.1 | 5.8 | 51.5 | 544.1 | 8.1 | 12.9 | 4.1 |
| 4.2 | 5.8 | 50.2 | 530.7 | 8.2 | 11.7 | 4.1 |
| 4.3 | 5.8 | 51.3 | 538.3 | 8.3 | 11.7 | 4.2 |
| 4.4 | 5.6 | 51.4 | 549.9 | 8.1 | 11.4 | 4.0 |
| 4.5 | 5.5 | 50.4 | 535.1 | 7.8 | 11.3 | 3.9 |
| Average | 5.8 | 52.0 | 544.2 | 9.0 | 12.3 | 4.3 |
| max | 6.0 | 59.6 | 619.1 | 11.5 | 14.4 | 5.0 |
| min | 5.5 | 44.7 | 469.2 | 6.6 | 11.0 | 3.9 |

Treatments (1.1 - 4.5) are each rate of mineral N fertiliser (five rates) combined with each organic fertiliser (four variants: no organic, fym1, fym2, straw). Treatment “0”: control, no fertilisation.

**Table S 2.3.** Results from analysis of variance (ANOVA) with Eta squared between fertilizer, year (annual weather) and spring barley yields. Dependent Variable : yield of spring barley (t $ha^{-1}$)

| Source | Type III Sum of Squares | df | Mean Square | F | Sig. | Eta Squared ($h^2$) |
|---|---|---|---|---|---|---|
| Corrected Model | 1604.999[a] | 188 | 8.537 | 16.838 | 0.000 | |
| Intercept | 8406.239 | 1 | 8406.239 | 16579.928 | 0.000 | |
| Treatment | 234.968 | 20 | 11.748 | 23.172 | 0.000 | 11 |
| Year | 1173.631 | 8 | 146.704 | 289.349 | 0.000 | 55 |
| Treatment * Year | 164.358 | 160 | 1.027 | 2.026 | 0.000 | 8 |
| Error | 543.011 | 1071 | 0.507 | | | 26 |
| Total | 10747.722 | 1260 | | | | |
| Corrected Total | 2148.009 | 1259 | | | | |

a. R Squared = 0.747 (Adjusted R Squared = 0.703)

The Table S 2.3 shows result of the two-way ANOVA – namely, whether either of the two independent variables (treatment and year) or their interaction are statistically significant. The "Sig." column showed that the treatment, year and their interaction have a statistically significant effect on grain yield of spring barley. Besides, “Eta squared ($h^2$)” column showed proportion of total variance that is attributed to an effect. In this case, the yields of SB were significantly affected by treatment, representing fertiliser application (11%), affected by year, representing annual weather conditions (55%), and the year × treatment interaction (8%); 26% of the variation was due to error (other factors), and the adjusted R-squared was 0.703.

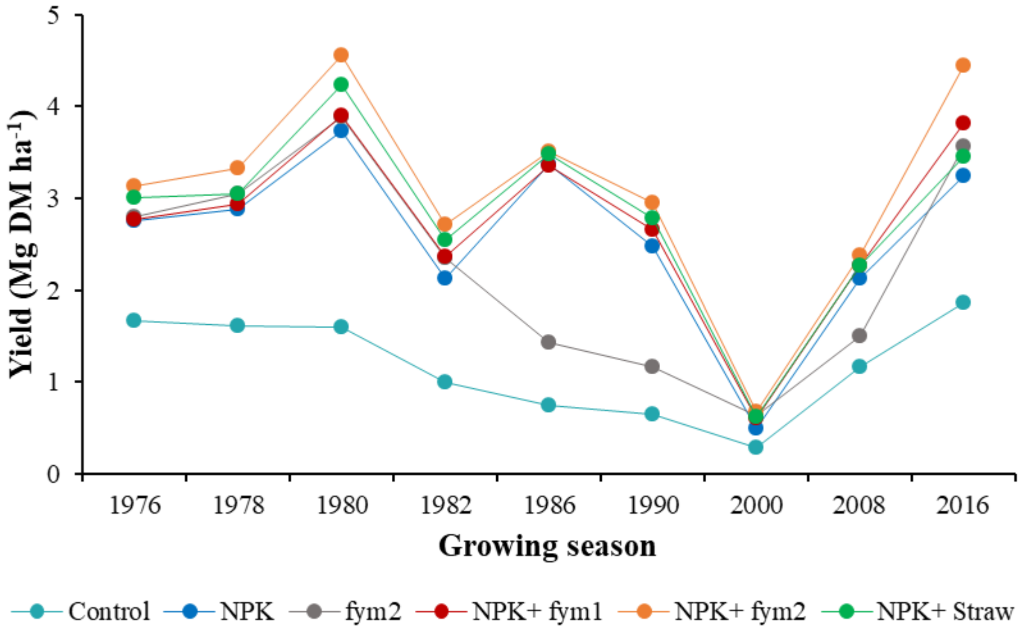



**Figure S 2.1.** Temporal dynamic of spring barley yield during the long term field experiment period.

NPK: no organic fertilisation, the average value of rate of mineral N from N1- N4; NPK + fym1: NPK + 1.2 t DM $ha^{-1}$ $a^{-1}$ farmyard manure; NPK + fym2: NPK + 3.2 t DM $ha^{-1}$ $a^{-1}$ farmyard manure; fym2: only 3.2 t DM $ha^{-1}$ $a^{-1}$ farmyard manure; NPK + straw: NPK + 2 t $ha^{-1}$ $a^{-1}$ straw.

# Chapter 3

## M5P machine learning algorithm for analysis of winter rye yield in a long-term field experiment in Müncheberg, Northeast Germany

Thi Huyen Thai, Sravya Mamidanna, Dietmar Barkusky, Sonoko Dorothea Bellingrath-Kimura

*European Journal of Agronomy (under review)*

# MP5 algorithm for analysis of winter rye yield in a long-term field experiment in Müncheberg, Germany

Thi Huyen Thai[a,b,c*], Sravya Mamidanna[a], Dietmar Barkusky[a], Sonoko Dorothea Bellingrath-Kimura[a,b]

[a]Leibniz Centre for Agricultural Landscape Research (ZALF), Eberswalder Str. 84, 15374 Müncheberg, Germany

[b]Institute of Agriculture and Horticulture, Faculty of Life Science, Humboldt University of Berlin, Albrecht-Thaer-Weg 5, 14195 Berlin, Germany

[c]Faculty of Agronomy, University of Agriculture and Forestry, Hue University, 102 Phung Hung street, Hue, Vietnam

[*]Correspondence: *thaihuyen@hueuni.edu.vn*; *Tel.: +49-(0)-33432-82-313*

## Abstract

Identifying major factors and investigating the relationships that account for crop yield and yield variability is important to understand how to maximize crop yields and minimize yield fluctuations facing the uncertainties under climate change. Data from a long-term fertilizer field experiment "V140," which was established in 1963 in sandy soil in Northeast Germany, were used to evaluate the grain yield of winter rye. Decision trees, machine learning techniques were used to explore the relationships to provide tailor-made agronomic recommendations in conjunction with seasonal weather forecasts to extend to agencies and farmers. Our results reveal that the winter rye grain yield response to fertilizer management comprises complex relationships among climatic dependence, crop rotation, and soil characteristics. The most important determinant of winter rye yield was mineral nitrogen application. The following was weather condition in the early stage of the growing season (in autumn), especially the temperature in September, precipitation in November; and the intensity and duration of extreme temperatures in the summertime (harvest year), especially the number of days recorded with a maximum temperature above 30°C in July and the temperature in May. Additionally, farmyard manure application and the preceding year's crop were also significant variables explaining the yield variability of winter rye. Soil parameter, in particular total carbon although less present in the model than other but also involved a role that influence WR yield variation once NF application more than $70kg^{-1}$ and temperature in September higher than 13.7°C. This finding implies that the strategies for improving yields or reducing the year-to-year yield variability of winter rye in sandy soil must consider the role of supply mineral fertilizer and organic fertilizer, in particular mineral

nitrogen and farmyard manure fertilizer, seasonal weather forecasts and crop rotation such preceding crop.

**Keywords:** Winter rye, long-term field trial, nitrogen fertilizer, organic fertilizer, machine learning algorithm.

### 3.1. Introduction

Major drivers of crop yields and their variability include weather conditions, soil properties, weeds, diseases, and pests (Gregory et al., 2009; Silungwe et al., 2019). In addition, management practices related to decisions such as fertilizer applications, crop rotations, irrigation and tillage, result in year-to-year crop yield variability (Brisson et al., 2010; Silungwe et al., 2018). Crop yield and yield variability therefore are the result of complex interdependencies and interactions among different factors. Identifying both the major factors and relationships that account for crop yield and yield variability is important to understanding how to maximize crop yields and minimize yield fluctuations. This is a challenge that requires long-term investigation. Long-term field experiments provide the necessary data and insights into identifying such factors and relationships and their influence on crop yields. This study is a follow-up to the previous work of the authors that addressed the effect of long-term fertilizer management and weather on cereal yields. Thai et al. (2019) investigated the yield response of spring barley and provided evidence that yield was a primary product of the relationship between fertilizer regimes and weather. In addition, the preceding crop was noted in the study as an important factor that could also have influenced the yield variation of spring barley. To unlock the complexities of factors interdependencies in influencing different crop yields, this study focuses on analyzing and evaluating the responses of grain yield of winter rye (WR) to weather and agronomic factors interactions using data collected as part of the V140 experiment in Müncheberg, Germany.

Rye (*Secale cereal* L.) is an important cereal crop in Europe, accounting for more than 75% of global rye production (FAO, 2019). Rye is almost exclusively cultivated as a winter crop in marginal locations with infertile, light soil with low water-holding capacity in Central and Eastern Europe such as Russia, Belarus, Ukraine, Poland, and Germany (Miedaner et al., 2012). Germany is a leading producer of rye with a global production share of 18.3% (FAO, 2019). In 2017, Germany recorded the highest rye production (approximately 2.7 million tons); however, its average grain yield is lower (at 5 t $ha^{-1}$) compared with neighboring European countries such as Sweden (6.7 t $ha^{-1}$), Denmark (6.5 t $ha^{-1}$), and Switzerland (6.2 t $ha^{-1}$).

Improving the average yield per hectare is indispensable; however, such achievement can be achieved only if the complex interdependencies of the factors affecting it are understood. In this case, classical analytical techniques such as analysis of variance, regression and parametric correlation are commonly used to evaluate yield response and differentiate between various management regimes and environmental and agronomic factors in long-term agronomic experiments (Gauch, 2006; Wu and Hamada, 2011). However, as noted earlier, the factors determining crop yield response are interdependent, and their interactions often lead to reinforcing loops or thresholds that indicate the existence of highly non-linear relationships (Krupnik et al., 2015; Lobell et al., 2005; Tack et al., 2015). The analysis of such interactions necessitates an ability to determine such non-linear patterns and requires application of multivariate analysis methods as opposed to traditional methods (Hastie et al., 2009; Lobell et al., 2005; Loh, 2006; Zheng et al., 2009). A range of multivariate and non-parametric analytical methods exist which can be applied to uncover such non-linear dependencies to model yield response; they include generalized linear and mixed effect models (Bolker et al., 2009), machine learning algorithms such as decision trees, Bayesian rule, neural networks, ensemble (e.g., random forest) (Breiman et al., 1984; Loh, 2011), among many others. Recently, decision tree algorithm (Song and Ying, 2015) has become popular in agriculture research that has been used to assess yield response to agronomic (Sileshi et al., 2010; Zhang et al., 2012), environmental (Dacko et al., 2016; Vagh and Xiao, 2012) and management (Delmotte et al., 2011; Zheng et al., 2010) factors. The decision tree model not only can apply for large data sets without a hypothesis but also can apply for small datasets from a designed experiment as well (Loh, 2006). Decision trees algorithm is conceptually simple yet powerful analytical tools which is one of the most effective and widely used classification methods for data mining (De'ath and Fabricius, 2000; Trajanov et al., 2019; Zhang, 2006). Unlike other supervised learning algorithms, the decision tree algorithm can be used not only for classification systems but also can be used for solving regression problems with multiple covariates as well. This study was designed to train a decision tree model to explore the input/output relationship in a dataset from a long-term field experiment in Müncheberg, Northeast Germany.

To this end, the aims of this work were as follows:

(i) To analyze long-term effect of fertilization regimes on the WR grain yield and estimate factors influencing the yield variation.

(ii) To identify driving factors of the WR yield variation in sandy soil from the long-term field experiment using decision trees method.

### 3.2. Materials and methods

#### 3.2.1. Site description and experimental design

The data were collected from an agricultural long-term field experiment (LTFE) ("V140", Figure 3.1). V140 is a long-term fertilizer field experiment the original aim of which was to demonstrate how fertilizers affected soil fertility. The LTFE was established in 1963 at ZALF, Müncheberg, Germany. It is located approximately 50 km east of Berlin. The site has recently been described in a study by Thai et al. (2019). This is a dry region with sandy soil, low nutrient inventory and low annual precipitation rates. Further information on the LTFE is presented in Table 3.1.

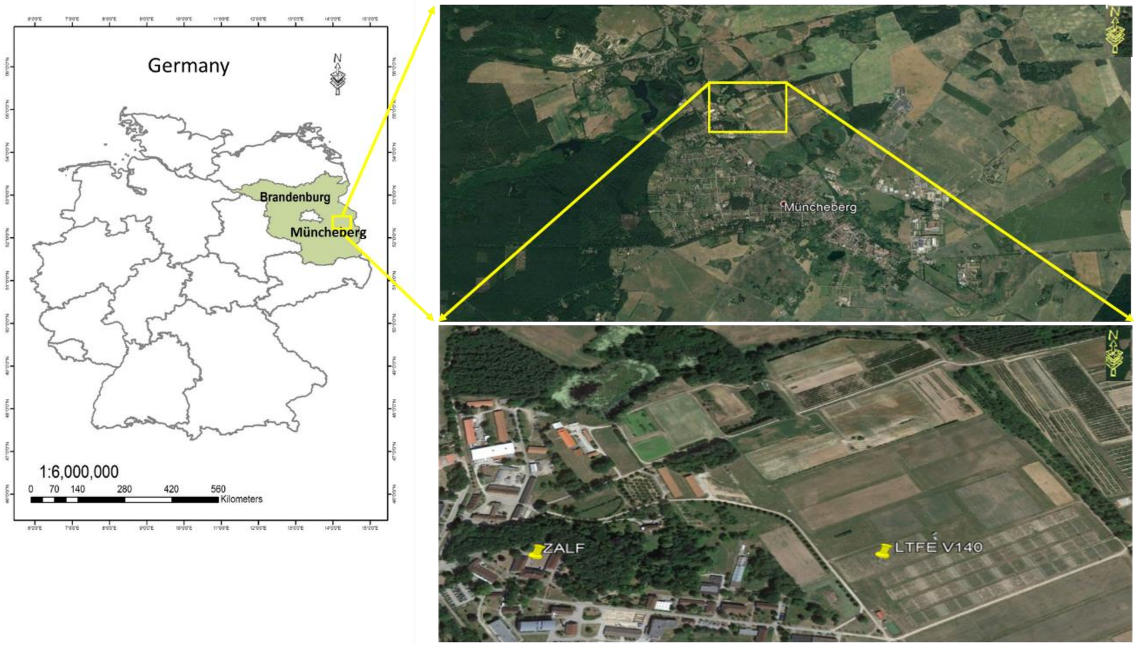



**Figure 3.1.** Map of the long-term experimental locations in Müncheberg, Germany

**Table 3.1.** The long-term fertilizer field experiment V140.

| Characteristics | Description | Information |
|---|---|---|
| Soil type | Podzoluvisol to Arenosol | |
| Soil characterize | | |
| | $pH_{KCl}$ | 5.4 – 5.9 |
| | Clay | 50 g kg$^{-1}$ |
| | Silt | 210 g kg$^{-1}$ |
| | Sand | 740 g kg$^{-1}$ |
| | Carbon content | 4.3 - 5.2 g kg$^{-1}$ |
| | CEC | 31.5-35.6 mmol$_c$ kg$^{-1}$ |
| Annual temperature | 1971-2014 | 8.8°C (6.5-10.4°C) |
| Annual precipitation | 1971-2014 | 545 mm (343-817 mm) |

Ellerbrock et al. (1999) adapted

The field experiment was designed as a randomized complete block of 21 treatments (including a control, Table 3.2), each with eight replicates. The trial consists of 168 individual plots, and the size of each plot is 30 m2 (6.0 m × 5.0 m). Five different rates of mineral fertilizer (MF) base on different five rates of mineral nitrogen fertilizer (MN) were applied with four organic fertilizer (OR) regimes: i) no OR, ii) 1.2 t dry mass (DM) ha$^{-1}$ a$^{-1}$ farmyard manure (= fym1), iii) 3.2 t DM ha$^{-1}$ a$^{-1}$ farmyard manure (= fym2), and iv) 2.0 t DM ha$^{-1}$a$^{-1}$ straw (Table 3.2). In addition to the control (no fertilization), five MN rates were 35, 70, 105, 140, and 175 kg N ha$^{-1}$, which are referred to as N0, N1, N2, N3, N4, and N5, respectively. In the regime with no OR, fym1, and straw the N1, N2, N3, N4, and N5 rates were included, whereas in the fym2 regime N0, N1, N2, N3, and N4 were included separately. Together with the control treatment, 21 treatments were included in this experiment. The treatments were grouped as shown in Table 3.2. Phosphorus and potassium fertilizer were applied the same for all plots, except the control.

**Table 3.2.** Description of the experimental treatments.

<table>
<tr><th>Treatment Code</th><th>Group treatment</th><th>Mineral nitrogen fertilizer-MN (kg ha$^{-1}$)</th><th>Organic fertilizer</th><th>Fertilizer application</th></tr>
<tr><td>0</td><td>Control</td><td>0</td><td>0</td><td>0</td></tr>
<tr><td>1.1</td><td rowspan="4">NPK</td><td>35</td><td rowspan="5">0</td><td rowspan="5">MF</td></tr>
<tr><td>1.2</td><td>70</td></tr>
<tr><td>1.3</td><td>105</td></tr>
<tr><td>1.4</td><td>140</td></tr>
<tr><td>1.5</td><td></td><td>175</td></tr>
<tr><td>2.1</td><td rowspan="4">NPK+fym1</td><td>35</td><td rowspan="5">1.2 t ha$^{-1}$ year$^{-1}$ DM farmyard manure</td><td rowspan="5">fym1</td></tr>
<tr><td>2.2</td><td>70</td></tr>
<tr><td>2.3</td><td>105</td></tr>
<tr><td>2.4</td><td>140</td></tr>
<tr><td>2.5</td><td></td><td>175</td></tr>
<tr><td>3.1</td><td>PK+fym2</td><td>0</td><td rowspan="5">3.2 t ha$^{-1}$ year$^{-1}$ DM farmyard manure</td><td rowspan="5">fym2</td></tr>
<tr><td>3.2</td><td rowspan="4">NPK+fym2</td><td>35</td></tr>
<tr><td>3.3</td><td>70</td></tr>
<tr><td>3.4</td><td>105</td></tr>
<tr><td>3.5</td><td>140</td></tr>
<tr><td>4.1</td><td rowspan="4">NPK+Straw</td><td>35</td><td rowspan="5">2.0 t ha$^{-1}$ year$^{-1}$ DM straw</td><td rowspan="5">straw</td></tr>
<tr><td>4.2</td><td>70</td></tr>
<tr><td>4.3</td><td>105</td></tr>
<tr><td>4.4</td><td>140</td></tr>
<tr><td>4.5</td><td></td><td>175</td></tr>
</table>

Treatments codes (1.1 - 4.5): each rate of mineral nitrogen fertilizer (five levels NF: 35, 70, 105, 140, 175 kg ha$^{-1}$, respectively) with organic fertilizer (four variants: no organic, 1.2 t dry mass (DM) ha$^{-1}$ a$^{-1}$ farmyard manure (fym1), 3.2 DM ha$^{-1}$ a$^{-1}$ farmyard manure (fym2), and 2.0 t DM ha$^{-1}$ a$^{-1}$ straw (straw). Treatment code "0" or control: no fertilizer inputs. Group treatments "NPK": sole mineral fertilizer applied at 35, 70, 105, 140 kg ha$^{-1}$ N; NPK + fym1: 1.2 t DM farmyard manure ha$^{-1}$ a$^{-1}$ applied at 35, 70, 105, 140 kg ha$^{-1}$ N; NPK + fym2: 3.2 t DM farmyard manure ha$^{-1}$ a$^{-1}$ applied at 35, 70, 105, 140 kg ha$^{-1}$ N; fym2: only 3.2 t DM farmyard manure ha$^{-1}$ a$^{-1}$; NPK + straw: 2.0 t DM ha$^{-1}$ a$^{-1}$ straw applied at 35, 70, 105, 140 kg ha$^{-1}$ N.

### 3.2.2. Crop management

WR was sown between the end of September and early October. The harvest was conducted between the end of July and the beginning of August depending on the weather conditions. The cropping system was conventional tillage with plowing, usually in autumn. The crop rotation was not fixed and consisted of winter wheat, winter rye, spring barley, potato, sugar beet, maize, flax, and pea (supplementary Table S 3.1). During each growing season in the experiment, only one of those crops was cultivated. The WR growing seasons cultivated in V140 during the period of the experiment from 1971-2014 were considered in this study. The crop preceding WR was maize (in 1971, 1995, 2003, and 2011) and flax (in 1997, 2005 and 2013). The seedbed was prepared immediately before sowing in autumn. Farmyard manure (fym) was applied every two years until 1994 in the autumn before planting maize, potato, or sugar beet; after that, fym was applied every four years in the autumn before planting maize or potato. The average nutrient contents of the dry mass manure used in the period of the experiment were 2.5% N, 1.0% P2O5, 2.4% K2O, 1.3% Mg and 60% organic matter. Straw was applied every two years throughout the experimental periods (using the straw from the preceding harvested cereals). The average nutrient contents of the dry mass straw used in the experiment were 0.6% N, 0.1% P2O5, 1.5% K2O, and 0.08% Mg. The plowing, cultivation, sowing, and liming methods and seeding rates were the same for all plots. The phosphorus and potassium fertilization rates (20-30 kg P2O5 $ha^{-1}a^{-1}$ and 100 kg K2O $ha^{-1}a^{-1}$) were the same for all plots (except the control) since 1980. Mineral nitrogen fertilizer was applied twice annually during WR growth: at the beginning period (the middle or end of March or early-April) and one month later between shooting to full blooming (the end of April or early May). The WR varieties changed over time; however, the "Minollo" variety was used in 2012 and 2014. Weeds were controlled with a post-emergence herbicide. WR was harvested plot by plot at the time of technological maturity by a plot harvester each year. The straw from each crop was removed from the experimental plots after harvest.

### 3.2.3. Data description

The data in the LTFE used for analysis consisted of the experiment period from 1971-2014 where winter rye was cultivated and experimental design stability.

Crop yield: The WR grain yield data (Mg DM $ha^{-1}$) were obtained from every plot in the seven years of the experiment. The DM yield of maize and flax (preceding crop) in every treatment and every replication was used to estimate the effect of the preceding crop on the WR yield.

Climate variables: Meteorological data used in the analysis were obtained from an adjacent climate station of the German Meteorological Service (DWD, 2019). For every year of WR planting, daily data of air median temperature, maximum temperature, minimum temperature and precipitation were recorded during the LTFE. Monthly parameters of average median temperature, cumulative precipitation, cumulative number of days recorded having temperatures above 30°C (days Tmax > 30°C), cumulative freezing days, and cumulative growing degree days during the growing season were calculated from these data. Maximum and minimum temperatures were used to calculate growing degree days (GDD). A freezing day was defined as when a daily median temperature (daily mean temperature) was below 0°C (32°F), and the cumulative monthly freezing days was the total number of freezing days.

Soil variables: During the WR period of the experiment, the soil chemical properties were measured in 1996, 1998, 2004, and 2012 in every plot of the experiment. Soil sampling carried out after harvest the WR. The average content of soil chemical properties is presented in Table S 3.2, supplement. Soil variables were selected for input of model such as total nitrogen (total N), total carbon (total C), plant-available phosphorus (plant-available P), plant-available potassium (plant-available K). The selected chemical soil parameters with eight replications in each treatment in those years were included in the study to estimate their effect on WR yields.

### 3.2.4. Data analysis

#### *3.2.4.1. General linear model (GLM)*

GLM was used to test for analysis of variance (ANOVA) of WR yield to estimate the yield variation due to the effects of treatment, annual effects (weather condition, others), and the size of their effects. In case of a significant ANOVA result, Tukey's HSD post hoc test was used to assess the differences in mean yields among treatments every year and over the years. The treatment effects were declared as significant with $P < 0.05$. When the WR yield data was evaluated over the years, fertilizer application was included as a fixed factor, and WR planting year was included as a random factor in the model. The analysis was performed using SPSS version 22.

#### *3.2.4.2. Decision tree analysis*

Decision trees are hierarchical models that recursively partition the data space and fit a prediction model within each partition that is graphically represented as inverted trees (Breiman et al., 1984; Loh, 2011). The tree contains a root node, internal nodes, leaf nodes or terminal nodes, splitting, and branches. The root node represents a choice that is entire population and will result in the subdivision of all records into two or more mutually exclusive subsets. The internal nodes are called chance nodes or also called decision nodes in which with each node, the value of a variable is tested and compared to a constant value. The branches or sub-trees are split from the root node and internal nodes correspond to the outcome of the test. The terminal nodes contain the predictions of the target variable that apply to all samples. When the values of the target variable are numeric, terminal nodes of the tree can be constant values in which decision trees are called regression trees. The terminal nodes of the tree can be pie-wise linear regression equations, in this case, the decision trees are termed as model trees (Azzeh, 2011; Song and Ying, 2015). Such constructed trees provide a simple and transparent structure for depicting complex interactions, enabling the end-user to intuitively understand relationships. The hierarchical representation shows the most important factors appear at the top node that influences the target variable (the WR yield). The importance of the various other independent variables in explaining the target variable decreases as you move towards the lower nodes of the tree.

