Statistical analysis and modelling of crop yield and nitrogen use efficiency in China

Yingxia Liu

Propositions

1. Space-time analysis and modelling of nitrogen use efficiency is meaningful for agricultural productivity, environmental sustainability and economic profit (this thesis).

2. Soil properties have a principal effect on nitrogen use efficiency (this thesis).

3. Food security is associated with chemical fertilizers.

4. The overuse of nitrogen fertilizer leads to soil acidification.

5. The random forest model has superior performance over the linear model.

6. The nutrient expert system is an efficient way to improve land management for smallholder farmers in China.

7. Climate change threats migration of wild animals.

8. People increasingly distrust news reports.

Propositions belonging to the thesis entitled:

Statistical analysis and modelling of crop yield and nitrogen use efficiency in China

Yingxia Liu

Wageningen, 19 May 2022

Statistical analysis and modelling of crop yield and nitrogen use efficiency in China

Yingxia Liu

Thesis committee

Promotors
Prof. Dr Gerard B.M. Heuvelink
Special Professor Pedometrics and Digital Soil Mapping
Wageningen University & Research

Prof. Dr Ping He
Professor of Nutrient Management
Chinese Academy of Agricultural Sciences, China

Co-promotor
Dr Zhanguo Bai, ISRIC – World Soil Information, the Netherlands

Other members
Prof. Dr Oene Oenema, Wageningen University & Research
Prof. Dr Shiwei Guo, Nanjing Agricultural University, China
Prof. Dr Kurt-Christian Kersebaum, The Leibniz Centre for Agricultural Research (ZALF), Germany
Dr Evert Jan Bakker, Wageningen University & Research

This research was conducted under the auspices of the C.T. de Wit Graduate School for Production Ecology and Resource Conservation (PE&RC).

Statistical analysis and modelling of crop yield and nitrogen use efficiency in China

Yingxia Liu

Thesis
submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof. Dr A. P. J. Mol,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Thursday 19 May 2022
at 11 a.m. in the Omnia Auditorium of Wageningen University

Yingxia Liu
Statistical analysis and modelling of crop yield and nitrogen use efficiency in China
189 pages.

PhD thesis, Wageningen University, Wageningen, the Netherlands (2022)
With references, with summary in English

ISBN https://doi.org/10.18174/565336

DOI 978-94-6447-124-3

Table of contents

Chapter 1

General introduction

1.1. Effects of nitrogen fertilizer on food security and environment

1.1.1. Contribution of nitrogen fertilizer application to food security

It has been and will continue to be a global challenge to fulfil the food demands of a growing population. It is predicted that the global food production will be 4.2 billion tons by 2050, which is an increase of 26% compared to that in 2020, and is relatively lower than the global food demand (Tian et al., 2021). The gap between food supply and demand is predicted to gradually increase from 390 million tons in 2020 to 1260 million tons in 2050 (Tian et al., 2021). In China, the peak food demand is predicted to occur in 2036 with the amount of 760 million tons (Du et al., 2020). The COVID-19 pandemic even makes thing worse, and 2.4 billion people suffered from food insecurity in the world in 2020. Half of these (1.2 billion) are in Asia; one-third (800 million) in Africa; and 11% (270 million) in Latin America & the Caribbean (Figure 1.1; FAO, 2021). Therefore, effective measures need to be taken to increase food production. Nitrogen (N) application is one of the most potential methods to accomplish this goal, since N plays a key role in crop growth and the introduction of N fertilizer realized a substantial increase in food production (Latiri-Souki et al., 1998). It has been proven that N contributed to more than 50% of the world's food production increase (Zhang et al., 2015; Dimkpa et al., 2020). Conversely, the excessive N application caused diminishing returns for increased crop yields and N removals (Zhang et al., 2015). That is to say, the nitrogen use efficiency (NUE) decreased. China applied the highest N rate in agriculture since 1990, even though it had a short decrease after 2015 due to the "Zero Growth in Chemical Fertilizer Use by 2020" policy ushered by the Chinese government in 2015 (Figure 1.2). Therefore, NUE in China needs our careful attention (Zhu et al., 2018).

Although China's one-child policy reduced population growth, it has now a large population base of 1.43 billion residents; since May 2021, the policy has been changed to allow a couple to have three children. Securing food demands of the population is a top priority of the government. The government has also drawn a 'red line' to protect the 1.2 billion ha of high-quality agricultural land for food security, and scientific communities have been looking for ways to boost yields through crop breeding, e.g., hybrid rice breeding by Longping Yuan, and advanced agronomic management to improve yield and resource use efficiency, including NUE. Zhang et al. (2016) indicated that sustainable high yield and economic gains could be achieved by a Science and Technology Backyard platform. Xu et al. (2016a) indicated that the Nutrient Expert tool could improve maize yield by 0.9 t ha^{-1} compared with common farmers' practices.