There are different statistical algorithms for building decision trees available such as AID, THAID, CART (classification and regression trees), C4.5, CHAID (Chi-Squared automatic interaction detection), M5, and M5P. In which, CART algorithm is a basic version and the most commonly used for building decision tree model. The M5P is one later classification tree algorithm (Behnood et al., 2017; Blockeel and Struyf, 2002) expanded version of M5 algorithm that was originally discovered by Quinlan (1992). Using M5P is getting more advantages when compared to those early algorithms especially in modifications of the decision tree pruning process and smoothing process (Azzeh, 2011; Behnood et al., 2017). It uses multivariate linear models and chooses the variables at the partition nodes in a way that maximizes the expected error reduction as a function of the standard deviation of the target variable (Behnood et al., 2017; Zhang, 2006). Training decision tree by the M5P algorithm consists of four main steps. In the first step, the input data is split into several sub-spaces to build a tree. After building the tree, in the second step, a linear regression model is developed in each of the sub-space using data associate with sub-tree. Then, a pruning technique is applied to overcome the over-training problem. The final step is a smoothing process. M5P can be used for both categorical and continuous dependent variables, and it handles missing values. We used the M5P algorithm offered by WEKA (Waikato Environment for Knowledge Analysis) software (Witten et al., 2016) to identify complex relationships among WR crop yields and fertilizer as well as crop management, weather conditions, and soil characteristics. Two models were created to predict the WR yield responses in two scenarios: (1) with soil information and (2) without soil information.

Although advantage of the M5P algorithm in building decision tree is elementary well in the pruning process and smoothing process to reduce overfitting problem, however, the pruning process can cause sharp discontinuities between the adjacent linear models. To avoid this, we used both pre-pruning and post pruning approaches (Patel and Upadhyay, 2012; Song and Ying, 2015). For pre-pruning, Pearson correlation test with weather variables was used to removed highly correlated variables and through ANOVA test for soil variables to estimate the soil parameters variation due to the effects of treatment and selected interest soil variable. During training the decision trees, a common practice is to use tenfold cross-validation as a standard technique to measure the predictive performance of such decision tree models (Witten et al., 2016; Zhang, 2006). In cross-validation, the dataset was split into n approximately equal partitions (folds). Each fold is used for testing while the remaining folds are used for training or building the model. This procedure was repeated n times, and at the end, the correlation coefficients obtained in the different iterations were averaged to obtain the overall correlation coefficients of the models (Trajanov et al., 2019; Witten et al., 2016). After generating a full decision tree, backward pruning was used to remove branches in a manner that improves the accuracy of the overall classification when applied to the validation dataset. There are various methods for post-pruning. In which, a method that considers the proportion of records with error prediction is one of common methods of selecting the best possible sub-tree from several candidates (Song and Ying, 2015). We also used correlation coefficients and root mean square error (RMSE) to assess the performance of our models. RMSE is a measure that presents the average magnitude of the error, and it is the square root of the average of squared differences between the values of prediction yield and actual yield. These analyses were performed using R version 3.44 and WEKA.

### 3.2.5. Input variables

All data mentioned above, including the yield, fertilizer, plant, and soil laboratory data, produced one aggregated file in preparation for the data mining analyses. We assembled two subsets based on seven WR year-seasons without soil parameters and four WR year-seasons with soil parameters. Different variables in the subsets used in the decision tree analysis are given in Table 3.3.

**Table 3.3**. Factors and different variables analyzed for their effects on the grain yield of winter rye by the decision tree model

| Factor and variable name | Unit | Over four growing seasons of WR (with soil information) | | | Over seven growing seasons of WR (without soil information) | | |
|---|---|---|---|---|---|---|---|
| | | Mean | Min | Max | Mean | Min | Max |
| **Fertilizers applied** | | | | | | | |
| 1. Mineral nitrogen | kg ha$^{-1}$ | | 0 | 175 | | 0 | 175 |
| 2. Farmyard manure | tons | | 0 | 3.2 | | 0 | 3.2 |
| 3. Straw | tons | | 0 | 2 | | 0 | 2 |
| **Monthly weather during growing season** | | | | | | | |
| 4. Temperature | °C | 7.96 | -4.7 | 17.8 | 8.2 | -4.7 | 23.4 |
| 5. Cumulative freeze day | day | 4.8 | 0 | 28 | 4.5 | 0 | 28 |
| 6. Days Tmax > 30°C | day | 0.3 | 0 | 3 | 0.5 | 0 | 20 |
| 7. Cumulative precipitation | mm | 46.4 | 0.9 | 136.5 | 45 | 0.9 | 136.5 |
| **Soil** | | | | | | | |
| 8. Total N | mg/100 g soil | 47.4 | 17.8 | 82.4 | | | |
| 9. Total C | mg/100 g soil | 501.7 | 175.8 | 845.8 | | | |
| 10. Plant-available P | mg/100 g soil | 9.5 | 3.9 | 16.3 | | | |
| 11. Plant-available K | mg/100 g soil | 12.5 | 3.8 | 28.2 | | | |
| **Crop yield** | | | | | | | |
| 12. Winter rye | Mg DM ha$^{-1}$ | 6.3 | 1 | 9.7 | 5.6 | 1 | 9.7 |
| 13. Maize | Mg DM ha$^{-1}$ | 13.2 | 1.1 | 29.9 | 12 | 1.1 | 29.9 |
| 14. Flax | Mg DM ha$^{-1}$ | 1.4 | 0.4 | 3.1 | 1.4 | 0.4 | 3.2 |
| **Preceding crop** | | | | | | | |
| 15. Maize | None | | | | | | |
| 16. Flax | None | | | | | | |

Source: own study; winter rye yield is the target variable

### 3.3. Results

### 3.3.1. Winter rye yield and yield variability

#### *3.3.1.1. Winter rye yield*

There was a significant effect of MF and OR on the grain yield of WR. However, the grain yield increased mainly due to an increase of MN-dose but not above 105 kg N $ha^{-1}$ (Figure 3.2, Table 3.4). The WR yield in all fertilizer treatment under four fertilization applications (NPK, fym1, fym2, and straw) was significantly higher than that in the control treatment (Figure 3.2). The effect of organic fertilizer on the yield was remarked at low mineral fertilizer supply. At a low MN rate of 35 kg N $ha^{-1}$ supply, the application of fym1 showed no significant effect on yield while the application of fym2 and straw had a significant effect on yield. At 70 kg N $ha^{-1}$ supply, the application of fym1 and fym2 had a significant effect on yield while the straw application showed no significant effect on yield. The highest WR yields were obtained when 105 kg N $ha^{-1}$ was applied in all four fertilization applications.

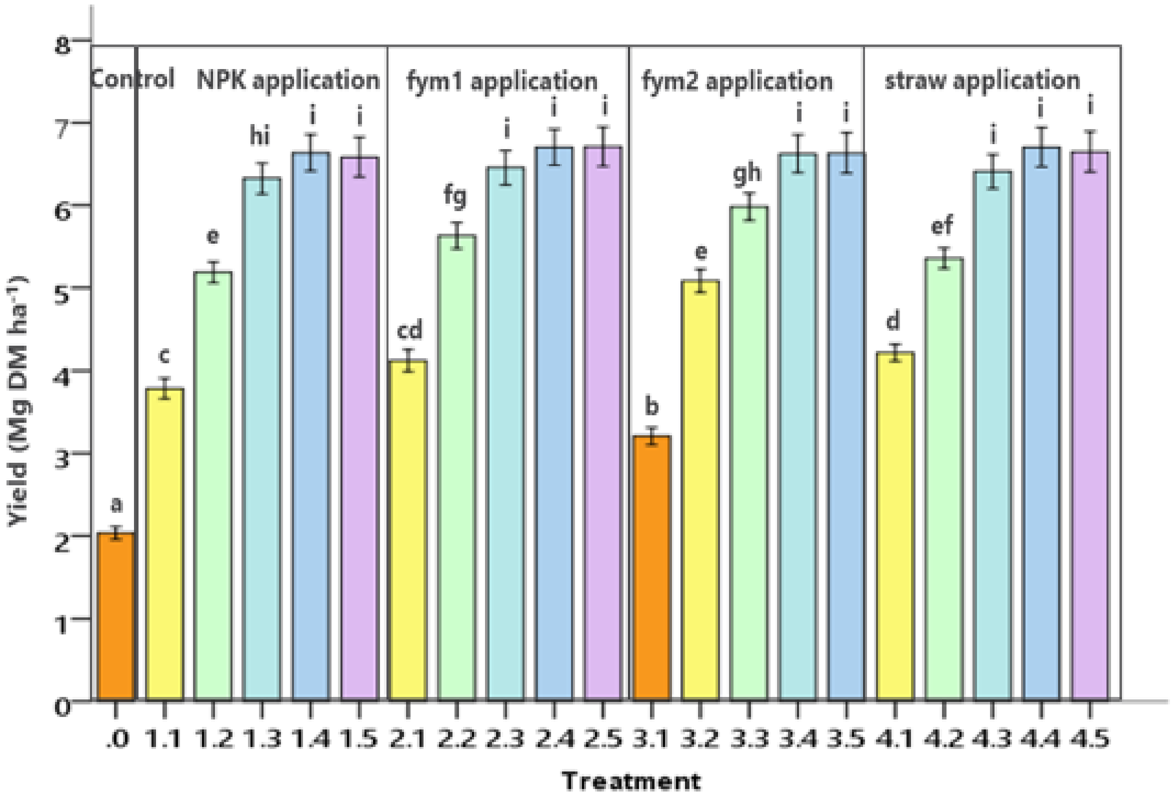



**Figure. 3.2.** Effect of fertilizer on winter rye grain yields through seven growing seasons in every treatment. Vertical lines indicate standard error (SE). A significant difference in winter rye by individual treatment over the year was analyzed by Tukey's test. Treatments sharing the same letter are not significantly different (P<0.05). Treatment numbers are given in Table 3.2. The same color showed the same rate of mineral nitrogen fertilizer supply.

The average WR yield in the most productive treatment (NPK+fym2) was approximately 6.1 t $ha^{-1}$ $a^{-1}$. The average yield in the control treatment was approximately 2 t ha$^{-1}$ a$^{-1}$, which was approximately 67% lower than in the NPK+fym2 treatment (Table 3.4). There was a significantly greater effect of combined applications of MF and OR (NPK+fym1 or NPK+fym2 or NPK+ straw) on WR yields compared with the others. The effect of the combined application resulted in a higher WR yield of approximately 1.8 to 2 times compared with the control treatment, whereas NPK and fym2 increased the WR yield to approximately 1.7 times and 0.6 times compared with the control treatment, respectively. The combined application increased the WR yield up to 0.9 times compared with the sole fym2 treatment and approximately 0.1 times compared with the NPK treatment. The difference in yields was not clear significant between the treatments NPK and NPK + straw. The coefficients of variation (Cvs) and Se value in the WR yields among the different combined applications and NPK applications were higher than those in the control and the sole fym2.

**Table 3.4.** Yield and yield variation of winter rye between group-treatments based on the level of mineral nitrogen and organic fertilizer application in the long-term field experiment.

| Group treatment | Yield (Mg DM ha$^{-1}$) | ±Se | CV | Yield increase compare to control (%) | Yield increase compare to NPK (%) | Yield increase compare to fym2 (%) |
|---|---|---|---|---|---|---|
| Control | 2.03$^{a}$ | 0.14 | 0.18 | - | - 63 | - 37 |
| NPK | 5.48$^{c}$ | 0.42 | 0.20 | 170 | - | 71 |
| fym2 | 3.21$^{b}$ | 0.22 | 0.18 | 58 | - 42 | - |
| NPK+ fym1 | 5.73$^{d}$ | 0.45 | 0.21 | 182 | 5 | 79 |
| NPK + fym2 | 6.08$^{e}$ | 0.53 | 0.23 | 199 | 11 | 90 |
| NPK + straw | 5.67$^{cd}$ | 0.43 | 0.20 | 179 | 3 | 77 |

Group treatments are given in Table 3.2. Mg DM: Mega gram dry mass; Se: standard error; CV: coefficient of variation. Different letters in the same column indicate that the difference is significant at $P < 0.05$.

#### *3.3.1.2. Winter rye yield variability*

The highest yields were observed in treatment NPK+fym2 for most of the cropping years except 1971/72 and 2005/06. The lowest yields were observed in the control ($p<0.05$) for all cropping years (Figure 3.3). Sole fym2 showed a significantly lower yield than all of the other fertilizer treatments, except in the 1971/72 season.

There was a significant yield variability between years. However, the yields in 1997/98 and 2013/14 or 2011/12 and 2013/14 showed similarities. The highest mean yield was observed in 1995/96, and the lowest yield was observed in 1971/72.

The descriptive data analysis (sum of squares, type III) showed that the WR variation was significantly ($P<0.05$) affected by treatment, representing fertilizer application (48%), followed by year, representing the annual weather conditions (32%), and the year × treatment interaction (11%); 9% of the variation was due to error (other factors), and the adjusted R-squared was 0.895 (supplementary Table S 3.3).

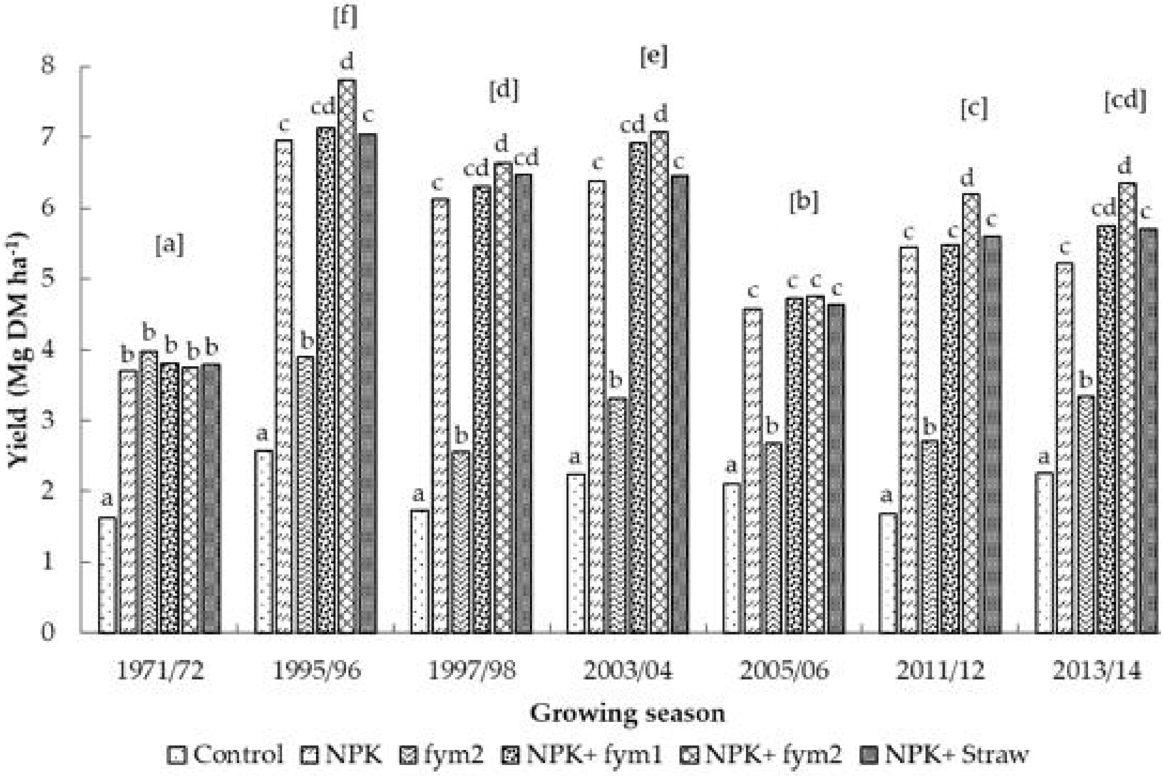



**Figure 3.3.** Winter rye yield (Mg DM $ha^{-1}$) under different fertilizer applications (group treatments) and yield variability across the years. Significant differences for the mean winter rye yield of group treatments or by averaging all treatments (include control) in a certain year were evaluated by Tukey's test. Means sharing the same letters are not significantly different ($P<0.05$). Letters in square brackets at the top of bars compare the mean winter rye yield of all treatments between different years. Letters at the top (without square brackets) of bars compare the mean winter rye yield of group treatments within a year. Treatment groups are given in Table 3.2.

### 3.3.2 Factors driving winter rye yield variability

#### *3.3.2.1. Model 1: Effect of fertilizer, weather, soil, and preceding crop on winter rye yields over 4 cropping years*

The decision tree generated had five splits and 11 terminal nodes (Figure 3.4) and explained 91% of the variability in the data with $R^2 = 0.83$.

The MN supply was the primary splitting node implying that the amount of MN supply was the most important determinant of the WR grain yield variability across four growing seasons. This parameter was divided into two groups: non-MN and the lower level of MN supply (from 0-70 kg ha$^{-1}$) with an average yield of 4.78 Mg DM ha$^{-1}$ and a higher level of MN supply (from 105-175 kg ha$^{-1}$) with an average yield of 7.74 Mg DM ha$^{-1}$.

In the plots that were supplied higher levels of MN, the average temperature in September was the most important determinant of yield. This was also applicable for plots that received lower levels of MN supply (N = 35 kg ha$^{-1}$).

Model 1 also presented higher preceding crop yields had a negative impact on WR yield across all levels of MN supply from 70 to 175 kg ha$^{-1}$ and they were a significant variable explaining the WR yields. Higher of total C had a positive influence on WR yields across levels of MN supply from 70 to 175 kg ha$^{-1}$. It is worthwhile to observe that there were different critical thresholds of total C under different crop management scenarios; for example, the critical threshold for total C was 451.5 mg/100 g soil for plots with 70 kg N supply compared with the threshold of 535.6 mg/100 g soil for plots with a higher MN supply. While this pattern is interesting, it is not within the scope of this research to disentangle and explain such observed pattern. Fym supply was also a significant factor in explaining the WR yields in plots which received 35 kg N ha$^{-1}$.

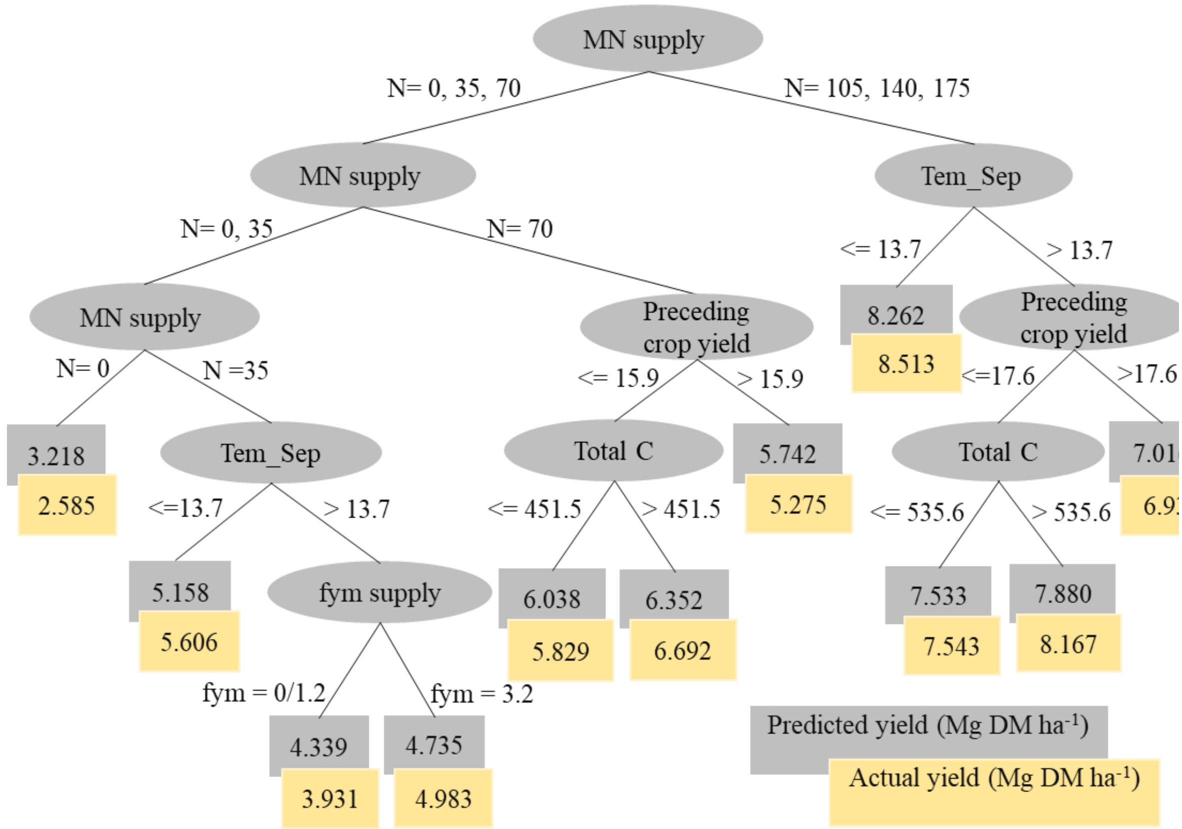



**Figure 3.4.** Decision tree explaining WR grain yield variation in the LTFE over 4 cropping years by fertilizer, weather, soil, and preceding crop variables. The target variable is the grain yield of winter rye. Predicted yield and actual yield values in mega gram dry mass ha$^{-1}$. Tem_Sep: temperature in September

#### *3.3.2.2. Model 2: Effect of fertilizer, weather and preceding crop on winter rye yields over 7 cropping years*

The decision tree generated had five splits and 20 terminal nodes (Figure 3.5) and explained 93% of the variability in the data with $R^2 = 0.86$.

With a similar result in model 1 as in model 2, the MN supply was the primary splitting node, indicating that the amount of MN supply was the most important determinant of the WR grain yield variability across seven growing seasons. This parameter was divided into two groups: the first group was non-MN and a low level of MN supply (0 or 35 kg $ha^{-1}$) with an average yield of 3.74 Mg DM $ha^{-1}$; the second group was a higher level of MN supply (more than 70 kg $ha^{-1}$) with an average yield of 6.31 Mg DM $ha^{-1}$.

In the plots that were supplied with higher levels of MN, the number of days recorded having a maximum temperature above 30°C in July (days Tmax > 30°C in July) was the most important determinant of yield. The application of fym played an important factor in explaining the WR yields only in the plots that received either low MN supply (equal 35 or 70 kg N $ha^{-1}$) or zero MN supply. Plots with zero MN but different levels of fym management interacted with two important weather variables, namely, the monthly precipitation in November and temperature in May which influenced the WR yields across years. Higher November precipitation had a positive impact while higher May temperature had a negative impact on WR yields. In plots with 35 kg N $ha^{-1}$ supply and higher November precipitation than 26.6 mm, different levels of fym management interacted with temperature in October. Furthermore, the preceding crop cultivated and its respective yield as well as the temperature in December, temperature in September were deemed important in explaining the WR yields in plots which received 70 to 175 kg N $ha^{-1}$.

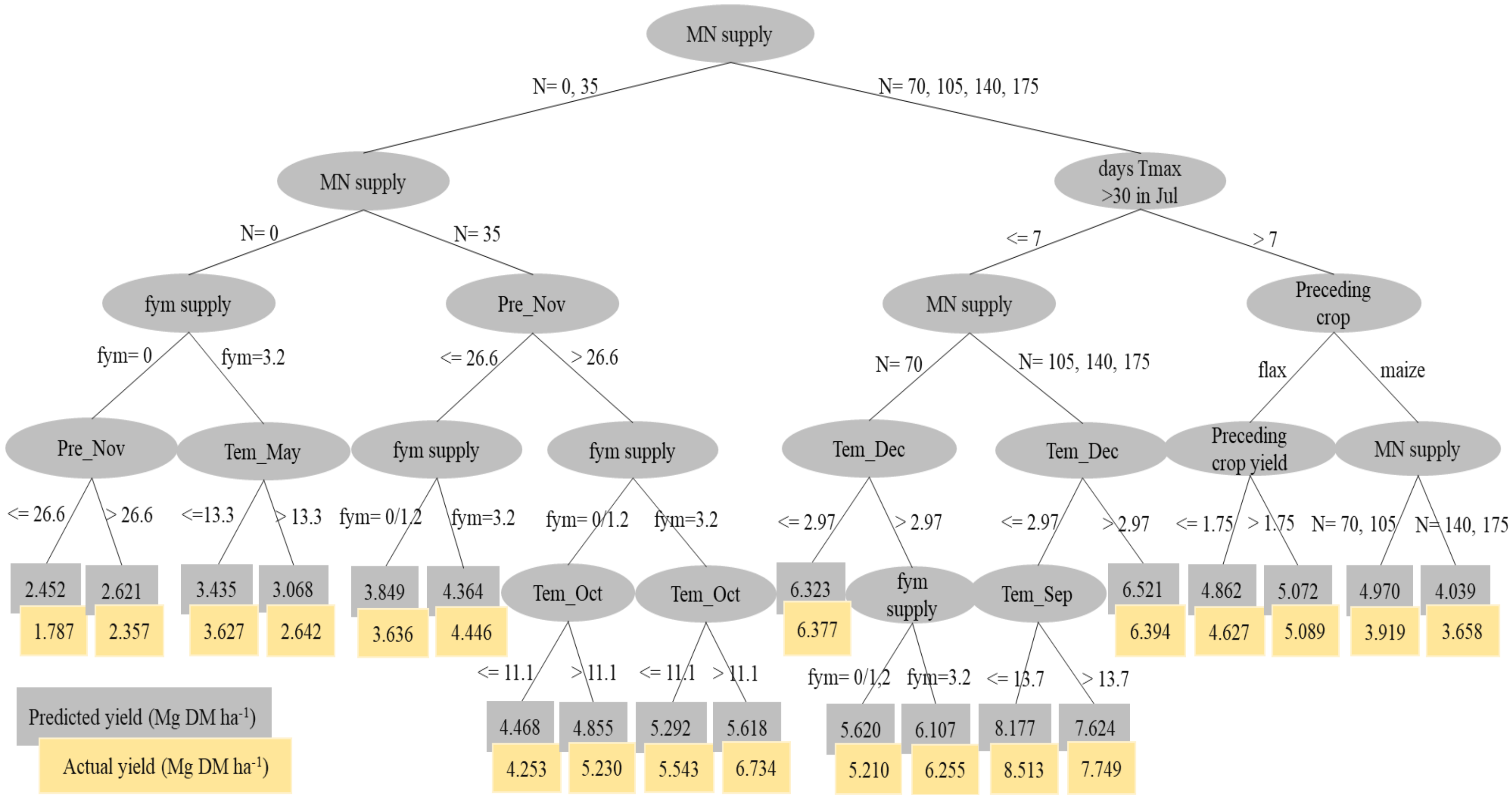



**Figure 3.5.** Decision tree explaining WR grain yield variation in the LTFE over 7 cropping years by fertilizer, weather, and preceding crop variables. The target variable is the grain yield of winter rye. The predicted yield and actual yield values in mega gram dry mass ha$^{-1}$. Days Tmax > 30 °C in July: number of days recorded having a maximum temperature above 30°C in July, Pre_Nov: precipitation in November, Tem_Dec: temperature in December, Tem_Oct: temperature in October, Tem_Sep: temperature in September

The results of the model trees and regression trees for WR grain yield in the V140 LTFE under both scenarios 1 and 2 are presented in Table 3.5 in terms of correlation coefficients (r) and RMSE. The high correlation coefficients for both scenarios suggest that the predictions can be highly reliable. The RMSE in the model tree and regression tree of scenario 2 with approximately 0.6-0.7 t ha$^{-1}$, was smaller than that in scenario 1 with approximately 0.6-0.8 t ha$^{-1}$. Therefore, the predictive performance was better for the models in scenario 2 than those in scenario 1.

**Table 3.5.** Predictive performance in terms of correlation coefficient (r) and root mean square error (RMSE) of the model and regression trees obtained for the WR grain yield in the LTFE

| | | Scenarios | |
|---|---|---|---|
| | | 1 (with soil variables) | 2 (without soil variables) |
| | Coefficient | | |
| Model tree | $R^2$ | 0.88 | 0.89 |
| | RMSE (t ha$^{-1}$) | 0.67 | 0.61 |
| Regression tree | $R^2$ | 0.83 | 0.86 |
| | RMSE (t ha$^{-1}$) | 0.80 | 0.70 |

Source: own study

The results from the decision trees of both scenarios 1 and 2 (Figure 3.4 and Figure 3.5) provide a synthetic assessment of the importance of yield predictors, including different variables of management practices such as fertilizer management, preceding year crop, and environmental conditions such as weather and soil properties.

**Table 3.6.** Ranking of predictors by importance for WR grain yield in the LTFE

| | | Validity category | |
|---|---|---|---|
| **Predictor** | Unit | Scenario 1 | Scenario 2 |
| Mineral nitrogen fertilizer | kg ha$^{-1}$ | Key | Key |
| Days Tmax > 30°C in July | day | | Key |
| Temperature in September | °C | Key | Less important |
| Farmyard manure | tons | Less important | Very important |
| Precipitation in November | mm | | Very important |
| Type of preceding crop | maize, flax | | Very important |
| Preceding crop yield | Mg DM ha$^{-1}$ | Very important | Important |
| Temperature in December | °C | | Important |
| Temperature in May | °C | | Important |
| Total C | mg/100 g soil | Important | |
| Temperature in October | °C | | Less important |

Source: own study (in tree models)

According to the hierarchy split in the trees, the ranking of important predictors that explained WR yield response is presented in Table 3.6, in which MN supply was the most important driving factors for WR yields. The following, temperatures in September, days Tmax > 30°C in July were also critical determinants for WR yields. The preceding crop type, preceding crop yield, fym application, and precipitation in November were very important variables for WR yield response. Also, the temperature in December, temperature in May, and total C in soil were important variables that can explain the WR yields. In addition, the temperature in October can explain the variation in WR yields as well.

## 3.4. Discussion

### 3.4.1. Long-term effect of fertilization regimes on winter rye yield

It is well known that WR has the best nutrient absorption of all grain crops due to its extensive root system. However, as it is mostly grown in locations that are poor in nutrients, a good fertilizer application supply remains an important factor in achieving high yields. In this study, we also found that fertilizer application is one of the important factors governing the WR yield variation (supplementary Table S 3.3).

Similar to the previous study of Thai et al. (2019) for spring barley, in this LTFE we found that WR yield was influenced by both MF application and OR application as well as the combination of MF and OR application during the experimental period. The yield over the years increased with increasing MN supply, but not above optimum 105 kg N ha$^{-1}$. However, the rate of increase in yield differed according to the OR application. The highest combined effect of MF and OR was found for the fym2 application, whereas there were no clear different effects among the other OR applications, such as between fym1 and straw (Table 3.4 and Figure 3.2).

When different fertilizer group treatments were compared, the treatment with sole fym2 resulted in lower yields than the treatment with NPK alone, and the N rate in fym was low and with a slow-release, which led to low utilization efficiency of the plant. The treatment with NPK+fym1 did not have a clear advantage compared with NPK at a low MN rate of 35 kg N ha$^{-1}$ supply, whereas the treatment with NPK + straw did not have a clear advantage compared with NPK at the MN rate of 70 kg N ha$^{-1}$ supply. The reason may be that there was an insufficient amount of fym1 or straw to improve soil fertility in the sandy experimental soil. However, the highest yield was obtained in the NPK+fym2 treatment; this was due not solely to the greater rate of additional fym but also a benefit long-term fym supply. As observed by Mazur and Mazur (2015) and Kulhánek et al. (2014), positive effects on chemical soil properties, such as soil total elements and plant-available nutrient forms, particularly nitrogen, phosphorus, potassium and magnesium, can be assumed to be a result of long-term manure fertilization. This effect was also evident in our study, as shown in the table of selected chemical soil contents in supplement (Table S 3.2). In addition, there are also positive effects of manure on the physical soil and soil biological activity due to long-term organic fertilization, the results of which have also been confirmed by other studies (Barzegar et al., 2002; Holík et al., 2019). These findings were reported by other studies stating that balanced MF application and incorporation of OR can improve crop yield (Wei et al., 2016; Yang et al., 2015). The positive effect of the fym2 application was notable at a lower MN supply rate (N1 and N2) (Figure 3.2). Hence, supplying MN in conjunction with fym2 to WR at low rates can lead to high-efficiency use of fertilizer and higher yields.

### 3.4.2. Factors driving winter rye yield variability

MN has been established as one of the most important nutrients used worldwide to increase and maintain crop production (Fixen and West, 2002). Our findings corroborate this knowledge, and our results show MN supply to be the main factor driving WR yields and yield variability. Plots with MN supply had an average yield approximately 2 times higher than plots with no MN. Furthermore, our study also revealed that the effect of MN supply on WR yields is linear and depends on several other variables and their interactions. Weather conditions, fym supply, preceding year crop, and soil properties form complex associations with MN management and influence WR yields.

Despite being one of the most winter-hardy and drought-resistant crops (Schittenhelm et al., 2014; Starzycki and Bushuk, 1976), WR when grown in marginal locations is vulnerable to drought events (Schlegel, 2013). The studies by Chmielewski (1992) and Chmielewski and Köhn (1999) report that WR yield is influenced by both temperature and precipitation factors. Our results reveal how specific weather variables affect WR yields.

Weather variables in autumn (September to November) were the most important determinants of WR yields in this experiment (Table 3.6). September to November is the sowing phase of the crop, featuring germination, emergence, tillering and initiation of differentiation of the growth apex, and the beginning of spikelet formation, which determines crop density, the number of kernels per ear and subsequent crop yield (Blecharczyk et al., 2016; Chmielewski, 1992; Chmielewski and Köhn, 2000; Meier, 2001). Sufficient moisture levels and moderate temperatures are favorable conditions for growth and development during this phase. Our results suggest that the September temperature was the most important determinant with plots that received an MN supply. This finding was robust across both models, with the plots receiving higher MN supply (105 to 175 kg ha$^{-1}$) recording a higher yield of 8.5 t ha$^{-1}$ (Figure 3.4 and Figure 3.5). We observed that beyond a critical value of 13.7°C, the temperatures in September negatively influence the WR yields despite a similar MN supply. This finding was confirmed through individual data points in the study: in 2005/06 lower mean yields were observed when high September temperatures were recorded versus higher mean yields when lower September temperatures were recorded in 1995/96. Higher temperatures in September could lead to higher evapotranspiration-related water saturation in the soil, leading to soil moisture stress affecting germination and emergence of the plant, especially in conditions such as sandy soils in dry regions such as the location of this LTFE. After the emergence stage, other weather variables continue to influence the growth of WR. Model 2 with excluded soil characteristics revealed detailed interactions between several weather variables and management regimes in this phase. The precipitation in November seemed to be the most important factor influencing yields in plots with unfertilized and low MN supply alongside temperatures in October. November precipitation below the critical threshold of 26.6 mm decreased the average yield by 24% in unfertilized plots and by 21% in plots supplied with 35 kg N $ha^{-1}$. Additionally, higher October temperature than the critical threshold of 11.1°C and higher precipitation in November than 26.6 mm resulted in higher WR yields. This also substantiates the importance of favorable weather conditions for post-germination and emergence in enabling tillering for higher crop density. It also promoted the differentiation of the growth apex and the beginning of spikelet formation and subsequent higher yields (Chmielewski and Köhn, 2000).

As a winter-hardy crop WR can tolerate low temperatures near freezing during the winter rest phase. However, cold winters with dry air masses that lead to high potential evaporation are not favorable for WR (Chmielewski and Köhn, 2000; Starzycki and Bushuk, 1976). Our findings support the above as we observed that the monthly temperatures in December influence WR yields under different fertilizer scenarios.

The days Tmax > 30°C in July was the main factor driving variation of WR yield when applying high MN (70 to 175 kg N $ha^{-1}$) (Figure 3.5, Table 3.6). When days Tmax > 30°C in July exceeded more than 7, the average yield significantly decreased by approximately 38%. Closer to the harvest season, higher temperatures in July influence WR yields. The quantity of available assimilates determines the size and weight of grains after flowering. Thus, the maintenance of the assimilative leaf area as long as possible is a prerequisite for a high kernel weight (Spiertz, 1971). High temperatures and strong evaporative demand in the atmosphere led to accelerated aging of leaves (Römer, 1988) as well as reduced grain filling period and thus a low kernel weight.

Organic fertilizer management practices such as the supply of fym and straw, organic fertilizers are important for soil health and influence yields over the long run. Our results show that fym application has significant effects on WR yields on plots with zero or low MN supply. In the zero-MN supply plots, higher fym application resulted in a higher average yield (up to 58%) (Figure 3.5). However, this yield increase from higher fym varied with the critical temperatures observed in May. When the temperature in May was lower than 13.3°C, average yields were higher by 37%. This again highlights moisture stress or higher evaporative demand leading to a reduction in yield.

The preceding crop type and the preceding crop yields are considered important factors that determine yield and yield variability in LTFEs (Hejcman and Kunzova, 2010; Kunzová and Hejcman, 2009, 2010). Our results also support this; preceding year´s crop and their yields were important variables that explained the WR yield when modeled alongside soil variables or exclusion of soil variable (Figure 3.4 and Figure 3.5). In plots with the preceding crop as maize with higher yield than preceding crop as flax recorded lower successive crop yields. Maize as a preceding crop decreased the WR yields by 23%. In contrast, flax as a preceding crop increased the yield by 30%. The preceding crop type and their yields resulted in nutrients and moisture uptake in soil, which related the growth of WR as a succeeding crop.

Soil characteristics and nutrients are significant yield-limiting factors. Fertilizer management influences crop yields by improving soil nutrients and fertility. The soil-related variables during the WR period of the experiment showed in Table S 3.2 in the supplementary material. One study by Pasley et al. (2019) reported that MN enhances the uptake of non-N nutrients and increases the soil availability of essential nutrients. Our findings support this finding. In plots that received from 70-175 kg $ha^{-1}$ MN supply, total C presented as prediction variable was an important variable explaining the WR yield response across years. There was a positive influence of total C on the WR yield with the yield increasing by 8% at a critical threshold or 535.6 Ct mg/100 g soil and 5% at a critical threshold or 451.5 Ct mg/100 g soil. While the critical thresholds for weather variables remained the same across both models and across different splits and nodes (i.e., different levels of MN supply), the soil variable revealed different critical thresholds against different levels of MN supply. This reflects the complexity of soil dynamics and their influence on crop yields.

Overall, our findings suggest that crop yield response to fertilizer management is not linear and depends on complex interactions between weather, soil fertility and other management practices. This implies that strategies to improve yields or reduce year-to-year yield variability through crop management must consider the impact of fertilizer management, climatic dependence, crop rotation, and soil characteristics. The results presented here indicated that a suitable combined application of MF and fym is an effective way to obtain higher WR yields than the application of either MF alone or sole fym. In addition, flax as a preceding crop was a good example in crop rotation for improving the WR yields. The importance of seasonal weather forecasts in aiding management decisions is strongly acknowledged (Girma et al., 2007). Notable in this study was the weather condition in the early stage of the growing season (temperature and precipitation in autumn) and weather conditions in the summertime (temperature in May and July). Furthermore, fertilizer management practices such as supplying organic fertilizer is considered a good way to enhance nutrients and reduce water stress in sandy soil. Similarly, precision agriculture is gaining traction in terms of defining a support system for agriculture practice with the goal of increasing crop yields and optimizing returns on inputs while preserving resources. Decision tree methods such as the one used in this study can help in identifying non-linear relationships and critical thresholds for both environmental variables (soil, weather variables) and management practice variables (fertilization regimes, crop rotation variables) which can be useful to providing customized agronomic recommendations in conjunction with seasonal weather forecasts to agencies and farmers. However, it is also important to note that the techniques we employ here are not without caveats. Overfitting is a pervasive and hard problem to solve, especially in a small data-sets such as the one we present in our study. And although the regression coefficients remain unaffected by serial correlation, standard errors may be underestimated (and corresponding significances overestimated) in our study (Durbin and Watson, 1950); we acknowledge this and carefully interpret our results.

The results from the decision tree models in this study explain the variability observed for the cereal crop yield responses in the V140 LTFE better than the statistical study previously conducted by Thai et al. (2019) on the same data. The results indicated that the data from V140 can be organized into an understandable and intuitive structure that highlights the interactions and critical thresholds of explanatory variables. Concurrently, the models demonstrated high predictive performance for the crop yield in different agricultural practice scenarios. This is consistent with the findings from Trajanov et al. (2019), Zheng et al. (2009) and Lobell et al. (2005) who reported that the decision tree model generated was very suitable and reliable for predicting primary productivity in a LTFE in Austria, soybean yield in Northeast China and wheat yield variation in Yaqui Valley, Mexico, respectively.

### 3.5. Conclusions

The winter rye grain yield response to fertilizer management involves complex relationships among climatic dependence, crop rotation (preceding year´s crop), and soil characteristics. Decision tree model by M5P machine learning algorithm has superior predictive performance as a further analysis after general linear model to (i) unravel linear, non-linear interactions and combined effects in a complex dataset such as V140 LTFE, and (ii) identify critical thresholds of explanatory the variables and their influence on winter rye yields. Our results reveal that the most important variable effect on winter rye yields were mineral nitrogen application. The following was weather condition in the early stage of the growing season (in autumn), especially the temperature in September and precipitation in November; and the intensity and duration of extreme temperatures in the summertime (harvest year), especially number of days recorded having temperatures above 30°C in July and the temperature in May. Additionally, farmyard manure application and the preceding year´s crop were also significant variables explaining the yield variability of winter rye. Soil parameter, in particular total carbon although less present in the model than other but also involved a role that influence WR yield variation once NF application more than 70kg$^{-1}$ and temperature in September higher than 13.7°C. This finding implies that strategies to improve yields or reduce year-to-year yield variability of winter rye in sandy soil must consider the role of supplying mineral nitrogen and farmyard manure fertilizer which relates to enhancing nutrients and reducing water stress in the soil. Furthermore, seasonal weather forecasts are important in adjusting the crop management strategy. Flax as a preceding crop was considered a good example in crop rotation to support improved winter rye yields.

Due to their ability to represent complex relationships in a visually simple yet powerful way, decision tree by algorithm M5P are useful supplementary tools for agronomists to devise different crop management intervention strategies such as fertilizer regimes, crop rotation to adapt to fluctuating weather conditions and dynamic soil fertility parameters over time. However, these techniques also pose considerable challenges to model fitting and subsequent interpretation. To meet the challenge of climate change, LTFE data should be analyzed more in detail by further statistical methods to devise suitable suggestions to support researcher-farmer-advisor dialogue on productivity management and the development and adoption of precision agriculture recommendations.

**Acknowledgments**

This work was financially supported by the Vietnamese Government through the 911 project. We would like to thank the Müncheberg Experimental Station, ZALF, for providing the data that were used in this analysis. The authors thank Dr. Carsten Hoffmann for his useful comments on this paper.

**References**

Azzeh, M., 2011. Software effort estimation based on optimized model tree. In "Proceedings of the 7th International Conference on Predictive Models in Software Engineering", pp. 1-8.

Barzegar, A., Yousefi, A., and Daryashenas, A., 2002. The effect of addition of different amounts and types of organic materials on soil physical properties and yield of wheat. Plant Soil. 247, 295-301.

Behnood, A., Behnood, V., Gharehveran, M. M., and Alyamac, K. E., 2017. Prediction of the compressive strength of normal and high-performance concretes using M5P model tree algorithm. Constr and Build Mater. 142, 199-207.

Blecharczyk, A., Sawinska, Z., Małecka, I., Sparks, T. H., and Tryjanowski, P., 2016. The phenology of winter rye in Poland: an analysis of long-term experimental data. Int. J. of biometeorol. 60, 1341-1346.

Blockeel, H., and Struyf, J., 2002. Efficient algorithms for decision tree cross-validation. J Mach Learn Res. 3, 621-650.

Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., and White, J.-S. S., 2009. Generalized linear mixed models: a practical guide for ecology and evolution. Trends Ecol. Evol. 24, 127-135.

Breiman, L., Friedman, J., Olshen, R., and Stone, C., 1984. Classification and regression trees. Wadsworth Int. Group. 37, 237-251.

Brisson, N., Gate, P., Gouache, D., Charmet, G., Oury, F.-X., and Huard, F., 2010. Why are wheat yields stagnating in Europe? A comprehensive data analysis for France. Field Crops Res. 119, 201-212.

Chmielewski, F.-M., 1992. Impact of climate changes on crop yields of winter rye in Halle (southeastern Germany), 1901 to 1980. Clim. Res., 23-33.

Chmielewski, F.-M., and Köhn, W., 1999. The long-term agrometeorological field experiment at Berlin-Dahlem, Germany. Agricultural and Forest Meteorology 96, 39-48.

Chmielewski, F. M., and Köhn, W., 2000. Impact of weather on yield components of winter rye over 30 years. Agr Forest Meteorol. 102, 253-261.

Dacko, M., Zając, T., Synowiec, A., Oleksy, A., Klimek-Kopyra, A., and Kulig, B., 2016. New approach to determine biological and environmental factors influencing mass of a single pea (Pisum sativum L.) seed in Silesia region in Poland using a CART model. Eur J Agron. 74, 29-37.

De'ath, G., and Fabricius, K. E., 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. Ecology 81, 3178-3192.

Delmotte, S., Tittonell, P., Mouret, J.-C., Hammond, R., and Lopez-Ridaura, S., 2011. On farm assessment of rice yield variability and productivity gaps between organic and conventional cropping systems under Mediterranean climate. Eur J Agron. 35, 223-236.

Durbin, J., and Watson, G. S., 1950. Testing for serial correlation in least squares regression: I. Biometrika 37, 409-428.

DWD, S. (2019). DWD station number 03376 via the link https://opendata.dwd.de/climate_environment/CDC/observations_germany/climate/daily/kl/historical.

Ellerbrock RH, Höhn A, Rogasik J. 1999. Functional analysis of soil organic matter as affected by long-term manurial treatment. Eur J Soil Sci. 50:65-71.

FAO, 2019. FAOSTAT online database, available at link http://faostat.fao.org/. Accessed on October 2019.

Fixen, P. E., and West, F. B. (2002). Nitrogen fertilizers: meeting contemporary challenges. AMBIO: J Hum. Environ. 31, 169-177.

Gauch, H. G., 2006. Statistical analysis of yield trials by AMMI and GGE. Crop sci. 46, 1488-1500.

Girma, K., Holtz, S. L., Arnall, D. B., Fultz, L. M., Hanks, T. L., Lawles, K. D., Mack, C. J., Owen, K. W., Reed, S. D., and Santillano, J., 2007. Weather, fertilizer, previous year yield, and fertilizer levels affect ensuing year fertilizer response of wheat. Agronomy J. 99, 1607-1614.

Gregory, P. J., Johnson, S. N., Newton, A. C., and Ingram, J. S., 2009. Integrating pests and pathogens into the climate change/food security debate. J . Exp. Bot. 60, 2827-2838.

Hastie, T., Tibshirani, R., and Friedman, J., 2009. "The elements of statistical learning: data mining, inference, and prediction," Springer Science & Business Media.

Hejcman, M., and Kunzova, E., 2010. Sustainability of winter wheat production on sandy-loamy Cambisol in the Czech Republic: Results from a long-term fertilizer and crop rotation experiment. Field Crops Res. 115, 191-199.

Holík, L., Hlisnikovský, L., Honzík, R., Trögl, J., Burdová, H., and Popelka, J., 2019. Soil Microbial Communities and Enzyme Activities after Long-Term Application of Inorganic and Organic Fertilizers at Different Depths of the Soil Profile. Sustainability 11, 3251.

Krupnik, T. J., Ahmed, Z. U., Timsina, J., Yasmin, S., Hossain, F., Al Mamun, A., Mridha, A. I., and McDonald, A. J., 2015. Untangling crop management and environmental influences on wheat yield variability in Bangladesh: an application of non-parametric approaches. Agric. Syst. 139, 166-179.

Kulhánek, M., Balík, J., Černý, J., Vašák, F., and Shejbalová, Š., 2014. Influence of long-term fertilizer application on changes of the content of Mehlich-3 estimated soil macronutrients. Plant Soil Environ. 60, 151-157.

Kunzová, E., and Hejcman, M., 2009. Yield development of winter wheat over 50 years of FYM, N, P and K fertilizer application on black earth soil in the Czech Republic. Field Crops Res. 111, 226-234.

Kunzová, E., and Hejcman, M., 2010. Yield development of winter wheat over 50 years of nitrogen, phosphorus and potassium application on greyic Phaeozem in the Czech Republic. Eur J Agron. 33, 166-174.

Lobell, D. B., Ortiz-Monasterio, J. I., Asner, G. P., Naylor, R. L., and Falcon, W. P., 2005. Combining field surveys, remote sensing, and regression trees to understand yield variations in an irrigated wheat landscape. Agronomy J. 97, 241-249.

Loh, W.-Y., 2006. Regression tree models for designed experiments. In "Optimality", pp. 210-228. Institute of Mathematical Statistics.

Loh, W. Y., 2011. Classification and regression trees. Wires Data Min Knowl. 1, 14-23.

Mazur, Z., and Mazur, T., 2015. Influence of Long-Term Fertilization on Phosphorus, Potassium, Magnesium, and Sulfur Content in Soil. Pol. J. Environ. Stud. 24.

Meier, U., 2001. BBCH-monograph: growth stages of mono-and dicotyledonous plants. Federal Biological Research Centre for Agriculture and Forestry, 130-133.

Miedaner, T., Koch, S., Seggl, A., Schmiedchen, B., and Wilde, P., 2012. Quantitative genetic parameters for selection of biomass yield in hybrid rye. Plant breed. 131, 100-103.

Pasley, H. R., Cairns, J. E., Camberato, J. J., and Vyn, T. J., 2019. Nitrogen fertilizer rate increases plant uptake and soil availability of essential nutrients in continuous maize production in Kenya and Zimbabwe. Nutr. Cycl. Agroecosys. 115, 373-389.

Patel, N., and Upadhyay, S., 2012. Study of various decision tree pruning methods with their empirical comparison in WEKA. Int. J. Comput. Appl. 60.

Quinlan, J. R., 1992. Learning with continuous classes. In "5th Australian joint conference on artificial intelligence", Vol. 92, pp. 343-348. World Scientific.

Römer, G., 1988. Die phänologischen Phasen bei Hafer, Sommergerste, Winterroggen und Mais und ihre Beziehungen zur Witterung und Ertragsbildung, Institut für Nutzpflanzenforscchung der Technischen Universität Berlin.