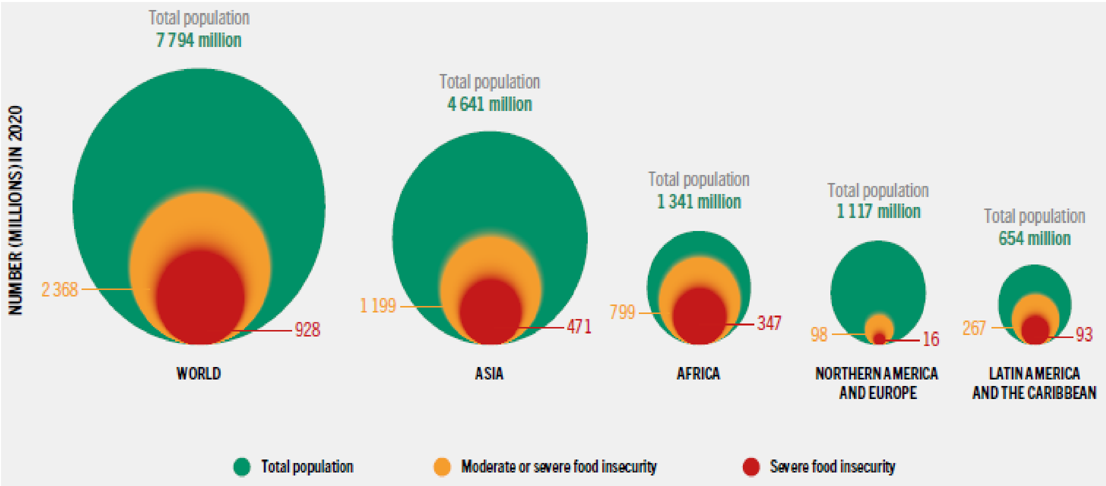


Figure 1.1: The concentration and distribution of food insecurity by severity differs greatly across the regions of the world (FAO, 2021).

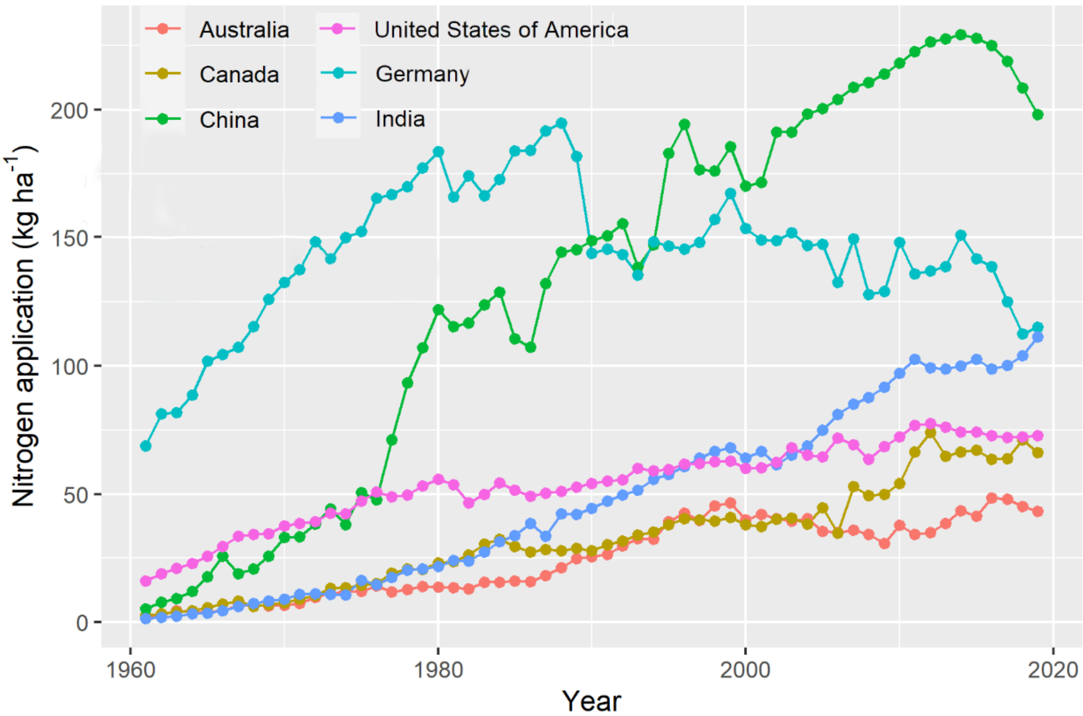


Figure 1.2: Nitrogen application in six main agricultural countries from 1960 to 2019 (FAO, 2021).