Schittenhelm, S., Kraft, M., and Wittich, K.-P., 2014. Performance of winter cereals grown on field-stored soil moisture only. Eur J Agron. 52, 247-258.

Schlegel, R. H., 2013. "Rye: genetics, breeding, and cultivation," Crc Press.

Sileshi, G., Akinnifesi, F. K., Debusho, L. K., Beedy, T., Ajayi, O. C., and Mong'omba, S., 2010. Variation in maize yield gaps with plant nutrient inputs, soil type and climate across sub-Saharan Africa. Field Crops Res. 116, 1-13.

Silungwe, F. R., Graef, F., Bellingrath-Kimura, S. D., Tumbo, S. D., Kahimba, F. C., and Lana, M. A., 2018. Crop upgrading strategies and modelling for rainfed cereals in a semi-arid climate-A review. Water 10, 356.

Silungwe, F. R., Graef, F., Bellingrath-Kimura, S. D., Tumbo, S. D., Kahimba, F. C., and Lana, M. A., 2019. Analysis of intra and interseasonal rainfall variability and its effects on pearl millet yield in a semiarid agroclimate: Significance of scattered fields and tied ridges. Water 11, 578.

Song, Y.-Y., and Ying, L., 2015. Decision tree methods: applications for classification and prediction. Shanghai Arch. Psychiatry 27, 130.

Spiertz, J., 1971. Relation between green area duration and grain yield in some varieties of spring wheat. Neth. J. agr. Sci. 19, 211-222.

Starzycki, S., and Bushuk, W., 1976. Diseases, pests, and physiology of rye. Rye: Production, Chemistry, and Technology. St. Paul, MN: American Association of Cereal Chemists, 27-61.

Tack, J., Barkley, A., and Nalley, L. L., 2015. Effect of warming temperatures on US wheat yields. Proc. Natl. Acad. Sci. 112, 6931-6936.

Thai, T. H., Bellingrath-Kimura, S. D., Hoffmann, C., and Barkusky, D., 2019. Effect of long-term fertilizer regimes and weather on spring barley yields in sandy soil in North-East Germany. Arch. Agron. Soil Sci., 1-15.

Trajanov, A., Spiegel, H., Debeljak, M., and Sandén, T., 2019. Using data mining techniques to model primary productivity from international long-term ecological research (ILTER) agricultural experiments in Austria. Reg. Environ. Change. 19, 325-337.

Vagh, Y., and Xiao, J., 2012. A Data Mining Perspective of the Dual Effect of Rainfall and Temperature on Wheat Yield. Int J Comput Commun. 1, 358-364.

Wei, W., Yan, Y., Cao, J., Christie, P., Zhang, F., and Fan, M., 2016. Effects of combined application of organic amendments and fertilizers on crop yield and soil organic matter: An integrated analysis of long-term experiments. Agric. Ecosyst. Environ. 225, 86-92.

Witten, I. H., Frank, E., Hall, M. A., and Pal, C. J., 2016. "Data Mining: Practical machine learning tools and techniques," Morgan Kaufmann.

Wu, C. J., and Hamada, M. S., 2011. "Experiments: planning, analysis, and optimization," John Wiley & Sons.

Yang, J., Gao, W., and Ren, S., 2015. Long-term effects of combined application of chemical nitrogen with organic materials on crop yields, soil organic carbon and total nitrogen in fluvo-aquic soil. Soil Till Res. 151, 67-74.

Zhang, D., 2006. "Advances in machine learning applications in software engineering," Igi Global.

Zhang, J., Liu, Q., Xu, M., and Zhao, B., 2012. Effects of soil properties and agronomic practices on wheat yield variability in Fengqiu County of North China Plain. Afr. J. Agric. Res. 7, 1650-1658.

Zheng, H., Chen, L., Han, X., Ma, Y., and Zhao, X., 2010. Effectiveness of phosphorus application in improving regional soybean yields under drought stress: a multivariate regression tree analysis. Afr. J. Agric. Res. 5, 3251-3258.

Zheng, H., Chen, L., Han, X., Zhao, X., and Ma, Y., 2009. Classification and regression tree (CART) for analysis of soybean yield variability among fields in Northeast China: The importance of phosphorus application rates under drought conditions. Agric Ecosyst Environ. 132, 98-105.

**Supplement**

**Table S 3.1.** Crop rotation of the long-term experiment "V140". In bold: Winter rye, in grey: preceding crops.

| Harvest year | Crop | Year | Crop | Harvest year | Crop |
|---|---|---|---|---|---|
| 1963 | Maize | 1981 | Sugar beet | 1999 | Potato |
| 1964 | Winter rye | 1982 | Spring barley | 2000 | Spring barley |
| 1965 | Potato | 1983 | Potato | 2001 | Pea |
| 1966 | Winter rye | 1984 | Winter wheat | 2002 | Winter wheat |
| 1967 | Potato | 1985 | Sugar beet | 2003 | Maize |
| 1968 | Summer wheat | 1986 | Spring barley | 2004 | **Winter rye** |
| 1969 | Sugar beet | 1987 | Potato | 2005 | Flax |
| 1970 | Spring barley | 1988 | Winter wheat | 2006 | **Winter rye** |
| 1971 | Maize | 1989 | Sugar beet | 2007 | Potato |
| 1972 | **Winter rye** | 1990 | Spring barley | 2008 | Spring barley |
| 1973 | Potato | 1991 | Potato | 2009 | Pea |
| 1974 | Winter wheat | 1992 | Winter wheat | 2010 | Winter wheat |
| 1975 | Sugar beet | 1993 | Sugar beet | 2011 | Maize |
| 1976 | Spring barley | 1994 | Winter wheat | 2012 | **Winter rye** |
| 1977 | Sugar beet | 1995 | Maize | 2013 | Flax |
| 1978 | Spring barley | 1996 | **Winter rye** | 2014 | **Winter rye** |
| 1979 | Sugar beet | 1997 | Flax | 2015 | Potato |
| 1980 | Spring barley | 1998 | **Winter rye** | 2016 | Spring barley |

Thai et al. (2019) adapted

The position of winter rye within the crop rotation was changed during the long-term experiment but there were four rounds of this crop rotation "maize-winter rye" and three rounds of this "flax-winter rye". Between 1973- 1994 no winter rye was seeded.

**Table S3.2** Average content of selected chemical properties of the top soil (0-25 cm) in each treatment through four winter rye seasons (1996, 1998, 2004, and 2012).

| Treatment | pH (KCl) | Total N (mg/100g soil) | Total C (mg/100g soil) | Plant-available P (mg/100g soil) | Plant-available K (mg/100g soil) | Plant-available Mg (mg/100g soil) |
|---|---|---|---|---|---|---|
| 0 | $6.03^{a}$ | $40.2^{a}$ | $417^{a}$ | $6.6^{a}$ | $8.1^{a}$ | $5.6^{b-e}$ |
| 1.1 | $6.04^{a}$ | $42.0^{ab}$ | $446^{ab}$ | $9.7^{efg}$ | $13.2^{f-j}$ | $5.3^{a-d}$ |
| 1.2 | $5.98^{a}$ | $42.9^{abc}$ | $466^{abc}$ | $9.1^{b-g}$ | $12.0^{c-h}$ | $5.3^{b-e}$ |
| 1.3 | $5.96^{ab}$ | $43.3^{a-d}$ | $452^{a-d}$ | $8.3^{bcd}$ | $11.0^{b-e}$ | $5.0^{a-d}$ |
| 1.4 | $5.82^{b}$ | $44.0^{a-d}$ | $467^{a-d}$ | $7.8^{ab}$ | $10.1^{abc}$ | $5.0^{a-d}$ |
| 1.5 | $5.65^{c}$ | $44.6^{a-d}$ | $473^{a-d}$ | $8.5^{b-e}$ | $9.7^{ab}$ | $4.5^{a-d}$ |
| 2.1 | $6.02^{a}$ | $45.6^{a-d}$ | $483^{a-d}$ | $10.0^{fgh}$ | $13.8^{g-j}$ | $5.5^{b-e}$ |
| 2.2 | $6.98^{a}$ | $47.0^{b-e}$ | $498^{b-f}$ | $9.6^{d-g}$ | $13.3^{f-j}$ | $5.3^{b-e}$ |
| 2.3 | $5.82^{b}$ | $50.0^{d-h}$ | $519^{c-g}$ | $9.4^{d-g}$ | $12.6^{d-i}$ | $5.2^{a-d}$ |
| 2.4 | $5.86^{ab}$ | $48.3^{b-g}$ | $524^{d-g}$ | $9.2^{c-g}$ | $11.0^{bcd}$ | $5.0^{a-d}$ |
| 2.5 | $5.70^{bc}$ | $49.8^{d-h}$ | $504^{b-f}$ | $8.9^{b-f}$ | $10.9^{bcd}$ | $4.8^{ab}$ |
| 3.1 | $6.06^{a}$ | $49.4^{c-h}$ | $518^{c-g}$ | $11.9^{ij}$ | $16.4^{k}$ | $5.6^{b-e}$ |
| 3.2 | $6.08^{a}$ | $56.1^{h}$ | $583^{g}$ | $12.6^{j}$ | $16.8^{k}$ | $6.1^{e}$ |
| 3.3 | $6.04^{a}$ | $53.0^{e-h}$ | $553^{efg}$ | $11.2^{hi}$ | $15.0^{jk}$ | $5.7^{de}$ |
| 3.4 | $5.97^{ab}$ | $54.1^{fgh}$ | $577^{g}$ | $11.2^{hi}$ | $13.9^{hij}$ | $5.7^{cde}$ |
| 3.5 | $5.83^{b}$ | $54.7^{gh}$ | $558^{fg}$ | $10.3^{gh}$ | $13.1^{f-i}$ | $5.4^{b-e}$ |
| 4.1 | $6.03^{a}$ | $44.5^{a-d}$ | $487^{b-e}$ | $9.2^{c-g}$ | $14.1^{ij}$ | $5.5^{b-e}$ |
| 4.2 | $5.99^{a}$ | $44.8^{a-d}$ | $483^{a-d}$ | $9.2^{c-g}$ | $12.9^{e-i}$ | $5.2^{a-d}$ |
| 4.3 | $5.93^{ab}$ | $47.2^{bcd}$ | $519^{c-g}$ | $9.2^{d-g}$ | $11.9^{c-g}$ | $5.2^{a-d}$ |
| 4.4 | $5.85^{b}$ | $46.5^{a-e}$ | $508^{b-f}$ | $8.7^{b-e}$ | $11.4^{b-f}$ | $5.0^{a-d}$ |
| 4.5 | $5.60^{c}$ | $47.4^{b-f}$ | $502^{b-f}$ | $7.9^{bc}$ | $10.6^{bc}$ | $4.9^{abc}$ |

Treatments (1.1 - 4.5) are each rate of mineral fertiliser (base on five N rates) combined with each organic fertiliser (four variants: no organic, fym1, fym2, straw). Treatment "0": control, no fertilisation. Total nitrogen: total N, total carbon: total C, plant-available phosphorus: plant-available P, plant-available potassium: plant-available K, plant-available magnesium: plant-available Mg. A significant difference in each soil element by individual treatment over the year was analysed by Tukey's test. Means sharing the same letters in the same column are not significantly different ($P<0.05$). Treatments are given in Table 3.2.

Table S 3.2 shows the average values of the selected chemical properties of the topsoil (0-25 cm) in each treatment through four winter rye seasons in the experimental period.

The pH (KCl) value and Plant-available Mg content in Table S 3.2 did not show clear differences between fertilizer plots and control. This could be explained by the fact that all the plots were uniformly applied with lime.

Compared to control plot, soil elements such as total N and total C, plant-available P, plant-available K were significantly improved in almost combination fertilizer treatments (mineral and organic fertilizer) and in sole farmyard manure treatment (3.1). Moreover, plant-available P, plant-available K also were improved in NPK treatments (treatment 1.1-1.5) compared to control. Noticeably, the values of these soil elements were much higher in treatments amended with high farmyard manure.

**Table S 3.3**

Results from analysis of variance (ANOVA) with Eta squared between fertilizers, year (annual weather) and winter rye yields.

| Source | Type III Sum of Squares | df | Mean Square | F | Sig. | Eta Squared ($h^2$) |
|---|---|---|---|---|---|---|
| Corrected Model | 3947.303a | 146 | 27.036 | 69.357 | 0.000 | |
| Intercept | 36526.374 | 1 | 36526.374 | 93702.696 | 0.000 | |
| Treatment | 2089.127 | 20 | 104.456 | 267.966 | 0.000 | 48 |
| Year | 1401.034 | 6 | 233.506 | 599.022 | 0.000 | 32 |
| Treatment * Year | 457.142 | 120 | 3.810 | 9.773 | 0.000 | 11 |
| Error | 401.116 | 1029 | 0.390 | | | 9 |
| Total | 40874.792 | 1176 | | | | |
| Corrected Total | 4348.419 | 1175 | | | | |

a. R Squared = 0.908 (Adjusted R Squared = 0.895)
Dependent variable: grain yield of winter rye (t $ha^{-1}$)

The Table S 3.3 shows result of the two-way ANOVA – namely, whether either of the two independent variables (treatment and year) or their interaction are statistically significant. The "Sig." column showed that the treatment, year and their interaction have a statistically significant effect on grain yield of winter rye. Besides, “Eta squared ($h^2$)” column showed proportion of total variance that is attributed to an effect. In this case, the yields of WR were significantly affected by treatment, representing fertilizer application (48%), affected by year, representing annual weather conditions (32%), and the year × treatment interaction (11%); 9% of the variation was due to error (other factors), and the adjusted R-squared was 0.895.

# Chapter 4

## Statistical analysis versus the M5P machine learning algorithm to analyze the yield of winter wheat in a long-term fertilizer experiment

Thi Huyen Thai, Richard Ansong Omari, Dietmar Barkusky, Sonoko Dorothea Bellingrath-Kimura

*Agronomy (published. DOI: 10.3390/agronomy10111779)*

# Statistical analysis versus the M5P machine learning algorithm to analyze the yield of winter wheat in a long-term fertilizer experiment

Thi Huyen Thai [1,2,3,*], Richard Ansong Omari [1,2], Dietmar Barkusky [1] and Sonoko Dorothea Bellingrath-Kimura [1,2,*]

*[1]Leibniz Centre for Agricultural Landscape Research (ZALF), Eberswalder Str. 84, 15374 Müncheberg, Germany; Richard.Omari@zalf.de (R.A.O.); dbarkusky@zalf.de (D.B.)*

*[2]Faculty of Life Science, Humboldt University of Berlin, Albrecht-Thaer-Weg 5, 14195 Berlin, Germany*

*[3]Faculty of Agronomy, University of Agriculture and Forestry, Hue University, 102 Phung Hung street, 530000 Hue, Vietnam*

**Correspondence: thaihuyen@hueuni.edu.vn (T.H.T.); belks@zalf.de (S.D.B.-K.); Tel.: +49-(0)-33432-82-313*

**Abstract:** To compare how different analytical methods explain crop yields from a long-term field experiment (LTFE), we analyzed the grain yield of winter wheat (WW) under different fertilizer applications in Müncheberg, Germany. An analysis of variance (ANOVA), linear mixed-effects model (LMM), and MP5 regression tree model were used to evaluate the grain yield response. All the methods identified fertilizer application and environmental factors as the main variables that explained 80% of the variance in grain yields. Mineral nitrogen fertilizer (NF) application was the major factor that influenced the grain yield in all methods. Farmyard manure slightly influenced the grain yield with no NF application in the ANOVA and M5P regression tree. While sources of environmental factors were unmeasured in the ANOVA test, they were quantified in detail in the LMM and M5P model. The LMM and M5P model identified the cumulative number of freezing days in December as the main climate-based determinant of the grain yield variation. Additionally, the temperature in October, the cumulative number of freezing days in February, the yield of the preceding crop, and the total nitrogen in the soil were determinants of the grain yield in both models. Apart from the common determinants that appeared in both models, the LMM additionally showed precipitation in June and the cumulative number of days in July with temperatures above 30°C, while the M5P model showed soil organic carbon as an influencing factor of the grain yield. The ANOVA results provide only the main factors affecting the WW yield. The LMM had a better predictive performance compared to the M5P, with smaller root mean square and mean absolute errors. However, they were richer regressors than the ANOVA. The M5P model presented an intuitive visualization of

important variables and their critical thresholds, and revealed other variables that were not captured by the LMM model. Hence, the use of different methods can strengthen the statement of the analysis, and thus, the co-use of the LMM and M5P model should be considered, especially in large databases involving multiple variables.

**Keywords:** winter wheat yield; long-term field experiment; fertilizer; weather; linear mixed-effects models; M5P machine learning algorithm

### 4.1. Introduction

Winter wheat (WW) (Triticum aestivum L.) is an important cereal in Europe and accounts for over 32% of the total global production next to Asia [1]. In Germany, WW covers 3.2 million hectares, accounting for around one-third of the total arable land area. The average total WW production from 2014 to 2018 was 24.7 million tons, with an average yield of 7.7 t ha$^{-1}$ [2]. The grain yield of WW in Germany has increased in recent decades from an average of less than 3 t ha$^{-1}$ in the 1960s to around 8 t ha$^{-1}$ in the 2000s [1]. However, the grain yield of WW has fluctuated in recent years. Apart from crop breeding improvement, which has contributed dramatically to the wheat yield increase throughout the 20$^{th}$ century in Germany [3,4], several other factors, such as enhanced agronomic management, favorable weather conditions, and soil improvement, also played an important role in yield development and yield stability [5,6]. Thus, similar to other crops, yield variation in WW is the result of interdependencies and complex interactions among different factors. In this regard, identifying the major factors and their relationships that account for grain yield variation of WW is crucial to understanding how to maximize yields and minimize annual yield fluctuations each year.

Long-term field experiments (LTFEs) provide insight to unravel the factors that influence crop yield dynamics in different cropping systems and thus serve as a means to assess the sustainability of agricultural practices over time [7]. Northeast Germany is one of the driest regions in central Europe, with loamy and sandy soils being the two dominant soil types [8,9]. Cereals are one of the main crops cultivated in this region. Previous cereal crop-related long-term experiments in the region have focused on crop yields [10,11], tillage [12], and the soil organic carbon (SOC) [13]. Moreover, irrespective of the underlying drought stress and poor water holding conditions of the prevailing soils in large areas of the northeast, concerns over climate change have driven research in these environments [14,15].

In analyzing data from designed experiments, classical parametric methods such as the analysis of variance (ANOVA), parametric correlation, and regression have long been commonly used to assess crop yield [16]. However, these classical methods have limitations. For example, while ANOVA is best suited to identifying yield differences between treatments in designed experiments, it does not exhaustively account for the extraneous factors that influence yields [17,18]. Similarly, parametric correlations and linear regressions are less suited to handle missing, unbalanced, and higher-order data and nonlinear interactions [18,19]. Flexible and robust methods are now available for dealing with multivariate, unbalanced data that account for nonlinear, higher-order interactions. For instance, statistical models such as the linear mixed-effects model (LMM), generalized linear models [20,21], and machine learning (ML) models such as random forest, artificial neural networks, and decision tree algorithms [22–24] can be applied to handle these challenges. Unlike ANOVA models, LMMs cover nontreatment variables and random factors that tend to mask the treatment effects, thereby improving the reliability and interpretation of experimental results [18]. While Piepho [25] stated that the most common variables affecting the yield could be determined in the LMM framework when environmental effects and treatment effects are considered random and fixed factors, other reports advised that the LMMs and linear regression models have limitations because the analytical interpretation and pattern prediction can be confounded due to significant high autocorrelation or missing data [26–28]. ML models such as the classification and regression tree (CART) and M5P algorithm-based decision tree have been employed in agricultural research [29,30]. These regression tree models are most useful in handling complex databases with a high number of attributes and high dimensions collected from observational experiments [31] but can also profitably be applied for small datasets from designed experiments [32]. They are robust tools for dealing with missing data. Additionally, regression tree models can capture important nonlinear relationships and interactions between variables [31]. However, this method, by contrast, has not yet been widely used in analyzing data collected from LTFEs.

Statistical inference and prediction are two major goals in the study of agricultural experiments. While statistical models are designed to draw inferences of relationships between variables within assumptions, ML is a modeling tool for finding generalizable predictive patterns without hypotheses [33]. Limitations in the use of statistical inferences and ML are still subject to debate. We have not yet found clarity in the literature regarding the comparison of how different statistical analyses and the ML model explain the results of LTFE data. Moreover, it is important to know the best suited tools and methods for unraveling the important interconnected multiple variables that influence crop yields in the long term. Therefore, in this study, we tested the use of ANOVA in the general linear model and two nonparametric methods, the LMM and M5P models, to understand grain yield variations of WW in an LTFE (“V140”) in Müncheberg, Germany. The objectives of the study were to (1) identify the important variables that explain the grain yield variations of WW in the LTFE and (2) compare different analytical methods for explaining the WW yield variation.

### 4.2. Materials and Methods

#### 4.2.1. Experimental Site

The LTFE “V140” was established at the experimental station of the Leibniz Centre for Agricultural Landscape Research (ZALF), Müncheberg, Germany, in 1963 [34]. The site is located in the Märkisch-Oderland district, around 50 km east of Berlin. The area is characterized by a dry period, particularly during the early summer [35]. The mean annual precipitation in the area was 551 mm ± 121.6 standard deviation (s.d.), and the mean annual temperature was 8.7°C ± 0.9 s.d. during the cultivation period of WW (1973–2010). The soil in the area is classified as a Podzoluvisol to Arenosol. According to the German Guidelines for Soil Assessment (Bodenschätzung), the dominant soil texture classes are slightly loamy sand and sand (Sl4D and S4D) [34]. The site has recently been described in more detail in Thai et al. [36].

#### 4.2.2. Experimental Design and Management

The experiment was set up on a flat plain measuring 5712 $m^2$ involving 168 individual plots. The individual plots measured 6.0 m × 5.0 m, and a buffer zone of 1 m was allowed between the blocks. The experiment was arranged in a randomized complete block design (RCBD) comprising 21 treatments with eight blocks. The treatments included five levels of mineral N fertilizer (NF), each in combination with four organic fertilizers (ORF) (Table S 4.1, supplement). The five NF levels comprised 35, 70, 105, 140, and 175 N kg $ha^{-1}$, which are hereafter referred to as N1, N2, N3, N4, and N5, respectively. The ORF treatment included 0, 1.2, and 3.2 t dry mass (DM) $ha^{-1}$ farmyard manure (FYM) and 2.0 t DM $ha^{-1}$ straw, which is henceforth referred to as sole mineral fertilizer (nitrogen, phosphorus, and potassium combination), fym1, fym2, and straw applications, respectively (Table S 4.1). However, at a 3.2 t DM $ha^{-1}$ application rate of FYM, the NF levels applied were 0, 35, 70, 105, and 140 N kg $ha^{-1}$ (fym2 application). The control treatment received no fertilizer inputs. Due to the different NF levels in the fym2 application compared to other applications, group treatments were made to balance the NF rates among the different applications to compare the effects of different ORFs on the yield. The group treatments included control, NPK, NPK+fym1, NPK+fym2, PK+fym2, and NPK+straw (Table S 4.1, supplement). From 1980 onwards, phosphorus and potassium fertilizers were applied at 30 kg $P_2O_5$ $ha^{-1}$ and 100 kg $K_2O$ $ha^{-1}$, respectively, to all plots except the control treatment. NF was applied twice each year during the growth of WW, i.e., the basal amount was applied in the middle of April, and the remainder was applied a month later between shooting to full blooming (end of May or early June). The FYM was applied every two years from 1973 to 1994 in autumn before planting maize, potato, or sugar beets, depending on the cropping system in the year. After 1994, the FYM was applied every four years in autumn before planting maize or potatoes. The FYM used in each year contained, on average, 2.1% N, 0.7% $P_2O_5$, 2.1% $K_2O$, 0.4% Mg, and 55.4% organic matter. Straw from the preceding cereal crop was applied at 2 t $ha^{-1}$ every two years throughout the experimental period. The dry mass straw contained, on average, 0.7% N, 0.1% $P_2O_5$, 1.9% $K_2O$, and 0.1% Mg. Lime was uniformly applied to all the plots in all trial years.

Sowing of WW was performed at the end of September or in early to middle October in most years during the study period. The sowing densities were the same in all experimental years, but the WW varieties were changed over time. Harvesting was performed at the end of July or the beginning of August in most experimental years, depending on the weather conditions. The WW was harvested at the physiological maturity stage using a harvester. Weeds were controlled with a postemergence herbicide.

The crop sequence was not fixed and consisted of WW, winter rye, spring barley, potatoes, sugar beets, maize, flax, and peas (Table S 4.2, supplement). One of these crops was cultivated in the experimental site during each growing season. The crop preceding WW was potatoes in 1973, 1983, 1987, and 1991 and sugar beets in 1993. In 2001 and 2009, the crop preceding WW was peas. There were seven WW crop rotations, and the crop rotations were different in this experiment. The management practices of WW, such as plowing, harrowing, and fertilization, were the same in each experiment year.

### 4.2.3. Data Description

Crop yield: Long-term data from 1973 to 2010 were used for the analyses. The DM grain yield data of WW (Mg DM ha$^{-1)}$ were obtained from the seven years of WW cultivation. The DM yield of the preceding crop was obtained in each trial year to estimate its effects on the yield of WW.