1.1.2. Environmental risk of nitrogen fertilizer application

Due to the large profit of N fertilizer application, farmers tend to apply excessive N fertilizer to obtain high yields and productivity. However, excessive use of N has a negative impact on environmental sustainability (Zhu et al., 2018). In the past decades, N has led to a series of environmental problems, such as soil acidity (Tian & Niu, 2015), greenhouse gas emission (Kahrl et al., 2010; Liu et al., 2015b), and eutrophication of surface waters (Reidsma et al., 2012). Furthermore, N fertilizer is produced in a chemical process that has a high energy cost (especially of natural gas) (Kliopova et al., 2016). Ren et al. (2021) found that low fixed inputs (e.g., machinery and knowledge) was the key factor leading to over-fertilization in smallholder farms in China.

1.2. Nitrogen use efficiency in the world and China

1.2.1. Nitrogen use efficiency definition and indicators

NUE is a crucial index in agricultural production and environmental sustainability. It is defined as the ratio of N output to N input and reflects how efficient the N application is for crop production. Technically, low NUE implies high N loss and, in contrast, high NUE indicates low N loss. NUE indicators include recovery efficiency, physiological efficiency, internal utilization efficiency, agronomic efficiency, partial factor productivity, and partial nutrient balance (Dobermann, 2007; He et al., 2015). All these indicators have their optimal range and interpretation for specific problems.

In this thesis, I used the partial factor productivity of N (PFP_N, in kilograms of grain per kilogram of N applied) and partial nutrient balance of N (PNB_N, in kilograms of nitrogen removal by aboveground crop per kilogram of nitrogen applied) to explore the long-term trend and spatial distribution of NUE in China. N application sources include chemical fertilizer and organic fertilizer (straw return, cake fertilizer, manure from livestock). It is worth noting that if PNB_N is higher than 1, it will deplete soil N, leading to decline in soil fertility.

1.2.2. Nitrogen use efficiency in the world

The spatial and temporal variation of NUE varies among countries and regions in the world. Some countries and regions have high NUE and maintain a sustainable range, such as America and Europe; while other countries, such as China, have a low NUE due to excessive N application, while countries in Africa have a low NUE because of low yields. Globally, PFP_N decreased from 68 kg kg^{-1} yr^{-1} in 1961 to 45 kg kg^{-1} yr^{-1} in 1980 and kept steady at 47 kg kg^{-1} yr^{-1} from 1981 to 2009 (Lassaletta et al., 2014). NUE is also different among crops: soybean had the highest PNB_N (0.80 kg kg^{-1}), followed by cereal crops (maize: 0.46 kg kg^{-1}, wheat: 0.42 kg kg^{-1}, rice: 0.39 kg kg^{-1}, others: 0.53 kg kg^{-1}), sugar crops (0.19 kg kg^{-1}), and fruits and vegetables (0.14 kg kg^{-1}) in 2010 globally (Zhang et al., 2015). From

1987 to 2006, PNB_N increased in Africa (from 1.2 to 1.4 kg kg^{-1} yr^{-1}) and Europe (from 0.40 to 0.60 kg kg^{-1} yr^{-1}) and decreased in China (from 0.50 to 0.40 kg kg^{-1} yr^{-1}) (Brentrup & Pallière, 2010). Overall, PFP_N was higher in European countries (e.g., Netherlands: 78 kg kg^{-1}, France: 76 kg kg^{-1} and Greece: 69 kg kg^{-1}) than in USA (45 kg kg^{-1}), Brazil (40 kg kg^{-1}), Egypt (34 kg kg^{-1}), China (30 kg kg^{-1}) and India (26 kg kg^{-1}) in 2009 (Lassaletta et al., 2014). There are two major reasons to explain the low NUE: 1) low N application and low food productivity (Africa); and 2) poor management and field conditions (China and India). Simply reducing N application rate or blindly pursuing high NUE (soil depletion) is not the right way forward. Instead, the aim is to obtain high yield with reasonable and sustainable NUE. Therefore, exploring key factors that influence NUE and yield is an urgent scientific topic.