Meteorological data: The weather data used in the analysis were obtained from an adjacent climate station of the German Meteorological Service [37]. The daily mean air temperature, maximum temperature, minimum temperature, and precipitation during the growing period of WW were used to calculate the input weather variables for this study. The monthly mean temperature, cumulative precipitation, cumulative number of days recorded with mean temperatures above 30°C in every month (days Tmax > 30°C in a month), the cumulative number of days recorded with mean temperatures below 0°C or 32°F (freezing days in a month), and cumulative growing-degree days during the growing seasons were calculated. The maximum and minimum temperatures were used to calculate the growing degree days (GDD).

Soil variables: Soil chemical analyses were performed in the 1984, 1988, 1992, and 1994 trial years. The results of the soil analysis are presented in Table S 4.3 of the supplementary material. Selected soil variables such as the total N and SOC in each treatment were used as input data to estimate their effects on the yield of WW. All input variables considered in this study are presented in Table S 4.4 of the supplementary material.

### 4.2.4. Data Analysis

There are two main steps in the analysis: (1) exploring the WW grain yield and yield variability using descriptive analysis and ANOVA within fixed effects models-general linear model and (2) applying nonparametric methods involving the LMM (statistical model) and M5P (ML model) models for the grain yield response.

#### *4.2.4.1. ANOVA Test*

The effects of fertilization treatments on the grain yields were analyzed by one-way ANOVA using SPSS version 25, and the significance was determined by Tukey's post hoc test. A fixed-effects model, the general linear model, was used to evaluate the main and interaction effects among treatments (fertilizer) and years (annual or environmental effects) on the grain yields over the years. The effect sizes of the fertilizer, environmental factors, and their interactions were estimated based on the sum of squares-type III in the general linear model. The environmental factors considered in this study were weather, soil chemical properties, and preceding crops and their yields. Furthermore, soil data were analyzed by ANOVA to understand the changes in soil properties under long-term fertilization practices and then to select the important variables for developing models. To avoid the effects of collinearity in the statistical model and overfitting in the ML model, Pearson's correlation analysis was checked between the target variables (WW grain yield) and predictor variables and between predictor variables together. Based on Pearson's correlation coefficient, useful variables were maintained, while redundant variables were removed before developing the yield models. Statistical significance for the analyses was set at $p < 0.05$.

#### *4.2.4.2. Linear Mixed-Effects Models*

LMMs are an extension of the linear regression model and include both fixed and random effects as predictor variables via a restricted maximum-likelihood estimate (REML). The LMMs were fitted using the "lmer" function implemented in the "lme4" package [20] of the R statistical language, version 3.6.3 [38], to assess the WW yield as a function of the different factors. The LMM for the yield response is specified by Equation (1).

$$y = X\beta + Zu + \varepsilon \quad (1)$$

Here, y is the vector of the wheat yield (outcome/target variable: Mg $ha^{-1}$); X and β are the design matrix and the vectors of fixed effects, respectively; Z and u are the design matrix and the vectors of random effects; and ε represents the vector experimental error. In this study, the LTFE with RCBD was specified with experimental years and experimental blocks and plots as random effects on the yield. The fixed-effect variables were the levels of NF, type, and levels of ORF, selected weather parameters, preceding crops and their yields, and selected soil chemical properties. Regarding the model development process, after checking for normality on model residuals using quantile-quantile (Q–Q) plots, we first fitted the random effect model. Then, we added more predictors as fixed effects to the random effect model. From these LMMs, we performed gradual backward elimination of nonsignificant LMM effects, beginning with the random effects followed by the fixed effects. In this study, the LMM was fitted as a random intercept model at a 97.5% confidence interval (CI). Models were selected using the Akaike Information Criterion (AIC).

From the final LMM, we calculated the relative important variables using the Relaimpo package in R, version 3.6.3 [39]. This is a supplemental test to regression analysis to calculate the proportional contribution of each predictor variable to explaining variance in the LMM. The statistical tests were considered significant at the 0.05 probability level.

#### *4.2.4.3. Machine Learning Model*

We used the M5P algorithm, which is a recursive partitioning algorithm based on thresholds for developing a decision tree structure, to uncover the relationship and interaction between the WW yield and predictor variables. The M5P is a powerful implementation of Quinlan's M5 algorithm [40, 41] and an advantaged algorithm among decision tree algorithms for training an ML model. We implemented M5P in WEKA (Waikato Environment for Knowledge Analysis) software version 3.8.4. The rule of M5P is to recursively partition the data space and fit a prediction model within each partition. The results of the implementation are a binary regression tree model and are represented as an inverted tree, wherein the terminal nodes are the linear regression functions. The tree includes a root node (top node), internal nodes, and terminal nodes connected by edges. Additionally, branches or subtrees are split from the root node, and internal nodes correspond to the outcome of the test. The terminal nodes are the prediction values of the WW yield. When the values of the outcome at the terminal nodes are numeric, the terminal nodes of the tree can be constant values, and the tree is called a regression tree. In contrast, the tree is called a model tree once the terminal nodes of the tree are piecewise linear regression equations [42]. Before training the M5P model, we checked the correlation ranking between the selected input variables and yield by the attribute selection function in WEKA. Next, we used the split function in WEKA to randomly partition the preprocessed data into two subsets, including the training set (80%) and test set (20%). The training set was used to build the decision tree model (determine its parameters), the ten-fold cross-validation method was used to estimate the accuracy of the supervised learning algorithm, and the test set was used to evaluate the predictive performance of the trained model [43]. The coefficients of determination ($R^2$) and root mean square error (RMSE) were used to assess the performance of the models. After obtaining a final M5P model, we used bootstrap 1000-tree analysis by the Relaimpo package in R version 3.6.3 [39], to identify the relative importance of predictor variables. In this study, we present the results of the regression tree model as a piecewise constant function.

#### *4.2.4.4. Evaluation Metrics*

A good fitting model is generally one in which the results of the predicted values are close to the actual values for the selected model. Thus, the predicted grain yields of WW produced by LMM and M5P-based regression trees were compared to the actual yields observed in the LTFE. We employed standard statistical criteria such as $R^2$, RMSE, and mean absolute error (MAE) values to assess the predictive performance for WW grain yields by the selected models. The $R^2$ value indicates the fitness of the model for predicting the WW yield, while the RMSE and MAE are commonly used to measure the difference between the predicted and actual values. Furthermore, the RMSE can be used to evaluate the closeness of these predictions to the actual values, while the MAE can better represent the predictor error. Higher $R^2$ values and lower values of the RMSE and MAE indicate better estimation accuracies of the models [44]. Equations for the evaluation metrics are given in the supplementary material, EQ1 (Es1, Es2, Es3).

### 4.3. Results

#### 4.3.1. Grain Yield of Winter Wheat

The ANOVA results show that there was a significant effect of fertilization on the grain yield of WW (Figure 4.1, $p < 0.001$). Irrespective of ORF application, the mean grain yield increased significantly with increasing NF application rates until N3, except in fym1 and fym2. At zero NF, treatment 3.1 showed nearly 0.5-times higher increases in grain yield relative to the no input control. At N1, the highest significant mean grain yield was observed when coapplied with fym2, while no differences were observed among mineral fertilizer, fym1, and straw applications. There were no significant differences in mean grain yields among the four application regimes (mineral fertilizer, fym1, fym2, and straw application) at N2, N3, N4, and N5. Optimal mean grain yields of less than 5.0 Mg DM ha$^{-1}$ were obtained at N3 in the mineral fertilizer and straw application and at N2 in fym1 and fym2 applications. Optimal yields of 4.60 Mg DM ha$^{-1}$ and 4.16 Mg DM ha$^{-1}$ were observed at N3 and N2 in mineral fertilizer and fym1 applications, respectively. In fym2 and straw applications, optimal yields of 4.44 Mg DM ha$^{-1}$ and 4.69 Mg DM ha$^{-1}$ were obtained at N2 and N3, respectively. Therefore, FYM applications with both levels (fym1 and fym2) showed better effects on the grain yields compared to mineral fertilizer application or straw application.

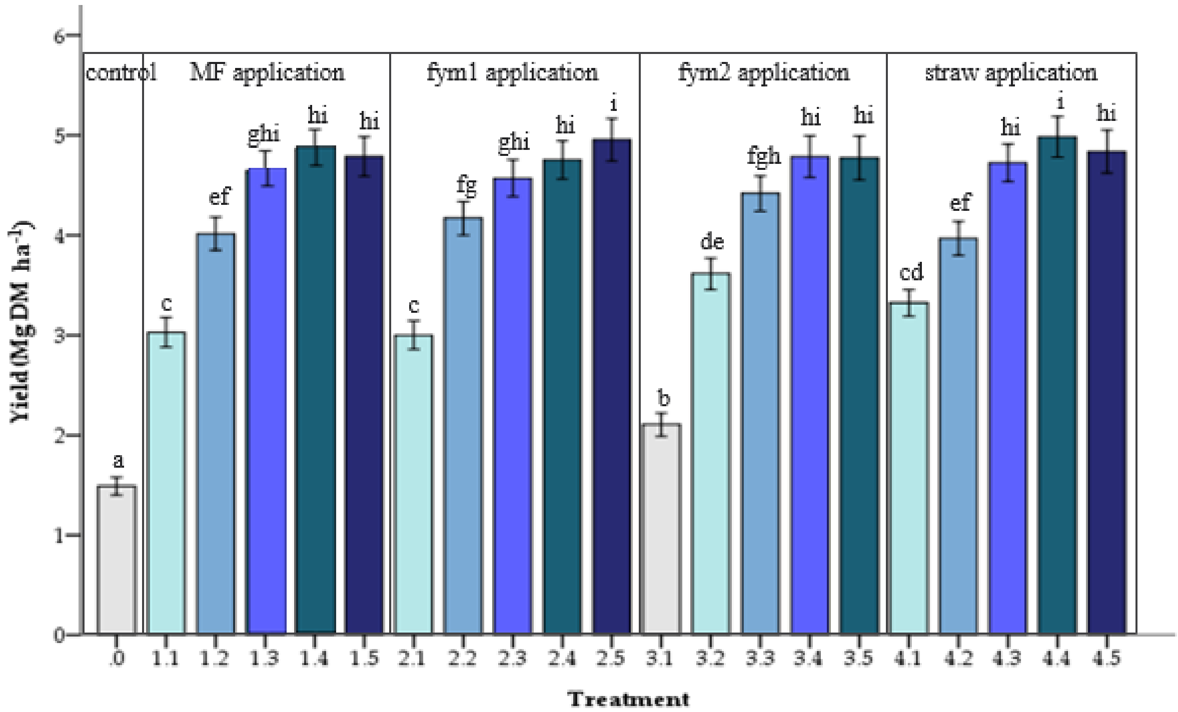


**Figure 4.1.** Mean grain yields (Mg DM $ha^{-1}$) of winter wheat (WW) under different fertilizer treatments and fertilization practices. Error bars indicate the standard errors (SE) of the means. Treatments sharing the same letter are not significantly different ($p < 0.05$). Treatment codes are given in Table S 4.1. MF: mineral fertilizer; fym: farmyard manure. Bars with the same color show the same rate of NF.

The mean WW grain yields of the group treatments are shown in Table 4.1. The average yields ranged from 1.48 Mg DM $ha^{-1}$ $year^{-1}$ in the control to 4.42 Mg DM $ha^{-1}$ $year^{-1}$ in the NPK + fym2 treatment. The average grain yields were not significantly different among the NPK, NPK + fym1, NPK + fym2, and NPK + straw treatments. The average grain yield in NPK+fym2 was twice as high as that in PK + fym2 and three times higher than that in the control. The coefficient of variation (CV) of grain yields for each group treatment ranged from 0.26 in NPK to 0.39 in the PK + fym2 treatment. Non-NF input treatments (control and fym2 + PK treatments) showed relatively higher CV values of 0.33 and 0.39, respectively, compared to NF-applied treatments. The NF-applied plots showed similar CV values of 0.26 and 0.27. In general, the results of ANOVA and descriptive statistics in Table 4.1 show that NF treatments can maintain a stable and higher WW grain yield compared to treatments without NF.

**Table 4.1.** Yield and yield variation of winter wheat for group-treatments in the long-term field experiment.

| Group Treatment | Yield (Mg DM $ha^{-1}$) | ±Se | CV | Percent Change in Yield Relative to Control (%) | Percent Change in Yield Relative to PK + fym2 (%) |
|---|---|---|---|---|---|
| Control | $1.48^{a}$ | 0.19 | 0.33 | - | −34 |
| NPK | $4.10^{c}$ | 0.41 | 0.26 | 179 | 85 |
| PK + fym2 | $2.23^{b}$ | 0.32 | 0.39 | 51 | - |
| NPK + fym1 | $4.11^{cd}$ | 0.40 | 0.26 | 179 | 85 |
| NPK + fym2 | $4.42^{d}$ | 0.45 | 0.27 | 200 | 99 |
| NPK + straw | $4.23^{cd}$ | 0.42 | 0.26 | 187 | 90 |

Group treatments are given in Table S 4.1; Mg DM: megagram dry mass; Se: standard error; CV: coefficient of variation. Different letters in the second column indicate a significant difference in the WW grain yield at $p < 0.05$.

The grain yield dynamics of WW in the tested years are shown in Figure 4.2. Irrespective of the group treatment application, there were significant differences in WW grain yields among the trial years. The WW grain yield variability was high among the years and ranged from 2.35 Mg DM ha−1 in 1991/92 to 5.39 Mg DM ha−1 in 1983/84. Except for the PK + fym2 treatment, there were significant increases in grain yield for all group treatments compared to the control in almost all years.

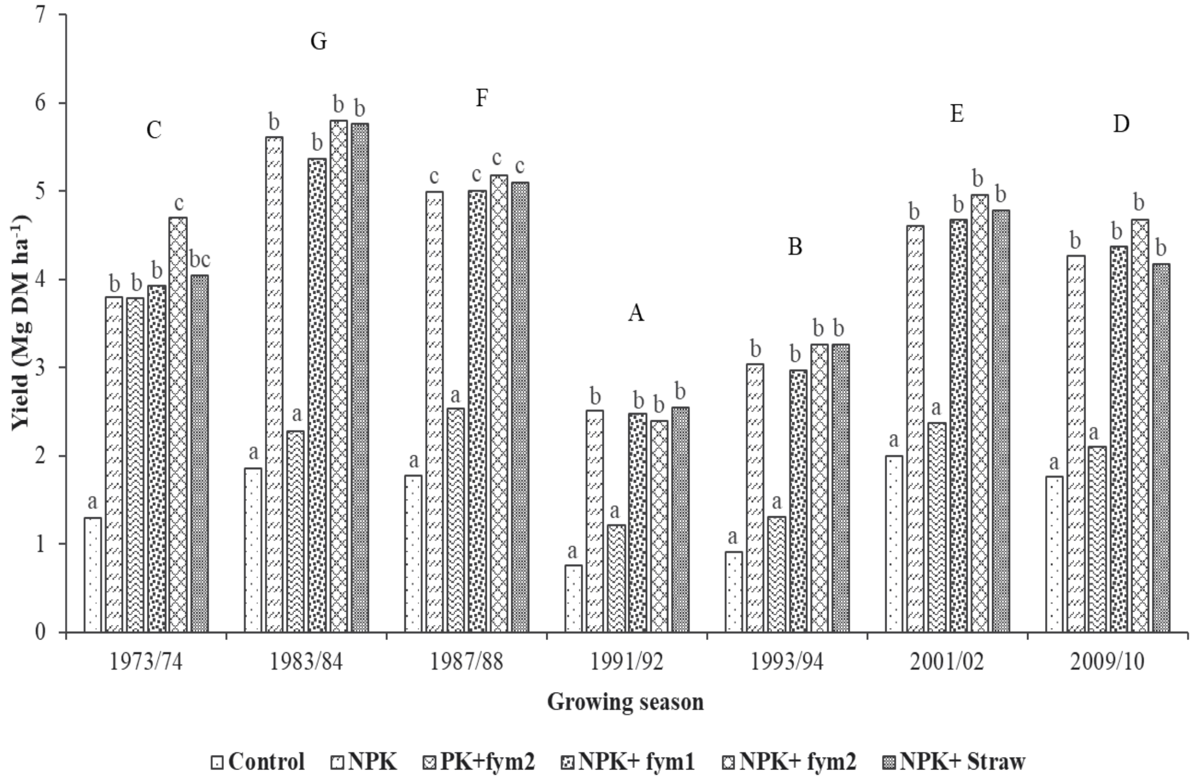



**Figure 4.2.** Grain yields of WW (Mg DM $ha^{-1}$) in all trial years under different group treatments. Means sharing the same letters are not significantly different ($p < 0.05$). Capital letters at the top of bars indicate a comparison of average grain yields among the trial years. Small letters denote a comparison of the yields among group treatments within a given year.

**Table 4.2.** Results of ANOVA and Eta squared between fertilizers and years (environment) for winter wheat yields.

| Source | Type III Sum of Squares | df | Mean Square | F | Sig. | Eta Squared ($h^2$) |
|---|---|---|---|---|---|---|
| Corrected Model | 2315.99 | 146 | 15.86 | 32.40 | *** | - |
| Intercept | 17142.05 | 1 | 17142.05 | 35009.21 | *** | - |
| Treatment | 920.13 | 20 | 46.01 | 93.96 | *** | 34 |
| Year | 1148.10 | 6 | 191.35 | 390.79 | *** | 42 |
| Treatment x Year | 154.63 | 120 | 1.29 | 2.63 | *** | 6 |
| Error | 462.71 | 945 | 0.49 | - | - | 17 |
| Total | 21013.49 | 1092 | - | - | - | - |
| Corrected Total | 2778.71 | 1091 | - | - | - | - |

Sig.: significant; *** = $p < 0.001$; $R^2 = 0.83$ (adjusted R squared = 0.80); dependent variable: yield of WW (Mg $ha^{-1}$); Eta squared ($h^2$): proportion of total variance that is attributed to an effect.

The results of ANOVA or general linear model analysis show that the grain yield of WW was significantly affected by the environment/year (42%), followed by the fertilization treatment (34%) and environment × fertilization (6%), with 17% of the variation attributed to error (other factors) (Table 4.2). These results explain 80% of the variance with an adjusted R squared value of 0.80 at $p < 0.001$.

### 4.3.2. Modeling and Predictors

#### *4.3.2.1. Linear Mixed-Effects Model*

The results of the LMM reveal that NF application, freezing days in December and in February, precipitation in June, the yield of the preceding crop, the temperature in October, the cumulative number of days in July with maximum temperatures above 30°C (days Tmax > 30°C in July) and the total N in the soil were fixed factors that influenced the grain yield of WW (Table 4.3).

**Table 4.3.** Estimate of the coefficients (β) and P-values in the linear mixed-effects model.

| Model | M0: intercept only | | | M: with predictors | | |
|---|---|---|---|---|---|---|
| | **Estimate (β, Mg ha$^{-1}$)** | **s.e.** | **P-values** | **Estimate (β, Mg ha$^{-1}$)** | **s.e.** | **P-values** |
| Fixed effects | | | | | | |
| Intercept | 4.081 | 0.443 | *** | −2.426 | 0.231 | *** |
| N fertilizer rate | - | - | - | 0.012 | 0.001 | *** |
| Freezing days in December | - | - | - | 0.144 | 0.011 | *** |
| Precipitation in June | - | - | - | 0.005 | 0.001 | *** |
| Freezing days in February | - | - | - | 0.134 | 0.007 | *** |
| Preceding crop yield | - | - | - | 0.157 | 0.015 | *** |
| Days Tmax > 30°C in July | - | - | - | −0.139 | 0.016 | *** |
| Temperature in October | - | - | - | 0.215 | 0.017 | *** |
| Total N in soil | - | - | - | 0.001 | 1E-04 | *** |
| $R_m^2$ | 0 | - | - | 0.73 | - | - |
| Random effects | Variance | SD | | Variance | SD | |
| Plot | 0.82 | 0.90 | *** | 0.09 | 0.31 | *** |
| Block | 0.09 | 0.31 | *** | 0.06 | 0.26 | *** |
| Year | 1.01 | 1.01 | *** | - | - | ns |
| Residual | 0.51 | 0.71 | - | 0.46 | 0.68 | - |
| Deviance | 2702.30 | - | - | 2516.9 | - | - |
| $R_c^2$ (Total) | 0.79 | - | - | 0.8 | - | - |

*** = $p < 0.001$; ns = $p > 0.05$; $R^2$m: marginal coefficient of determination for fixed factors alone; $R^2$c: conditional coefficient of determination for both fixed and random factors; SD:

standard deviation; freezing days in December/February: cumulative number of days in December/February with mean temperatures below 0°C (32°F); days Tmax > 30°C in July: cumulative number of days in July with maximum temperatures above 30°C.

The model indicated blocks and plots as random factors. The fixed effects explained 73% ($R^2m = 0.73$) of the variance in the grain yield, while the total of both the fixed and random effects explained 80% of the variance ($R^2c = 0.80$) at a 97.5% CI. In particular, NF application and freezing days in December showed the highest significant contribution to the grain yield, i.e., 21.7% and 17.3%, respectively (Table 4.4). However, the temperature in October (3.9%) and total N in the soil (3.3%), although significant, were less important predictors of the grain yield. The plots and blocks explained 15.2% and 10.5% of the variance in the grain yield, respectively.

#### *4.3.2.2. Machine Learning Model*

The M5P regression tree model generated five splits and 17 terminal nodes and explained 80% of the variability in the data (Figure 4.3). The hierarchy of the regression tree model, as well as the results from bootstrapping 1000 trees, indicated freezing days in December and the NF rate as the main determinants of the WW grain yield (Figure 4.3, Table 4.4). Freezing days in December and the NF rate accounted for 31.7% and 22.5%, respectively, of the contribution to the grain yield of WW. Other variables, such as the yield of the preceding crop, the temperature in October, freezing days in February, the total N in the soil, the SOC, and the FYM, were also determinants of the WW grain yield. The total N in the soil, SOC, and FYM showed a minimal influence on the grain yield of WW, with relative contributions of 3.0%, 2.3%, and 0.4%, respectively. The effects of the total N in soil and SOC on the grain yield were only evident in plots that received NF application. In contrast, the FYM slightly influenced the grain yield only in plots that received no NF application (Figure 4.3).

#### *4.3.2.3. Comparing Models and Model Fit*

The results of ANOVA indicate fertilizer application and the environment as the main factors that explained the grain yield, with an adjusted R squared of 0.80 (Table 4.2). Among the treatment inputs, NF application was the main variable that influenced the grain yield, while FYM slightly influenced the grain yield with no NF application. The effects of fertilizer and the environment on the grain yield were revealed in detail in the results of the LMM and M5P regression tree models. Both the LMM and M5P model identified the NF and freezing days in December as the most crucial variable that influenced the grain yield of WW (Table 4.4). However, the relative proportions of both variables differed hierarchically in both models. The NF rate was the most important variable, while freezing days in December was the second most important variable that explained WW yields in the LMM. Conversely, the results from the M5P model show the NF rate as the second most important variable and freezing days in December as the first most important predictor of WW yields.

Additionally, freezing days in February, the yield of the preceding crop, the temperature in October, and the total N in the soil were determinants of the grain yield in both models. Apart from the common determinants in both models, the LMM highlighted precipitation in June and days Tmax > 30°C in July, while the M5P model revealed SOC and FYM in plots that received zero NF input as variables that influenced the WW yield.

The results of the fitting model for the LMM ($R^2$ = 0.8, $p < 0.001$) and the training model for the M5P regression tree ($R^2$ = 0.8) (Table 4.4) show a generally good fit between the predicted yield and actual yield of WW. The evidence of good fit between the predicted yield (modeled) and actual yield (observed) for the LMM and M5P model are shown in Figure S 4.1 and Figure S 4.2 in the supplementary material. The LMM showed a better performance in predicting the grain yield compared to the M5P regression tree model, as reflected by its relatively smaller RMSE and MAE values. The LMM showed RMSE and MAE values of 0.68 and 0.54, respectively, while 0.74 and 0.58 were computed for the M5P regression tree model.

**Table 4.4.** Important variables indicated by the linear mixed-effect and M5P regression tree models as predictors of winter wheat yields in the LTFE.

| **LMM** | | **Relative Contributions with Confidence Intervals (%)** | | | **M5P Regression Tree** | | **Relative Contributions with Confidence Intervals (%)** | | |
|---|---|---|---|---|---|---|---|---|---|
| No. | Predictors | Relative important variables | Lower | Upper | No. | Predictors | Relative important variables | Lower | Upper |
| | Fixed effects | | | | | | | | |
| 1 | Nitrogen fertilizer rate | 21.7[a] | 19.2 | 24.3 | 1 | Freezing days in December | 31.7[a] | 29.2 | 34.2 |
| 2 | Freezing days in December | 17.3[b] | 15.7 | 19.0 | 2 | Nitrogen fertilizer rate | 22.5[b] | 19.7 | 25.6 |
| 3 | Precipitation in June | 8.2[cd] | 6.9 | 9.6 | 3 | Preceding crop yield | 7.9[c] | 6.5 | 9.2 |
| 4 | Freezing days in February | 7.6[cde] | 6.3 | 9.2 | 4 | Temperature in October | 5[de] | 3.9 | 6.3 |
| 5 | Preceding crop yield | 6.6[def] | 5.6 | 7.7 | 5 | Freezing days in February | 4.6[de] | 3.6 | 5.8 |
| 6 | Days Tmax > 30°C in July | 6.0[ef] | 5.3 | 6.8 | 6 | Total nitrogen in the soil | 3.0[f] | 2.4 | 3.8 |
| 7 | Temperature in October | 3.9[gh] | 3.0 | 4.9 | 7 | SOC | 2.3[g] | 2 | 2.8 |
| 8 | Total nitrogen in the soil | 3.3[gh] | 2.5 | 4.1 | 8 | FYM | 0.4[h] | 0.3 | 0.6 |
| | Random effects | | | | | | | | |
| 1 | Plot | 15.2 | - | - | - | | - | - | - |
| 2 | Block | 10.5 | - | - | - | | - | - | - |
| Statistical indicators | | | | | | | | | |
| $R^2$ | 0.8 | - | - | - | - | 0.8 | - | - | - |
| RMSE | 0.68 | - | - | - | - | 0.74 | - | - | - |
| MAE | 0.54 | - | - | - | - | 0.58 | - | - | - |

FYM: farmyard manure, SOC: soil organic carbon, $R^2$: coefficients of determination, RMSE: root mean square error, MAE: mean absolute error. Different letters in the same column indicate that the difference in predictor ranking is significant at 97.5%.

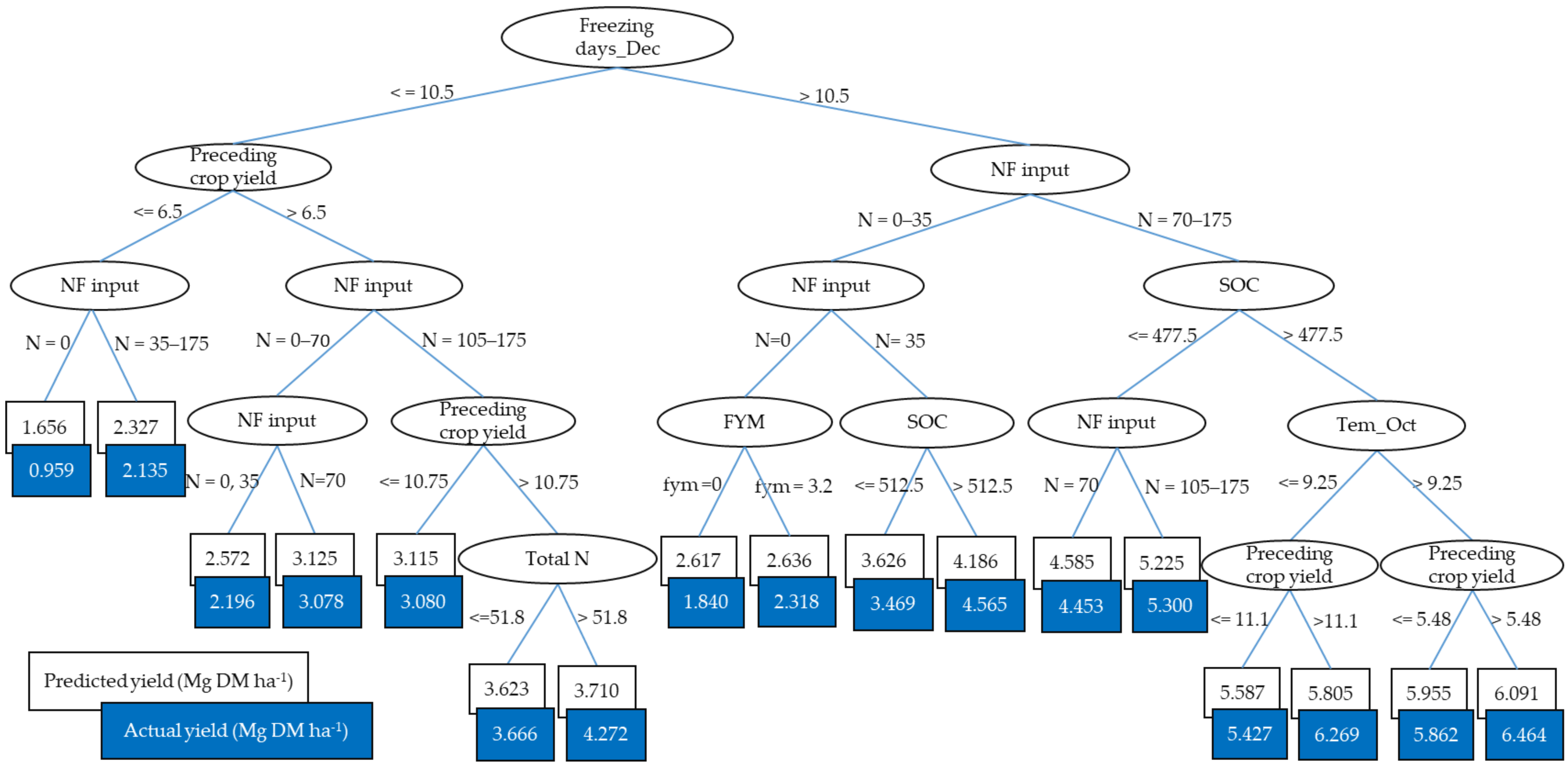



**Figure 4.3.** M5P regression tree model describing the grain yield of winter wheat (Mg DM ha$^{-1}$) in the LTFE as a function of the fertilizer, weather, soil, and preceding crop yield. The predicted yield and actual yield values are given in megagram dry mass ha$^{-1}$. Freezing days_Dec: cumulative freezing days in December; Tem_Oct: temperature in October; SOC: soil organic carbon; mineral nitrogen fertilizer: NF; farmyard manure; FYM.

### 4.4. Discussion

#### 4.4.1. Grain Yield of Winter Wheat and Treatment Effects

The optimum WW grain yield of less than 5.0 Mg DM ha$^{-1}$ observed in the current study was much lower than the national average yield of 7.7 Mg DM ha$^{-1}$ from 2014 to 2018 [2]. The yields of WW in this study markedly increased when NF or its combination with ORF was applied (Figure 4.1). Similar to the observations for spring barley in the experiment [36], we found that NF input was a major determinant of the grain yield of WW. Fixen and West [45] stated that plant-available N is one of the most important nutrients for increased yields of major food crops. In this study, the average grain yields of WW increased to optimal yields along with increasing NF application to a certain threshold. The optimal yields were obtained at N3 in the mineral fertilizer and straw applications and at N2 in the fym1 and fym2 applications (Figure 4.1). This reveals that the effects of NF application on the yield of WW were different under different ORF applications. The effect of the ORF application compared to mineral fertilizer application alone on the grain yield was evident at N1 only in the fym2 application, reiterating the importance of FYM amendment and its dose in attaining the optimum yields of crops. FYM amendment can reduce the need for the higher NF application rate demanded by wheat. This observation is similar to those in studies by Blanchet et al. [46] in Switzerland.

The combined application of FYM with mineral fertilizer has been reported to improve the grain yield of WW in Germany [11, 47]. The yield increase is attributed directly to the effects of additional N and indirectly to the improved soil conditions related to organic material applications [48]. In this study, with the same NF input in NPK, NPK + fym1, NPK + fym2, and NPK + straw, the average yields of these group treatments were similar, irrespective of the ORF application type and amounts (Table 4.1), implying the minimal influence of FYM or straw on the grain yield. Additionally, the FYM in plots that received zero NF input appeared in the M5P regression model, but the corresponding effect on the WW yield variation was very small. This is ascribed to the fact that organic inputs are usually low in nutrients and unable to satisfy the nutrient demands of cultivated crops [49].

Although the effects of FYM application on the grain yield variation were very small and the effect of straw on the yield was not clear in the present study, the combined application of ORF along with NF increased the grain yield stability of WW (Table 4.1). These findings are consistent with observations from WW grain yield stability studies in Giessen, Germany [5,6]. We observed higher grain yield stabilities of WW in all treatments with NF input, as shown by their lower CV values compared to the control or PK+ fym2 treatment. This observation is ascribed to the plant-available N input from NF, which aided in the vigorous growth of wheat plants and the development of greater resilience against environmental stress. NF was the main fertilizer factor that showed enhanced effects on wheat yields through improvements in plant growth and root development [50], and thus aided in the water and nutrient uptake capacity. Thus, NF application in the WW cultivation system could not only enhance grain yields but also reduce the yield variability year to year.

### 4.4.2. Environmental Effect on the Winter Wheat Yield

The temperature in October was an influencing variable for the WW grain yield in both models (Table 4.4). In this experiment, WW was sown at the end of September to mid-October. In general, the optimal temperature for wheat germination is 12 to 25 °C [51], while the average temperature in October throughout the trial years was 8.7 ± 1.8 °C (Table S 4.5a). Therefore, a warmer temperature in October is favorable for the germination, emergence, and initial growth of leaves, crowns, and secondary root systems of WW plants [52].

Many previous studies have reported on the effects of winter freezing temperatures on the grain yield of WW [53–55]. Our findings reveal that freezing days in December appeared to be the most crucial and consistent climate-based driver for the grain yield of WW in both the LMM and M5P model (Table 4.4). The freezing days in February were also a consistent determinant of grain yields in both models, although they showed only a 7.6% contribution in the LMM and a 4.6% contribution in the M5P model. Seedlings of wheat normally require a minimum of four to five leaves and at least one to two tillers to have enough energy reserves to survive the winter [56]. Thus, winter hardiness or cold tolerance is an extremely crucial physiological process that affects wheat survival during winter and its subsequent growth and development. According to Lollato et al. [56], wheat plants remain cold-hardy as long as the crown temperatures remain below 0 °C. The wheat crown is the most crucial organ for WW survival during winter [57], since viable crown tissue enables the regeneration of other plant organs damaged by freezing injuries. Hence, the survival of WW depends on the viability of the crown. WW will normally have reached its maximum level of cold hardiness by the time winter begins in December [56]. In this regard, more freezing days in December, which implies more exposure to freezing conditions, will support the cold hardiness process and WW survival and grain formation. On the other hand, wheat plants will experience a gradual loss of cold hardiness when the soil temperature around the crown rises above 10 °C. Once WW plants lose their maximum level of cold hardiness, there is the possibility to reharden during the winter, but they will not regain their maximum level of cold hardiness. Thus, having more freezing days in February is important for the subsequent growth and increased grain yield of WW plants. Furthermore, the climate in most parts of Germany is moderately continental and is characterized by an average daily temperature of 0 °C in winter [58]. During the trial years in this study, the average temperatures in December and February were 0.6 and 1.4 °C, respectively (Table S 4.5a), and more days of freezing temperatures potentially not only favored the survival of WW plants but also reduced plant disease inoculum and incidence during winter.

Generally, drought and high-temperature stress often occur simultaneously at anthesis and during the grain-filling period and/or at physiological maturity in wheat, causing significant yield losses [59,60]. An increased frequency of droughts, especially in early summer in Germany, has been suggested to affect wheat production, particularly in Northeast Germany, which is characterized by predominant sandy soils [9,61]. In this study, the LMM showed that more precipitation in June positively influenced the WW grain yield, whereas more days of Tmax > 30°C in July negatively influenced the grain yield (Table 4.3). This observation has previously been reported for both WW and spring wheat [62–64]. In addition, the second application of NF between shooting and full blooming (the end of May or early June), together with adequate precipitation in June and fewer days of Tmax > 30°C in July, was critical to grain yield development (Table S 4.5b and S 4.5c). This finding is consistent with the observations of Altenbach et al. [59] that fertilizer application at anthesis, drought, and high temperatures affects the grain development, kernel composition, and grain yield. When plants are grown without additional fertilizer at anthesis, coupled with exposure to drought and high-temperature stress, the duration of grain filling shortens, resulting in low kernel weights and low yields. Moreover, senescenced leaves appear much earlier under high temperatures and coincide with physiological maturity, which shortens the time to maximum growth, dry weight, and duration of starch accumulation [65].

Previous studies have reported the effects of the preceding crop type and preceding crop yield as important factors that influence WW yields in LTFEs [66–68]. Our results show that only the yield of the preceding crop was an important variable that explained the variance in the WW yield (Table 4.4). Nonetheless, once the yield of the preceding crop was included in the models explaining WW yield variation, the preceding crop type could be related. The type of preceding crop did not appear in the models, which was likely a result of the small replication of each preceding crop or small sample representation of the preceding crop type in this experiment. In the unfertilized control, the grain yield of WW was 1.3 and 1.9 t $ha^{-1}$ after root crops and peas, respectively. Therefore, peas could be considered a favorable preceding crop for WW in this experiment. More long-term trials with peas are required to verify this observation.

The total N in soil revealed minor influences on the grain yield of WW and yield variability in both models (explaining around 3%). This consistency in both models relates to the N input, which is an important nutrient for increased crop yield. Additionally, the visualized M5P regression tree model revealed a relationship between the total N and preceding crop yield and the WW yield variation (Figure 4.3). This result corroborates the previous report that the total N content in soil and the allocation of residual N of the preceding crop within the soil matrix affect the yield of the subsequent crop [69]. The SOC also appeared in the M5P regression model, with a small contribution (Table 4.4). The increase in the SOC content positively influenced the grain yields in plots that received NF input from 35 to 175 kg $ha^{-1}$. Similar positive correlations in the relationship between the grain yield and SOC were reported in other studies [70].

### 4.4.3. Comparing Models and Model Fits

The ANOVA test used in this work is a basic step in statistical inferences to understand yield differences between treatments using the F-test and p-value in the fixed model-general linear model. Therefore, the analysis only indicated fixed factors such as NF input and FYM as predictors of the WW grain yield. Nonetheless, when the trial years were considered as a fixed factor in the general linear model, the ANOVA result reveals the environment as an additional main determinant of the WW grain yield.

The LMM and M5P regression tree models were compared for their effectiveness in explaining the grain yield of WW. The LMM had better predictive performance compared to the M5P regression tree model, as indicated by its smaller RMSE and MAE (Table 4.4). This is because the LMM is an advanced statistical inference model that includes fixed and random factors and thus reduces experimental errors and increases the predictive performance. Second, the data used in this study were collected from well-designed experiments and thus suit a traditional model, such as LMM. Similar to the findings of this study, Krupnik et al. [17] observed that LMMs had better predictive performances compared to random forests and CART models for wheat grain yields in farm trials in Bangladesh. In contrast, our results slightly differ from the findings of Sihag et al. [71], whose field unsaturated hydraulic conductivity study revealed that M5P and random forest regression analyses provide better prediction efficiencies compared to the multiple nonlinear regression model. Additionally, the decision tree model generated by the M5P algorithm in a study by Trajanov et al. [30] achieved a better predictive performance of primary productivity in LTFEs compared to statistical studies previously carried out on the same data [72,73].

Although the M5P regression tree had a lower predictive performance than that of the LMM in this study, both models generally indicated a good fit with the actual yields (Table 4.4, Figure S 4.1 and S 4.2). The main results and factors identified in the LMM and M5P regression tree basically agree with each other. However, the M5P regression tree showed an intuitive visualization and interpretation of the main effects and interactions beyond their representations of single-degree of freedom contrasts [32]. Additionally, the M5P regression tree identified variables that were not captured by the LMM model. The SOC and FYM variables that showed up in the M5P regression tree analysis were not captured by the traditional statistical methods. Conversely, two important weather parameters in summer-precipitation in June and days Tmax > 30°C in July-were important variables that explained the grain yield variation in the LMM but did not appear in the M5P regression tree. In the soil matrix, organic matter decomposition is stimulated by increased temperature in summer, resulting in the release of nutrients locked up in the litter. Additionally, decomposition is also dependent on soil moisture, and litter breakdown will potentially be enhanced at warm temperatures, especially after a rainfall event. Thus, the two exclusive variables of each model were related to each other, especially in terms of decomposition and nutrient release to plants.

Thus, our study revealed that although the M5P regression tree offered less formal statistical inference compared to the LMM, it complemented the output derived from the LMM in analyzing the complex factors and mechanisms influencing the grain yield variation (Table 4.S6). According to Loh [32], the traditional statistical methods cannot account for variables that have more than two levels because their interactions cannot be fully represented by low-order contrasts.

Overall, our findings suggest that in addition to using the traditional ANOVA and the LMM to explain WW yields in LTFEs, as in earlier studies, the M5P regression tree could be used to produce a good prediction of WW yields as well. Thus, the co-use of these different analytical methods can strengthen the statement of the analysis by capturing other relevant variables overlooked by either of the models.

**4.5. Conclusions**

The grain yields of WW varied among the trial years, and an optimum grain yield of less than 5.0 Mg DM $ha^{-1}$ was observed. NF application and freezing days in December were identified as the main determinants of the WW grain yield. The combined fertilizer application with NF input enhanced the yield stability of WW. Additionally, the temperature in October, freezing days in February, precipitation in June, days Tmax > 30°C in July, the yield of the preceding crop, total N in the soil, SOC and FYM were important variables that explained the grain yield variation of WW.

The results of ANOVA provide the main factors affecting the WW yield. While the M5P showed a lower predictive performance compared to the LMM, it complemented the output from the LMM by revealing important yield predictors that were not captured by the LMM.

Thus, the co-use of different analytical methods such as ANOVA, LMM, and M5P model for the inference and prediction of yield responses in long-term studies should be considered, especially in those involving a larger database with multiple variables. The present finding adds more insights to the available literature by exhibiting the advantage of using various methods to analyze factors that affect the grain yield of WW in the LTFE. In addition, the results of this study indicate the need for adjustments in the management and exploration of appropriate preceding crops and/or the usage of appropriate wheat cultivars to adapt to year-to-year weather changes such as drought events and high temperatures in summer and winter. Further research with other crops and, ideally, with data obtained across many more years involving multiple variables is required to validate our observation.

**Supplementary Materials:**

The following are available online at www.mdpi.com/xxx/s1, Table S 4.1: Description of the experimental treatments, Table S 4.2: Cropping sequencing in the long-term experiment (LTFE) "V140", Table S 4.3: Selected chemical soil parameters in the topsoil (0–25 cm) of each treatment through four WW seasons (1984, 1988, 1992, and 1994), Table

S 4.4: Analyzed input variables for their effects on the grain yield of winter wheat by LMM and M5P models, Table S 4.5a: Average monthly temperature during WW growing season in the trial years, Table S 4.5b: Average monthly precipitation during WW growing season in the trial years, Table S 4.5c: Cumulative freezing days (freezing days) and the cumulative number of days recorded having temperatures above 30°C (days Tmax > 30°C) in selected months during winter wheat growing season in the trial years, Table S 4.6: Pearson's correlation of fixed effects in the LMM, Figure S 4.1: Modeled vs observed 1:1 scatter plots of the LMM, Figure S 4.2: Modeled vs observed 1:1 scatter plots of the M5P, EQ 4.1. Equations for the evaluation metrics.

**Author Contributions:** Conceptualization, T.H.T., S.D.B.-K.; data curation, D.B., T.H.T.; methodology, T.H.T., S.D.B.-K.; software, T.H.T.; formal analysis, T.H.T.; validation, T.H.T., S.D.B.-K., R.A.O.; investigation, T.H.T.; resources, S.D.B.-K.; writing–original draft preparation, T.H.T.; writing–review and editing, T.H.T., R.A.O., S.D.B.-K., and D.B., visualization, T.H.T., R.A.O.; supervision, S.D.B.-K., project administration, S.D.B.-K.; funding acquisition, T.H.T., S.D.B.-K. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was financially supported by the Vietnamese Government through the 911 project. The preparation and publication of this article was funded by the Leibniz Center for Agricultural Landscape Research (ZALF).

**Acknowledgments:** We would like to thank the Müncheberg Experimental Station, ZALF, for providing the data that were used in this analysis. The authors also acknowledge the support of the Strong Research Group Program of Hue University.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

**References**

1. FAO. FAOSTAT Online Database. Available online: http://faostat.fao.org/ (accessed on 10 December 2019).
2. Destatis. Available online: https://www.destatis.de/DE/PresseService/Presse/Pressemitteilungen/2016/12/PD16_470_412.html (accessed on 25 March 2020).
3. Ahrends, H.E.; Eugster, W.; Gaiser, T.; Rueda-Ayala, V.; Hüging, H.; Ewert, F.; Siebert, S. Genetic yield gains of winter wheat in Germany over more than 100 years (1895–2007) under contrasting fertilizer applications. Environ. Res. Lett. 2018, 13, 104003.
4. Laidig, F.; Piepho, H.-P.; Rentel, D.; Drobek, T.; Meyer, U.; Huesken, A. Breeding progress, variation, and correlation of grain and quality traits in winter rye hybrid

and population varieties and national on-farm progress in Germany over 26 years. Theor. Appl. Genet. 2017, 130, 981–998.

5. Macholdt, J.; Honermeier, B. Stability analysis for grain yield of winter wheat in a long-term field experiment. Arch. Agron. Soil Sci. 2019, 65, 686–699.

6. Macholdt, J.; Piepho, H.-P.; Honermeier, B. Mineral NPK and manure fertilization affecting the yield stability of winter wheat: Results from a long-term field experiment. Eur. J. Agron. 2019, 102, 14–22.

7. Rasmussen, P.E.; Goulding, K.W.; Brown, J.R.; Grace, P.R.; Janzen, H.H.; Körschens, M. Long-term agroecosystem experiments: Assessing agricultural sustainability and global change. Science 1998, 282, 893–896.

8. Wessolek, G.; Asseng, S. Trade-off between wheat yield and drainage under current and climate change conditions in northeast Germany. Eur. J. Agron. 2006, 24, 333–342.

9. Wechsung, F.; Gerstengarbe, F.W.; Lasch, P.; Lüttger, A. Die Ertragsfähigkeit Deutscher Ackerflächen Unter Klimawandel. Available online: www.pik-potsdam.de/glowa/pdf/bvvg/zusammenfassung11_9.pdf (accessed on 14 December 2019).

10. Ellmer, F.; Baumecker, M. 65 Years long-term experiments at Thyrow. Results for sustainable crop production at sandy soils. Arch. Agron. Soil Sci. 2002, 48, 521–531.

11. Ellmer, F.; Erekul, O.; Kohn, W. Influence of long-term different organic-mineral fertilization on yield, yield structure and bread-making quality of winter wheat. Arch. Agron. Soil Sci. 2001, 47, 423–444.

12. Verch, G.; Kächele, H.; Höltl, K.; Richter, C.; Fuchs, C. Comparing the profitability of tillage methods in Northeast Germany—A field trial from 2002 to 2005. Soil Tillage Res. 2009, 104, 16–21.

13. Ellerbrock, R.H.; Hohn, A.; Rogasik, J. Functional analysis of soil organic matter as affected by long-term manurial treatment. Eur. J. Soil Sci. 1999, 50, 65–71.

14. Chmielewski, F.M.; Kohn, W. Impact of weather on yield components of spring cereals over 30 years. Agric. For. Meteorol. 1999, 96, 49–58, doi:10.1016/s0168-1923(99)00047-7.

15. Chmielewski, F.M.; Köhn, W. Impact of weather on yield components of winter rye over 30 years. Agric. For. Meteorol. 2000, 102, 253–261, doi:10.1016/S0168-1923(00)00125-8.

16. Gauch, H.G. Statistical analysis of yield trials by AMMI and GGE. Crop Sci. 2006, 46, 1488–1500.

17. Krupnik, T.J.; Ahmed, Z.U.; Timsina, J.; Yasmin, S.; Hossain, F.; Al Mamun, A.; Mridha, A.I.; McDonald, A.J. Untangling crop management and environmental influences on wheat yield variability in Bangladesh: An application of non-parametric approaches. Agric. Syst. 2015, 139, 166–179.

18. Yang, R.-C. Towards understanding and use of mixed-model analysis of agricultural experiments. Can. J. Plant Sci. 2010, 90, 605–627.

19. Virk, D.; Witcombe, J. Evaluating cultivars in unbalanced on-farm participatory trials. Field Crops Res. 2008, 106, 105–115.

20. Bates, D.; Mächler, M.; Bolker, B.; Walker, S. Fitting linear mixed-effects models using lme4. J. Stat. Softw. 2015. , 67, 1-51, doi: 10.18637/jss.v067.i01.

21. Bolker, B.M.; Brooks, M.E.; Clark, C.J.; Geange, S.W.; Poulsen, J.R.; Stevens, M.H.H.; White, J.-S.S. Generalized linear mixed models: A practical guide for ecology and evolution. Trends Ecol. Evol. 2009, 24, 127–135.

22. Hopfield, J.J. Neural networks and physical systems with emergent collective computational abilities. PNAS 1982, 79, 2554–2558.

23. Breiman, L. Random forests. Mach. Learn. 2001, 45, 5–32.

24. Breiman, L.; Friedman, J.; Olshen, R.; Stone, C. Classification and regression trees. Wadsworth Int. Group 1984, 37, 237–251.

25. Piepho, H. Analysing disease incidence data from designed experiments by generalized linear mixed models. Plant Pathol. 1999, 48, 668–674.

26. Lobell, D.B.; Ortiz-Monasterio, J.I.; Asner, G.P.; Naylor, R.L.; Falcon, W.P. Combining field surveys, remote sensing, and regression trees to understand yield variations in an irrigated wheat landscape. Agron. J. 2005, 97, 241–249.

27. Zheng, H.; Chen, L.; Han, X.; Zhao, X.; Ma, Y. Classification and regression tree (CART) for analysis of soybean yield variability among fields in Northeast China: The importance of phosphorus application rates under drought conditions. Agric. Ecosyst. Environ. 2009, 132, 98–105.

28. Ferraro, D.O.; Rivero, D.E.; Ghersa, C.M. An analysis of the factors that influence sugarcane yield in Northern Argentina using classification and regression trees. Field Crops Res. 2009, 112, 149–157.

29. Chlingaryan, A.; Sukkarieh, S.; Whelan, B. Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. Comput. Electron. Agric. 2018, 151, 61–69.

30. Trajanov, A.; Spiegel, H.; Debeljak, M.; Sandén, T. Using data mining techniques to model primary productivity from international long-term ecological research (ILTER) agricultural experiments in Austria. Reg. Environ. Chang. 2019, 19, 325–337.

31. Song, Y.-Y.; Ying, L. Decision tree methods: Applications for classification and prediction. Shanghai Arch. Psychiatry 2015, 27, 130.

32. Loh, W.-Y. Regression tree models for designed experiments. In Optimality, Institute of Mathematical Statistics; 2006, 49, 210–228.

33. Bzdok, D.; Altman, N.; Krzywinski, M. Points of Significance: Statistics versus Machine Learning. Nat methods, 2018, 15, 233–234., doi:10.1038/nmeth.4642.

34. Barkusky, D. Long-Term Field Experiment V140 Muencheberg from 1963 to 2009-Plots (Version 1.0); BonaRes Data Centre (Leibniz Centre for Agricultural Landscape Research (ZALF)), Germany: 2018; doi:10.20387/BONARES-BSVY-R418.

35. Mirschel, W.; Barkusky, D.; Hufnagel, J.; Kersebaum, K.C.; Nendel, C.; Laacke, L.; Luzi, K.; Rosner, G. Coherent multi-variable field data set of an intensive cropping system for agro-ecosystem modelling from Müncheberg, Germany. Open Data J. Agric. Res. 2016, 2, 6-15, doi.org/10.18174/odjar.v2i1.15412.

36. Thai, T.H.; Bellingrath-Kimura, S.D.; Hoffmann, C.; Barkusky, D. Effect of long-term fertilizer regimes and weather on spring barley yields in sandy soil in North-East Germany. Arch. Agron. Soil Sci. 2020, 66, 1812–1826.

37. DWD. DWD Station Number 03376. Available online: https://opendata.dwd.de/climate_environment/CDC/observations_germany/climate/daily/kl/historical/ (accessed on 25 October 2019).

38. R Core Team. A Language and Environment for Statistical Computing; R Foundation for Statistical Computing: Vienna, Austria, 2020. URL: http://www.r-project.org/index.html

39. Grömping, U. Relative importance for linear regression in R: The package relaimpo. J. Stat. Softw. 2006, 17, 1–27.

40. Quinlan, J.R. Learning with continuous classes. In Proceedings of the 5th Australian Joint Conference on Artificial Intelligence, Hobart, Tasmania, 16–18 November 1992; pp. 343–348.

41. Wang, Y.W.; Ian, H. Inducing model trees for continuous classes. In Proceedings of the 9th European Conference on Machine Learning, Prague, Czech Republic, 23–25 April 1997.

42. Azzeh, M. Software effort estimation based on optimized model tree. In Proceedings of the 7th International Conference on Predictive Models in Software Engineering, Banff, AB, Canada, 20–21 September 2011; pp. 1–8.

43. Witten, I.H.; Frank, E.; Hall, M.A.; Pal, C.J. Data Mining: Practical Machine Learning Tools and Techniques, 4th ed.; Morgan Kaufmann (MK) Publishers is an imprint of

Elsevier: 30 Corporate Drive, Suite 400, Burlington, MA 01803, USA, 2016; ISBN 9780128042915.

44. Thi Kieu Tran, T.; Lee, T.; Shin, J.-Y.; Kim, J.-S.; Kamruzzaman, M. Deep learning-based maximum temperature forecasting assisted with meta-learning for hyperparameter optimization. Atmosphere 2020, 11, 487.

45. Fixen, P.E.; West, F.B. Nitrogen fertilizers: Meeting contemporary challenges. J. Hum. Environ. 2002, 31, 169–177.

46. Blanchet, G.; Gavazov, K.; Bragazza, L.; Sinaj, S. Responses of soil properties and crop yields to different inorganic and organic amendments in a Swiss conventional farming system. Agric. Ecosyst. Environ. 2016, 230, 116–126, doi:10.1016/j.agee.2016.05.032.

47. Barzegar, A.; Yousefi, A.; Daryashenas, A. The effect of addition of different amounts and types of organic materials on soil physical properties and yield of wheat. Plant Soil 2002, 247, 295–301.

48. Haynes, R.J.; Naidu, R. Influence of lime, fertilizer and manure applications on soil organic matter content and soil physical conditions: A review. Nutr. Cycl. Agroecosyst. 1998, 51, 123–137.

49. Omari, R.A.; Aung, H.P.; Hou, M.; Yokoyama, T.; Onwona-Agyeman, S.; Oikawa, Y.; Fujii, Y.; Bellingrath-Kimura, S.D. Influence of different plant materials in combination with chicken manure on soil carbon and nitrogen contents and vegetable yield. Pedosphere 2016, 26, 510–521, doi:10.1016/S1002-0160(15)60061-3.

50. Rasmussen, I.S.; Dresbøll, D.B.; Thorup-Kristensen, K. Winter wheat cultivars and nitrogen (N) fertilization-effects on root growth, N uptake efficiency and N use efficiency. Eur. J. Agron. 2015, 68, 38–49.

51. Acevedo, E.; Silva, P.; Silva, H. Wheat growth and physiology. In Bread Wheat, Improvement and Production; Fao. org, Rome, Italy, 2002; p. 30.

52. Evans, L. T.; Wardlaw, I.; Fischer, R.Crop Physiology: Some Case Histories; Evans, L., Ed.; the Syndics of Cambridge University Press: Cambridge, UK, 1975.

53. Zheng, D.; Yang, X.; Mínguez, M.I.; Mu, C.; He, Q.; Wu, X. Effect of freezing temperature and duration on winter survival and grain yield of winter wheat. Agric. For. Meteorol. 2018, 260, 1–8.

54. Bergjord Olsen, A.; Persson, T.; de Wit, A.; Nkurunziza, L.; Sindhøj, E.; Eckersten, H. Estimating winter survival of winter wheat by simulations of plant frost tolerance. J. Agron. Crop Sci. 2018, 204, 62–73.

55. Barlow, K.; Christy, B.; O'leary, G.; Riffkin, P.; Nuttall, J. Simulating the impact of extreme heat and frost events on wheat crop production: A review. Field Crops Res. 2015, 171, 109–119.

56. Lollato, R.; Knapp, M.; Redmond, C. Factors to Consider in Winter Survival of Wheat. Agronomy eUpdates-Kansas State University. 2019. Available online: https://webapp.agron.ksu.edu/agr_social/article_new/factors-to-consider-in-winter-survival-of-wheat-362-1 (accessed on 7 June 2020).

57. Persson, T.; Bergjord Olsen, A.; Nkurunziza, L.; Sindhöj, E.; Eckersten, H. Estimation of crown temperature of winter wheat and the effect on simulation of frost tolerance. J. Agron. Crop Sci. 2017, 203, 161–176.

58. WCG. World Climate Guide—Climate Germany. Available online: https://www.climatestotravel.com/climate/germany#:~:text=In%20most%20of%20Germany%2C%20the,F)%20in%20July%20and%20August (accessed on 21 April 2020).

59. Altenbach, S.; DuPont, F.; Kothari, K.; Chan, R.; Johnson, E.; Lieu, D. Temperature, water and fertilizer influence the timing of key events during grain development in a US spring wheat. J. Cereal Sci. 2003, 37, 9–20.

60. Pradhan, G.P.; Prasad, P.V.; Fritz, A.K.; Kirkham, M.B.; Gill, B.S. Effects of drought and high temperature stress on synthetic hexaploid wheat. Funct. Plant Biol. 2012, 39, 190–198.

61. Wiebe, K.; Lotze-Campen, H.; Sands, R.; Tabeau, A.; van der Mensbrugghe, D.; Biewald, A.; Bodirsky, B.; Islam, S.; Kavallari, A.; Mason-D'Croz, D. Climate change impacts on agriculture in 2050 under a range of plausible socioeconomic and emissions scenarios. Environ. Res. Lett. 2015, 10, 085010.

62. Lobell, D.B.; Sibley, A.; Ortiz-Monasterio, J.I. Extreme heat effects on wheat senescence in India. Nat. Clim. Change. 2012, 2, 186–189.

63. Meng, T.; Carew, R.; Florkowski, W.J.; Klepacka, A.M. Analyzing temperature and precipitation influences on yield distributions of canola and spring wheat in Saskatchewan. J. Appl. Meteorol. Climatol. 2017, 56, 897–913.

64. Shimoda, S.; Hamasaki, T.; Hirota, T.; Kanno, H.; Nishio, Z. Sensitivity of wheat yield to temperature changes with regional sunlight characteristics in eastern Hokkaido. Int. J. Climatol. 2015, 35, 4176–4185.

65. Spiertz, J. Relation between green area duration and grain yield in some varieties of spring wheat. Neth. J. Agr. Sci. 1971, 19, 211–222.

66. Hejcman, M.; Kunzova, E. Sustainability of winter wheat production on sandy-loamy Cambisol in the Czech Republic: Results from a long-term fertilizer and crop rotation experiment. Field Crops Res. 2010, 115, 191–199, doi:10.1016/j.fcr.2009.11.004.

67. Kunzová, E.; Hejcman, M. Yield development of winter wheat over 50 years of FYM, N, P and K fertilizer application on black earth soil in the Czech Republic. Field Crops Res. 2009, 111, 226–234.

68. Kunzová, E.; Hejcman, M. Yield development of winter wheat over 50 years of nitrogen, phosphorus and potassium application on greyic Phaeozem in the Czech Republic. Eur. J. Agron. 2010, 33, 166–174.

69. Evans, J.; Scott, G.; Lemerle, D.; Kaiser, A.; Orchard, B.; Murray, G.; Armstrong, E. Impact of legume 'break' crops on the yield and grain quality of wheat and relationship with soil mineral N and crop N content. Aust. J. Agric. Res. 2003, 54, 777–788.

70. Seremesic, S.; Milosev, D.; Djalovic, I.; Zeremski, T.; Ninkov, J. Management of soil organic carbon in maintaining soil productivity and yield stability of winter wheat. Plant Soil Environ. 2011, 57, 216–221.

71. Sihag, P.; Karimi, S.M.; Angelaki, A. Random forest, M5P and regression analysis to estimate the field unsaturated hydraulic conductivity. Appl. Water Sci. 2019, 9, 129.

72. Lehtinen, T.; Dersch, G.; Söllinger, J.; Baumgarten, A.; Schlatter, N.; Aichberger, K.; Spiegel, H. Long-term amendment of four different compost types on a loamy silt Cambisol: Impact on soil organic matter, nutrients and yields. Arch. Agron. Soil Sci. 2017, 63, 663–673.

73. Spiegel, H.; Sandén, T.; Dersch, G.; Baumgarten, A.; Gründling, R.; Franko, U. Soil organic matter and nutrient dynamics following different management of crop residues at two sites in Austria. In Soil Management and Climate Change; Elsevier: Amsterdam, 2018; pp. 253–265.

**Supplement**

**Table S 4.1.** Description of the experimental treatments

| Treatment Code | Group treatment | Mineral nitrogen fertilizer-NF (kg $ha^{-1}$) | Organic fertilizer-ORF | Fertilizer application |
|---|---|---|---|---|
| 0 | Control | 0 | 0 | 0 |
| 1.1 | NPK | 35 | 0 | MF |
| 1.2 | | 70 | | |
| 1.3 | | 105 | | |
| 1.4 | | 140 | | |
| 1.5 | | 175 | | |
| 2.1 | NPK+fym1 | 35 | 1.2 t $ha^{-1}$ $year^{-1}$ DM farmyard manure | fym1 |
| 2.2 | | 70 | | |
| 2.3 | | 105 | | |
| 2.4 | | 140 | | |
| 2.5 | | 175 | | |
| 3.1 | PK+fym2 | 0 | 3.2 t $ha^{-1}$ $year^{-1}$ DM farmyard manure | fym2 |
| 3.2 | NPK+fym2 | 35 | | |
| 3.3 | | 70 | | |
| 3.4 | | 105 | | |
| 3.5 | | 140 | | |
| 4.1 | NPK+Straw | 35 | 2.0 t $ha^{-1}$ $year^{-1}$ DM straw | straw |
| 4.2 | | 70 | | |
| 4.3 | | 105 | | |
| 4.4 | | 140 | | |
| 4.5 | | 175 | | |

Treatment codes (1.1-1.5; 2.1-2.5; 4.1-4.5): each rate of mineral nitrogen fertilizer-NF (five levels NF: 35, 70, 105, 140, 175 kg $ha^{-1}$, respectively) with organic fertilizer-ORF (three types: no ORF, 1.2 t dry mass (DM) $ha^{-1}$ farmyard manure (FYM) and 2.0 t DM $ha^{-1}$ straw). Treatment codes (3.1-3.5): each 3.2 t DM $ha^{-1}$ FYM with each NF level (five levels: 0, 35, 70, 105, and 140 kg $ha^{-1}$, respectively). Treatment code "0" or control: no fertilizer inputs. Fertilizer application (MF application: sole mineral fertilizer applied at 35, 70, 105, 140 and 175 kg $ha^{-1}$ N; fym1: FYM applied at 35, 70, 105, 140 and 175 kg $ha^{-1}$ N; fym2: FYM applied at 0, 35, 70, 105 and 140 kg $ha^{-1}$ N; straw: straw applied at 35, 70, 105, 140 and 175 kg $ha^{-1}$ N.

**Table S 4.2.** Cropping sequencing in the long-term experiment (LTFE) "V140". In bold: Winter wheat (WW), in grey: preceding crops

| Harvest year | Crop | Harvest year | Crop | Harvest year | Crop |
|---|---|---|---|---|---|
| 1963 | Maize | 1981 | Sugar beet | 1999 | Potato |
| 1964 | Winter rye | 1982 | Spring barley | 2000 | Spring barley |
| 1965 | Potato | 1983 | Potato | 2001 | Pea |
| 1966 | Winter rye | **1984** | **Winter wheat** | **2002** | **Winter wheat** |
| 1967 | Potato | 1985 | Sugar beet | 2003 | Maize |
| 1968 | Summer wheat | 1986 | Spring barley | 2004 | Winter rye |
| 1969 | Sugar beet | 1987 | Potato | 2005 | Flax |
| 1970 | Spring barley | **1988** | **Winter wheat** | 2006 | Winter rye |
| 1971 | Maize | 1989 | Sugar beet | 2007 | Potato |
| 1972 | Winter rye | 1990 | Spring barley | 2008 | Spring barley |
| 1973 | Potato | 1991 | Potato | 2009 | Pea |
| **1974** | **Winter wheat** | **1992** | **Winter wheat** | **2010** | **Winter wheat** |
| 1975 | Sugar beet | 1993 | Sugar beet | 2011 | Maize |
| 1976 | Spring barley | **1994** | **Winter wheat** | 2012 | Winter rye |
| 1977 | Sugar beet | 1995 | Maize | 2013 | Flax |
| 1978 | Spring barley | 1996 | Winter rye | 2014 | Winter rye |
| 1979 | Sugar beet | 1997 | Flax | 2015 | Potato |
| 1980 | Spring barley | 1998 | Winter rye | 2016 | Spring barley |

Thai *et al.* (2019) adapted

The position of WW within the crop system was changed during the LTFE. There were different WW crop rotations: potato-WW-sugar beet (04 rounds), sugar beet-WW-maize (01 round), pea-WW-maize (02 rounds). Between 1975-1982, no WW was seeded.

**Table S 4.3**. Selected chemical soil parameters in the topsoil (0-25 cm) of each treatment in four WW seasons (1984, 1988, 1992, and 1994)

| Treatment code | pH (KCl) | Total N (mg/100g soil) | SOC (mg/100g soil) | P (mg/100g soil) | K (mg/100g soil) | Mg ($CaCl_2$) (mg/100g soil) |
|---|---|---|---|---|---|---|
| 0 | 6.3$^{g}$ | 39.7$^{a}$ | 440.4$^{a}$ | 6.5$^{a}$ | 9.5$^{a-e}$ | 5.0$^{ab}$ |
| 1.1 | 6.2$^{fg}$ | 42.8$^{ab}$ | 470.2$^{ab}$ | 8.0$^{ab}$ | 11.3$^{b-f}$ | 5.0$^{ab}$ |
| 1.2 | 6.1$^{efg}$ | 42.9$^{ab}$ | 486.6$^{a-d}$ | 7.9$^{ab}$ | 10.0$^{a-e}$ | 4.7$^{ab}$ |
| 1.3 | 6.1$^{c-g}$ | 43.5$^{abc}$ | 480.6$^{abc}$ | 7.5$^{ab}$ | 9.4$^{abc}$ | 4.7$^{ab}$ |
| 1.4 | 6.0$^{a-e}$ | 44.2$^{a-d}$ | 490.8$^{a-d}$ | 7.6$^{ab}$ | 8.7$^{a}$ | 4.6$^{ab}$ |
| 1.5 | 5.9$^{a-d}$ | 45.0$^{a-d}$ | 498.1$^{a-e}$ | 8.0$^{ab}$ | 9.2$^{ab}$ | 4.4$^{a}$ |
| 2.1 | 6.3$^{g}$ | 46.8$^{bcd}$ | 507.1$^{b-e}$ | 9.0$^{bdc}$ | 11.4$^{b-g}$ | 4.8$^{ab}$ |
| 2.2 | 6.1$^{b-g}$ | 46.9$^{bcd}$ | 523.0$^{b-e}$ | 8.5$^{bc}$ | 10.2$^{a-e}$ | 4.5$^{ab}$ |
| 2.3 | 6.0$^{a-e}$ | 50.3$^{d-h}$ | 559.1$^{e-h}$ | 9.0$^{b-e}$ | 10.0$^{a-e}$ | 4.6$^{ab}$ |
| 2.4 | 6.0$^{a-e}$ | 49.3$^{b-g}$ | 548.8$^{d-g}$ | 8.9$^{bdc}$ | 10.0$^{a-e}$ | 4.5$^{ab}$ |
| 2.5 | 5.8$^{a}$ | 49.9$^{c-g}$ | 548.6$^{d-g}$ | 8.8$^{bcd}$ | 9.3$^{abc}$ | 4.4$^{a}$ |
| 3.1 | 6.1$^{d-g}$ | 50.5$^{d-h}$ | 548.2$^{d-g}$ | 10.8$^{ef}$ | 13.7$^{g}$ | 4.6$^{ab}$ |
| 3.2 | 6.1$^{d-g}$ | 56.8$^{h}$ | 618.3$^{h}$ | 11.0$^{f}$ | 13.0$^{fg}$ | 4.8$^{ab}$ |
| 3.3 | 6.1$^{d-g}$ | 55.4$^{gh}$ | 604.1$^{fgh}$ | 10.8$^{ef}$ | 11.9$^{d-g}$ | 5.0$^{ab}$ |
| 3.4 | 6.0$^{a-e}$ | 54.4$^{fgh}$ | 608.4$^{gh}$ | 10.2$^{c-f}$ | 11.6$^{c-g}$ | 4.8$^{ab}$ |
| 3.5 | 5.8$^{ab}$ | 54.0$^{e-h}$ | 596.1$^{fgh}$ | 10.4$^{def}$ | 10.8$^{a-f}$ | 4.5$^{ab}$ |
| 4.1 | 6.1$^{d-g}$ | 48.5$^{b-f}$ | 539.1$^{c-f}$ | 8.3$^{ab}$ | 11.9$^{efg}$ | 5.2$^{b}$ |
| 4.2 | 6.1$^{d-g}$ | 46.8$^{bcd}$ | 522.5$^{b-e}$ | 8.4$^{bc}$ | 10.5$^{a-e}$ | 4.7$^{ab}$ |
| 4.3 | 6.0$^{a-f}$ | 47.7$^{b-e}$ | 545.2$^{c-g}$ | 8.1$^{ab}$ | 10.3$^{a-e}$ | 4.8$^{ab}$ |
| 4.4 | 5.9$^{a-e}$ | 48.6$^{b-f}$ | 549.9$^{d-g}$ | 8.3$^{ab}$ | 9.8$^{a-e}$ | 4.7$^{ab}$ |
| 4.5 | 5.9$^{abc}$ | 48.0$^{b-f}$ | 543.5$^{c-g}$ | 8.2$^{ab}$ | 8.7$^{a}$ | 4.5$^{ab}$ |

Total nitrogen: total N; organic carbon: SOC; plant-available phosphorus: P (mg/100g soil); plant-available potassium: K (mg/100g soil); plant-available magnesium: Mg ($CaCl_2$) (mg/100g soil). Means sharing the same letters in the same column are not significantly different ($P < 0.05$). Treatments codes are given in Table S 4.1.

**Table S 4.4:** Analysed input variables for their effects on the grain yield of winter wheat by LMM and M5P models

| Input variable name | Unit |
|---|---|
| Fertilizers applied | |
| 1. Mineral nitrogen fertilizer | kg $ha^{-1}$ |
| 2. Farmyard manure fertilizer | tons |
| 3. Straw | tons |
| Monthly weather during the growing season | |
| 4. Monthly mean temperature | °C |
| 5. Cumulative freeze days in a month, | day |
| 6. Cumulative days Tmax > 30°C in a month | day |
| 7. Cumulative precipitation | mm |
| 8. Growing degree days (GDD). | GDD |
| Soil | |
| 9. Total nitrogen in soil | mg/100 g soil |
| 10. Soil organic carbon | mg/100 g soil |
| Crop yield | |
| 11. Winter rye | Mg DM $ha^{-1}$ |
| 12. Potatoes | Mg DM $ha^{-1}$ |
| 13. Sugar beets | Mg DM $ha^{-1}$ |
| 14. Pea | Mg DM $ha^{-1}$ |
| Preceding crop | |
| 15. Potatoes | None |
| 16. Sugar beets | None |
| 17 Pea | None |

**Table S 4.5a.** Average monthly temperature during WW growing season in the trial years

| Year | Sep | Oct | Nov | Dec | Jan | Feb | Mar | Apr | May | Jun | Jul |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1973/74 | 13.6 | 7 | 3.1 | 0 | 2.5 | 3.1 | 4.6 | 7.6 | 11.3 | 14.8 | 16 |
| 1983/84 | 14.2 | 9.3 | 3.6 | 0 | 1 | -0.6 | 1.9 | 7.6 | 12.6 | 14.1 | 16.1 |
| 1987/88 | 14 | 9.2 | 5.5 | 1.9 | 2.9 | 2.5 | 2.3 | 8.2 | 15.2 | 15.8 | 18.3 |
| 1991/92 | 14.7 | 7.9 | 3.6 | 0.8 | 0.4 | 2.9 | 4.4 | 8.3 | 14.9 | 18.8 | 19.9 |
| 1993/94 | 12 | 7.6 | -1.1 | 2.4 | 2.5 | -2.1 | 5 | 8.4 | 12.2 | 15.4 | 21.4 |
| 2001/02 | 12.5 | 12.4 | 3.8 | -0.4 | 1.5 | 4.8 | 4.9 | 8.1 | 15.1 | 17.1 | 19.1 |
| 2009/10 | 14.9 | 7.5 | 6.9 | -0.5 | -5.9 | -0.8 | 4.3 | 8.7 | 11 | 16.8 | 21.5 |
| Mean | 13.7 | 8.7 | 3.6 | 0.6 | 0.7 | 1.4 | 3.9 | 8.1 | 13.2 | 16.1 | 18.9 |
| SD | 1.1 | 1.8 | 2.5 | 1.1 | 3.0 | 2.5 | 1.3 | 0.4 | 1.8 | 1.6 | 2.3 |
| CV | 0.08 | 0.21 | 0.68 | 1.91 | 4.35 | 1.82 | 0.32 | 0.05 | 0.14 | 0.10 | 0.12 |
| Max | 14.9 | 12.4 | 6.9 | 2.4 | 2.9 | 4.8 | 5.0 | 8.7 | 15.2 | 18.8 | 21.5 |
| Min | 12.0 | 7.0 | -1.1 | -0.5 | -5.9 | -2.1 | 1.9 | 7.6 | 11.0 | 14.1 | 16.0 |

Sep to Jul is the short form of the month from September to July. SD: standard deviation; CV: coefficient variation; Max: maximum, Min: minimum.

**Table S 4.5b:** Average monthly precipitation during WW growing season in the trial years

| Year | Sep | Oct | Nov | Dec | Jan | Feb | Mar | Apr | May | Jun | Jul |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1973/74 | 24.7 | 60.2 | 47.8 | 49 | 29.9 | 42.9 | 10.5 | 17.4 | 46.6 | 74.8 | 62.4 |
| 1983/84 | 28.4 | 35.7 | 43.2 | 60.9 | 65.6 | 29.2 | 4.2 | 36.7 | 61.1 | 80.3 | 26.5 |
| 1987/88 | 72.4 | 10.2 | 52.5 | 47.6 | 44.4 | 83.4 | 65.7 | 1.6 | 23 | 116.9 | 33.6 |
| 1991/92 | 13.3 | 31.6 | 51.6 | 43.3 | 18 | 26.9 | 74.3 | 34.7 | 14 | 19.2 | 35.5 |
| 1993/94 | 73.3 | 22.8 | 37.4 | 105.8 | 81.3 | 7.3 | 88.1 | 48.3 | 88.5 | 35.5 | 45.4 |
| 2001/02 | 125.7 | 39.4 | 25.3 | 20.6 | 36.8 | 75.7 | 48.2 | 48.5 | 60.7 | 35.7 | 66.5 |
| 2009/10 | 38.3 | 79 | 63.8 | 42.4 | 12.7 | 13 | 33 | 22.9 | 85.6 | 5.4 | 131 |
| Mean | 53.7 | 39.8 | 45.9 | 52.8 | 41.2 | 39.8 | 46.3 | 30.0 | 54.2 | 52.5 | 57.3 |
| SD | 39.3 | 23.1 | 12.3 | 26.3 | 24.9 | 29.6 | 32.0 | 17.1 | 28.6 | 39.4 | 35.8 |
| CV | 0.73 | 0.58 | 0.27 | 0.50 | 0.60 | 0.74 | 0.69 | 0.57 | 0.53 | 0.75 | 0.62 |
| Max | 125.7 | 79.0 | 63.8 | 105.8 | 81.3 | 83.4 | 88.1 | 48.5 | 88.5 | 116.9 | 131.0 |
| Min | 13.3 | 10.2 | 25.3 | 20.6 | 12.7 | 7.3 | 4.2 | 1.6 | 14.0 | 5.4 | 26.5 |

Sep to Jul is the short form of the month from September to July. SD: standard deviation; CV: coefficient variation; Max: maximum, Min: minimum.

**Table S 4.5c:** Cumulative freezing days (freezing days) and the cumulative number of days recorded having temperatures above 30°C (days Tmax > 30°C) in selected months during winter wheat growing season in the trial years

| | Freezing days | | | | | Days Tmax > 30°C | |
|---|---|---|---|---|---|---|---|
| Year | Nov | Dec | Jan | Feb | Mar | Jun | Jul |
| 1973/74 | 6.0 | 12.0 | 7.0 | 4.0 | 0.0 | 0.00 | 0.00 |
| 1983/84 | 7.0 | 13.0 | 10.0 | 15.0 | 9.0 | 0.00 | 4.00 |
| 1987/88 | 0.0 | 14.0 | 5.0 | 3.0 | 6.0 | 0.00 | 2.00 |
| 1991/92 | 1.0 | 8.0 | 11.0 | 3.0 | 0.0 | 2.00 | 6.00 |
| 1993/94 | 14.0 | 4.0 | 3.0 | 15.0 | 1.0 | 2.00 | 14.00 |
| 2001/02 | 2.0 | 15.0 | 12.0 | 4.0 | 0.0 | 1.00 | 5.00 |
| 2009/10 | 0.0 | 14.0 | 30.0 | 17.0 | 8.0 | 0.00 | 11.00 |
| Mean | 4.3 | 11.4 | 11.1 | 8.7 | 3.4 | 0.7 | 6.0 |
| SD | 5.1 | 4.0 | 8.9 | 6.6 | 4.1 | 1.0 | 4.9 |
| CV | 1.20 | 0.35 | 0.80 | 0.75 | 1.19 | 1.33 | 0.82 |
| Max | 14.0 | 15.0 | 30.0 | 17.0 | 9.0 | 2.0 | 14.0 |
| Min | 0.0 | 4.0 | 3.0 | 3.0 | 0.0 | 0.0 | 0.0 |

Sep to Jul is the short form of the month from September to July. SD: standard deviation; CV: coefficient variation; Max: maximum, Min: minimum.

**Table S 4.6:** Pearson's correlation of fixed effects in the LMM

| | (Intr) | NF | N in soil | pcrp_y | Tmd_Oc | Tm30_J | Frz_Dc | Frz_Fb |
|---|---|---|---|---|---|---|---|---|
| NF | -0.460 | | | | | | | |
| N in soil | -0.237 | -0.074 | | | | | | |
| pcrop_yield | -0.283 | -0.156 | 0.432 | | | | | |
| Tmid_Oct | -0.383 | -0.067 | 0.375 | 0.421 | | | | |
| Tmax30_Jul | -0.012 | 0.105 | -0.273 | -0.674 | -0.591 | | | |
| Freeze_Dec | -0.143 | -0.006 | -0.490 | 0.065 | -0.500 | 0.368 | | |
| Freeze_Feb | -0.127 | -0.071 | 0.297 | 0.451 | 0.625 | -0.818 | -0.444 | |
| Preci_Jun | -0.014 | 0.112 | 0.004 | -0.730 | -0.353 | 0.787 | -0.105 | -0.480 |

NF: Nitrogen fertilizer rate; pcrop_yield: preceding crop yield; Tmid_Oct: Temperature in October; Tmax30_July: cumulative number of days in July with maximum temperatures above 30°C; Freeze_Dec: cumulative number of days in December with mean temperatures below 0°C (32°F); Freeze_Feb: cumulative number of days in February with mean temperatures below 0°C (32°F); Preci_Jun: precipitation in June

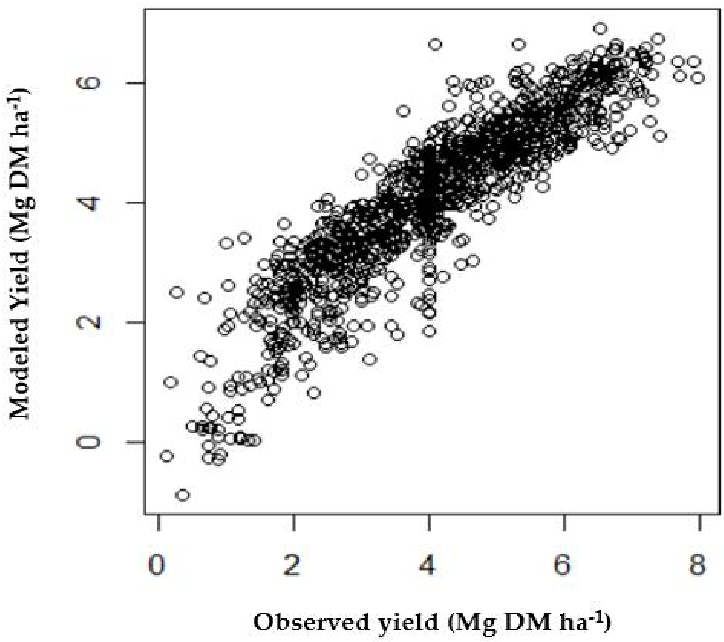



**Figure S 4.1:** Modeled vs observed 1:1 scatter plots of the LMM

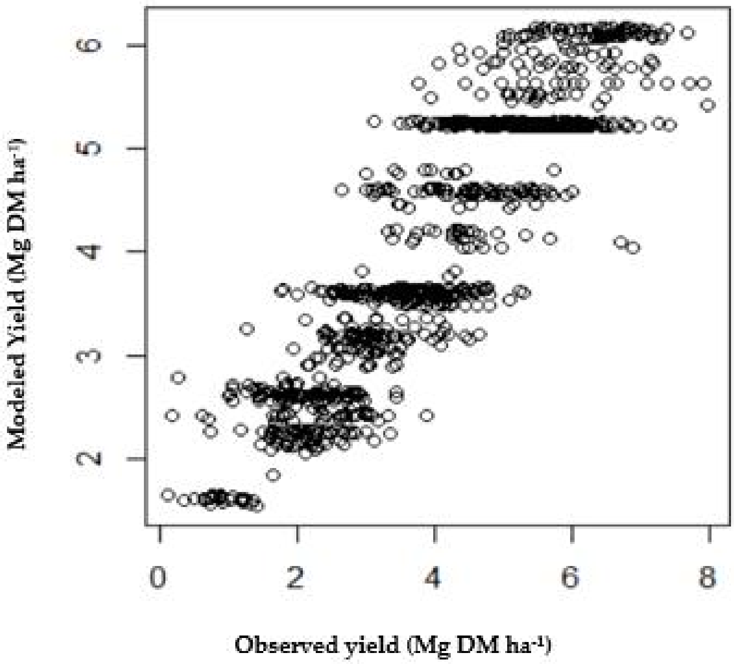



**Figure S 4.2:** Modeled vs observed 1:1 scatter plots of the M5P

**EQ 4.1. Equations for the evaluation metrics are given as:**

**1.1.** Equation of coefficient of determination ($R^2$) is given as:

$$R^2 = 1 - \frac{\sum(\bar{y} - \hat{y})^2}{\sum(\bar{y} - \hat{y})^2} \qquad \text{(Es1)}$$

**1.2.** Equation of root mean square error (RMSE)

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\bar{y} - \hat{y})^2} \qquad \text{(Es2)}$$

**1.3.** Equation of mean absolute error (MAE)

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|\bar{y} - \hat{y}| \qquad \text{(Es3)}$$

Where, ŷ = predicted value of y

ȳ = mean value of y

# Chapter 5

# General Discussion and Conclusion

## 5.1. Overview

Cereal crops are the major source of food and nutritional components for human and feed for livestock throughout the world (Maiti et al., 2014). The cereal yields and its variability among years are affected by weather conditions, soil properties, weeds, diseases, and pests (Gregory et al., 2009; Silungwe et al., 2019). In addition, management practices such as fertilizer applications, crop rotation, irrigation and tillage result in year-to-year yield variability (Brisson et al., 2010; Silungwe et al., 2018). Crop yield and yield variability therefore are the result of complex interdependencies and interactions among different factors. Thus, identifying both the major factors and relationships that account for crop yield and yield variability is important to understanding how to maximize crop yields and minimize yield fluctuations. This is a challenge that requires long-term investigation. Long-term field experiments (LTFEs) provide the necessary data and insights into identifying such factors and relationships and their influence on crop yield. Assessing annual yield variability of cereals requires the use of different analytical methods to clear the investigation. A comprehensive research employing the co-use of different analytical methods such as the statistical and machine learning models for the inference and prediction of the yield responses in long-term data is necessary.

## 5.2. Synthesis of findings

Figure 5.1 illustrates the overall thesis and how each objective is addressed in the scientific papers. All objectives of the thesis are addressed throughout the three papers (Thai et al., 2019; Thai et al., 2020a; Thai et al., 2020b) in chapter 2, 3 and 4.

Regarding objective 1, the findings revealed that yield response of the investigated cereals to fertilizer management involved complex relationships among climatic dependence, preceding crop, and soil characteristics (papers 1, 2 &3). Fertilizer was the most important factor determining winter rye (WR) yield (48%), the rates for winter wheat (WW) yield and spring barley (SB) yield were 34% and 11%, respectively (Table 5.1). Also, the study found out that the combined fertilizer application with chemical nitrogen fertilizer (NF) input can enhance the yield stability of cereal. A suitable combined application of NF and organic fertilizer produced better cereals yield than the application of either NF or organic fertilizer individually (papers 1, 2 &3), while NF application was identified as the main determinant of the grain yield of WR and WW (papers 2, 3). Besides NF input, farmyard manure fertilizer application was also a significant variable influencing the grain yield variation of the three tested cereals (papers 1, 2 &3)

Yield response of cereals

Analyze cereal grain yield responses to fertilizer management

Analyze sensitive timing of weather events

Compare different analysis models

Cereal yield response to fertilizer management involves complex relationships among climatic dependence, preceding crop, and soil characteristics (papers 1, 2 &3).

Weather at seeding and early growth stage of cereal (springtime for SB, autumn for WR, WW) influenced the grain yield. For winter cereals, the intensity and duration of extreme temperatures in the summertime (harvest year), especially the number of days recorded with a maximum temperature above 30°C in July was an important variable for the yield (papers 1, 2 &3).

Use of different analytical methods in the study can strengthen the statement of the analysis (papers 1, 2 &3).

**Synthesis 1**

- Fertilizer is the most important factor determining WR yield (48%), while the rates for WW yield and SB yield were 34 % and 11%, respectively (papers 1, 2, &3).
- The combined fertilizer application with NF input enhanced the yield stability of cereal. A suitable combined application of NF and organic fertilizer produced better cereal yield than the application of either NF or organic fertilizer alone (papers 1, 2 &3)
- For winter cereal: NF application was identified as the main determinants of the grain yield (papers 2 &3).

**Synthesis 2**

- Weather condition is the most important factor determining SB yield (55%) and WW yield (42%), while the weather condition influence WR yield by 32% (papers 1, 2 &3).
- For SB: precipitation rate in March and temperature in April negatively affected SB yield, while total precipitation during the growing season (April-July) positively affected SB yield when high NF application was supplied (paper 1).
- For WR: important variables for the yield were temperature in September and October, precipitation in November, temperature in December and May, cumulative days Tmax > 30°C in July (paper 2).
- For WW: temperature in October, the cumulative freezing days in December and February, precipitation in June, cumulative days Tmax > 30°C in July were important variables for the yield (paper 3).

**Synthesis 3**

- For spring barley: although soil and crop preceding variables have not been quantified in the BMA model, they could have an impact on SB yield (paper 1).
- For winter cereal (WR, WW): preceding crop, total N in the soil, and SOC were significant variables explaining the grain yield variation (papers 2 & 3).

**Synthesis 4**

- ANOVA results and the GLM provide only the target factor affecting cereal yield (papers 1, 2 & 3).
- BMA quantified in detail weather variables (main factor) influence SB yield, missed preceding crop, and soil variables in the model (paper 1).
- M5P model has well predictive performance as a further analysis after GLM to (i) unravel linear, non-linear interactions and combined effects on yield of winter cereal, and (ii) identify critical thresholds of explanatory the variables and their influence the cereal yield (papers 2 & 3).
- LMM showed a higher predictive performance compared to the M5P (paper 3).

**Figure 5.1.** Flow chart detailing the objectives and synthesis of results.

Regarding objective 2, the weather conditions at seeding and in the early growth stage of cereal (springtime for SB, autumn for WR, WW) influenced the grain yield (papers 1, 2 &3). For winter cereals, the intensity and duration of extreme temperatures in the summertime (harvest year), especially the number of days recorded with a maximum temperature above 30°C in July (cumulative days Tmax > 30°C in July) was an important variable that influenced the yield (papers 2, 3). Annual weather condition is the most important factor determining SB yield (55%) and WW yield (42%), while the weather condition influence WR yield by 32% (Table 5.1). For spring barley, BMA model showed that the precipitation rate in March and temperature in April negatively affected the grain yield. Meanwhile the total precipitation during the growing season (April-July) positively affected SB yield when high mineral NF application was supplied (paper 1). For WR, the M5P model showed that the important weather variables explaining the grain yield variation were temperature in September and October, precipitation in November, temperature in December and May, and cumulative days Tmax > 30°C in July. Among these, cumulative days Tmax > 30°C in July was the most important weather variable which influenced the WR yield (paper 2). For WW, the M5P and LMM showed that temperature in October, the cumulative number of freezing days in December and February, precipitation in June, and cumulative days Tmax > 30°C in July were important variables explaining the grain yield. Among these variables, the cumulative number of freezing days in December was the most important weather variable that influenced the WW yield (paper 3).

Along with fertilizer and weather condition, other agronomic factors such as soil parameter and preceding crop were also considered as factors that could be used in explaining the grain yield variation of cereals. The crop preceding and soil parameter could have an impact on grain yield of SB (paper 1), while the preceding crop type and the preceding crop yield, the total N in the soil, and the soil organic carbon were important variables that influenced grain yield of WR (paper 2) and winter wheat (paper 3).

Regarding objective 3, the study indicates that the results of ANOVA and GLM provided only the target factor affecting cereal yield (paper 1, 2 &3). BMA quantified in detail the weather variables (as main factor) which influenced SB yield, but missed important variables such as preceding crop and soil variables in the model (paper 1). M5P model has a well predictive performance as a further analysis after GLM to (i) unravel linear, non-linear interactions and combined effects on winter cereal yield, and (ii) identify critical thresholds of explanatory variables and their influence the winter cereal yield (paper 2, 3). LMM showed a higher predictive performance compared to the M5P (paper 3). However, co-use of different models such as ANOVA, LMM, and the M5P for the inference and prediction of yield responses of cereal in the LTFE can strengthen the statement of the analysis (paper 3).

The aforementioned findings revealed that all the analysis models employed in the three studies identified fertilizer regime application and annual weather conditions as the main factors that explained more than 60% variance in grain yield of each tested cereal (Table 5.1). In which the weather condition was the most important factor influencing grain yield of SB and WW in the experiment. While fertilizer is the most important factor explaining the grain yield of WR. This finding clear because usually, cool-season cereals such as barley and wheat have been reported as vulnerable plants to environmental stress factors such as heat stress, cold stress, drought, and water availability (Dolferus et al., 2011; Gooding et al., 2003; Jeyasri et al., 2021; Sallam et al., 2019). This point also discussed well in chapter 2 (paper 1) and chapter 4 (paper 3). Nutrient input was the main determinant of WR yield and confirms Schittenhelm et al., (2014), that WR is the most winter hardy and relatively drought tolerant crop with higher productivity compared to the other small grain crops when grown on the same soil. The soil condition in the trial site was sandy with poor nutrient profiles and low precipitation and thus there is a growing dependency on fertilization for increased cereal yield.

**Table 5.1.** Percentage of main factors effect on cereal grain yield

| **Cereal** | **Fertilizer (Fer.) (%)** | **Weather condition (Wea.) (%)** | **Fer. x Wea. (%)** | **Error (%)** |
|---|---|---|---|---|
| Spring barley | 11 | 55 | 8 | 26 |
| Winter rye | 48 | 32 | 11 | 9 |
| Winter wheat | 34 | 42 | 6 | 17 |

The study revealed that, assessing annual variability of cereals yield in long-term trial require using multiple analytical methods for robust analysis that will help strengthen the inferences. Chapter 2 reveals limitations in the adopted classical analytical methods (paper 1). For example, ANOVA showed only the main factors affecting the yields, but failed to exhaustively account for the other extraneous factors that influenced yields. Similarly, parametric correlations and linear regressions are less suited to handle missing, unbalanced, and higher-order data and nonlinear interactions (Krupnik et al., 2015; Virk and Witcombe, 2008; Yang, 2010). Thus, the study used a nonparametric BMA model to handle these challenges and address multiple weather variables that influenced SB yield. However, soil parameters and other agronomic factors such as preceding crops were missing in the model. Chapter 3 revealed the nonparametric M5P machine learning model to be a superior tool for inferring structural patterns from large, complex, and missing data as data usually in long-term experiments. Assessing grain yield variation of WR with the MP5 model primarily showed a positive approach to data mining. However, further statistical analysis would be required to obtain robust estimates for the designed experiment. The assumption as addressed in chapter 4 (paper 3) showed that the LMM had a better predictive performance compared to the M5P, with smaller root mean square and mean absolute errors. The LMM and M5P models were richer regressors than the ANOVA and the BMA. The M5P model presented an intuitive visualization of important variables and their critical thresholds and revealed other variables that were not captured by the LMM and vice versa. Hence, the use of different methods can strengthen the statement of the analysis. Thus, the co-use of the ANOVA/GLM, LMM, and M5P models should be considered especially in large databases involving multiple variables.

**5.3. Conclusion and recommendation**

This study concludes that seasonal weather forecasts and suitable sowing dates are important factors to consider for improving yields and reducing yield variability in SB and WW in sandy soil. It is thus essential to adjust the sowing dates to suitable times to ensure optimum growth of plants in spring. For SB, the study highlights that cold weather at the beginning of plant growth assures better SB yields in a temperate climate as well as seeds sowing before receiving heavy rains in the area. Otherwise, delayed sowing in spring due to soil saturation may lead to maturity of SB during the dry spells in summer, which will in turn affect its growth and yield. Also, the farmer needs to consider irrigation for SB during growth season if it is dry or has drought events. Similarly, there is a need to consider the weather conditions for WR seeding date and WW seeding date at the early growth stages. Besides the extreme weather in winter and summer also influenced the growth, development, and yield of the winter cereals. Therefore, it is necessary to adjust the management of appropriate preceding crops and/or the usage of appropriate wheat, and rye cultivars to adapt to year-to-year weather changes. Also, it is special to consider irrigation for WR and WW during summer if it is dry or having drought events.

On the other hand, supplying appreciable amounts of mineral NF and farmyard manure are important considerations for increased grain yield in WR. In addition, the current study highlighted need to consider the role and amount of nitrogen to choose the optimal amount and combination of mineral NF and farmyard manure in order to get a higher yield for each tested cereal in the sandy soil and dry region as the experimental site.

Decision trees by the M5P algorithm are useful supplementary tools for agronomists to devise different crop management intervention strategies such as fertilizer regimes, crop rotation, or seasonal arrangements to adapt to fluctuating weather conditions and dynamic soil fertility parameters over time. However, these techniques also pose considerable challenges to model fitting and subsequent interpretation. To meet the challenge, LTFE data should be analyzed more in detail by further statistical methods to devise suitable suggestions to support researcher-farmer-advisor dialogue on productivity management and the development and adoption of precision agriculture recommendations.

Therefore, the findings of the present studies suggest comprehensive co-use of different analysis models such as ANOVA/GLM, M5P, and LMM for the inference and prediction of yield responses of cereals. The finding adds more insights to the available literature by revealing the advantage of using various methods to analyze factors that affect the cereal grain yield from long-term data. There is the need for further research with other crops and, ideally, with data obtained across many more years involving multiple variables is required to validate our observation.

Overall, study contributes to the existing literature about comprehensive strategies for sustainable crop production with regard to climate change in the future.

**REFERENCES**

Brisson, N., Gate, P., Gouache, D., Charmet, G., Oury, F.-X., and Huard, F., 2010. Why are wheat yields stagnating in Europe? A comprehensive data analysis for France. Field Crops Research. 119, 201-212.

Dolferus, R., Ji, X., and Richards, R., 2011. Abiotic stress and control of grain number in cereals. Plant science. 181, 331-341.

Gooding, M.J., Ellis, R.H., Shewry, P.R., and Schofield, J. D., 2003. Effects of restricted water availability and increased temperature on the grain filling, drying and quality of winter wheat. 37, 295-309.

Gregory, P. J., Johnson, S. N., Newton, A. C., and Ingram, J. S., 2009. Integrating pests and pathogens into the climate change/food security debate. Journal of experimental botany 60, 2827-2838.

Jeyasri, R., Muthuramalingam, P., Satish, L., Pandian, S. K., Chen, J.-T., Ahmar, S., Wang, X., Mora-Poblete, F., and Ramesh, M., 2021. An Overview of Abiotic Stress in Cereal Crops: Negative Impacts, Regulation, Biotechnology and Integrated Omics. 10, 1472.

Krupnik, T. J., Ahmed, Z. U., Timsina, J., Yasmin, S., Hossain, F., Al Mamun, A., Mridha, A. I., and McDonald, A. J., 2015. Untangling crop management and environmental influences on wheat yield variability in Bangladesh: an application of non-parametric approaches. Agricultural Systems. 139, 166-179.

Maiti, R., Satya, P., 2014. Research advances in major cereal crops for adaptation to abiotic stresses. Biotechnology in Agriculture and the Food Chain. 5, 259-279.

Sallam, A., Alqudah, A. M., Dawood, M. F., Baenziger, P. S., and Börner, A., 2019. Drought stress tolerance in wheat and barley: advances in physiology, breeding and genetics research. Molecular Sciences. 20, 3137.

Schittenhelm, S., Kraft, M., and Wittich, K.-P., 2014. Performance of winter cereals grown on field-stored soil moisture only. European journal of agronomy. 52, 247-258.

Silungwe, F. R., Graef, F., Bellingrath-Kimura, S. D., Tumbo, S. D., Kahimba, F. C., and Lana, M. A., 2018. Crop upgrading strategies and modelling for rainfed cereals in a semi-arid climate-A review. Water. 10, 356.

Silungwe, F. R., Graef, F., Bellingrath-Kimura, S. D., Tumbo, S. D., Kahimba, F. C., and Lana, M. A., 2019. Analysis of intra and interseasonal rainfall variability and its effects on pearl millet yield in a semiarid agroclimate: Significance of scattered fields and tied ridges. Water. 11, 578.

Thai, T. H., Bellingrath-Kimura, S. D., Hoffmann, C., and Barkusky, D., 2019. Effect of long-term fertiliser regimes and weather on spring barley yields in sandy soil in North-East Germany. Archives of Agronomy and Soil Science. 66, 1-15.

Thai, T. H., Mamidanna, S., Barkusky, D., and Bellingrath-Kimura, S. D., 2020a. MP5 machine learning algorithm for analysis of winter rye yield in a long-term field experiment in Müncheberg, Germany. Under review. European Journal of Agronomy.

Thai, T. H., Omari, R. A., Barkusky, D., and Bellingrath-Kimura, S. D., 2020b. Statistical Analysis versus the M5P Machine Learning Algorithm to Analyze the Yield of Winter Wheat in a Long-Term Fertilizer Experiment. Agronomy. 10, 1779.

Virk, D., and Witcombe, J., 2008. Evaluating cultivars in unbalanced on-farm participatory trials. Field Crops Research. 106, 105-115.

Yang, R.-C., 2010. Towards understanding and use of mixed-model analysis of agricultural experiments. Canadian Journal of Plant Science. 90, 605-627.

**Statutory declaration**

I hereby declare that I completed the doctoral thesis on the topic: "Cereal grain yield responses to fertilizer management in sandy soil in a long-term fertilizer experiment in Northeast Germany", independently and used no other aids than those cited. In each individual case, I have clearly identified the source of the passages that are taken word for word or paraphrased from other works.

I equally declare that this dissertation has not been part of a doctoral procedure before and that this is my first doctoral program. I have not submitted the doctoral thesis or parts of it, to another academic institution and the thesis has not been accepted or rejected. I have not applied for a doctoral degree elsewhere and do not have a corresponding doctoral degree before.

I also hereby declare that I have carried out my scientific work according to the principles of good scientific practice in accordance with the current rules and regulations of Humboldt University of Berlin. Furthermore, I declare that no collaboration with commercial doctoral degree supervisors took place and that the principles of Humboldt-Universität zu Berlin for assuring Good Scientific Practice were abided by.

Berlin, 14 November 22

Thai, Thi Huyen

**Erklärung**

Ich erkläre, dass ich die Dissertation ausschließlich auf der Grundlage der angegebenen Hilfsmittel selbständig angefertigt habe.

Ich erkläre weiterhin, dass ich weder an der HU noch an einer anderen Universität bereits einen Promotionsantrag gestellt habe bzw. einen entsprechenden Doktorgrad besitze.

Ich erkläre außerdem, dass ich die dem angestrebten Verfahren zugrunde liegende

Promotionsordnung zur Kenntnis genommen habe.

Berlin, 14 November 2022

Thai, Thi Huyen