1.2.3. Nitrogen use efficiency in China

NUE in China is lower than in other developed countries. To improve NUE and obtain a sustainable range, it is beneficial to understand its spatial and temporal variation. Yan et al. (2022) indicated that PNB_N generally decreased and then remained at low levels (0.20-0.35 kg kg-1 ^{y}r-1$^{)}$ in the economically developed provinces (e.g., Guangdong and Zhejiang) and undeveloped provinces (e.g., Yunnan and Guizhou) from 2005 to 2014. There was a pronounced increasing trend in agricultural provinces with high economic development (e.g., Shandong and Jiangsu) and intermediate economic growth (Hunan and Hebei). NUE varied in different crops; even within the same crop, different varieties could have different NUE. PNB_N increased gradually from 0.48 kg kg^{-1} in 1980 to 0.83 kg kg^{-1} in 2012 for wheat and increased from 0.43 kg kg^{-1} in 1980 to 0.68 kg kg^{-1} in 2002 and then declined to 0.61 kg kg^{-1} in 2014 for maize (Zhang et al., 2017). For maize, PFP_N was higher in the northeast provinces (Heilongjiang: 61.2 kg kg^{-1} $yr^{-1;}$ Liaoning: 59.1 kg kg^{-1} $yr^{-1;}$ Jilin: 51.7 kg kg^{-1} $yr^{-1)}$ than in northcentral China (Shanxi: 50.5 kg kg^{-1} $yr^{-1;}$ Shandong: 44.3 kg kg^{-1} $yr^{-1;}$ Henan: 39.4 kg kg^{-1} $yr^{-1;}$ Hebei: 32.1 kg kg^{-1} $yr^{-1)}$ from 2010 to 2012 (Xu et al., 2014b). For irrigated rice, PFP_N was higher in Jiangsu (37 kg kg^{-1}) and Hunan (34 kg kg^{-1}) than in Zhejiang (32 kg kg^{-1}) and Guangdong (27 kg kg^{-1}) in 2002 (Peng et al., 2006).

Few studies have been carried out to explore the influential factors on spatial and temporal variation in NUE. Possible reasons for temporal and spatial variation of NUE are farmer's income, agricultural management practices (AMP, e.g., irrigation, machinery, fertilizer, crop types and variety) and environmental factors (e.g., soil, climate, topography). Most existing studies explored the influence of management practices on NUE, such as N application rate (Qiu et al., 2015; Li et al., 2016b), application time (Yousaf et al., 2016), and straw return (Zhang et al., 2017). Some studies focused on population growth and economy (Gao et al., 2019), soil types (Xu et al., 2014b), and climate (Liang et al., 2018). Few studies were carried out on the influence of integrated factors on NUE.

1.3. Methods to analyze causal factors of nitrogen use efficiency and prediction uncertainty

1.3.1. Summary statistics

Summary statistics are used to summarize a set of observations. It communicates the largest amount of information as simply as possible and typically includes the following descriptive statistics: arithmetic mean and median (a measure of location or reflecting central tendency); standard deviation (a measure of statistical dispersion); skewness and kurtosis (measures of the shape of the distribution); quantiles or percentiles, and maximum/minimum value of the observation dataset (Bullen, 2013; Piazza et al., 2013; Alexander et al., 2014). Summary statistics can also be displayed by histograms, box plots and scatter plots, depending on the objectives.

1.3.2. Stepwise multiple linear regression

Multiple linear regression (MLR) is an extension of simple linear regression, that uses just one explanatory variable (Kolehmainen & Knuutinen, 1981; Rencher & Christensen, 2012). It is a statistical technique that uses multiple explanatory variables to predict the outcome of a response variable. Normally there are two broad categories that define the use of MLR. On the one hand, if the goal is prediction, forecasting, or error reduction, a predictive MLR model can be built with a calibration data set of the response and explanatory variables (Sousa et al., 2007; Khan et al., 2013). On the other hand, the goal might be to gain understanding of how explanatory variables explain variation in the response variable. MLR can quantify the strength of the relationship between the response and the explanatory variables, especially to determine whether some explanatory variables have no linear relationship with the response or to identify which subsets of explanatory variables may contain redundant information about the response (Chen et al., 2014; Popescu et al., 2016; Chang & Xu, 2021). Stepwise MLR (SMLR) reduces multi-collinearity of explanatory variables (Khanal et al., 2018) by an iterative process that continues to add or remove variables from the regression equation until there is no improvement (Dutta et al., 2015). It is widely used in developing empirical models from large data sets (Peerlinck et al., 2018).

1.3.3. (Quantile) Random Forest

Random forests (RF) or random decision forests is an ensemble learning method for classification, regression and other tasks. It operates by constructing a multitude of decision trees (Svetnik, 2003; Thongkam et al., 2008). For classification tasks, the output of the random forest is the class selected by most trees (Rodriguez-Galiano et al., 2012; Alencar et al., 2014), while for regression tasks, the mean or average prediction of the individual trees is returned (Lie et al., 2012; Bettinger, 2021). The first algorithm for RF was created in 1995 by Tin Kam Ho using the random subspace method (Ho, 1995, 1998). An extension of the algorithm was developed by Breiman (2001). It sacrifices the intrinsic interpretability present in decision trees (without a specific equation it becomes a 'black box'), but it usually achieves higher accuracy (Pereira et al., 2017; Orlenko & Moore, 2020). Quantile regression forests (QRF) considers the spread of the response variable from which prediction intervals are constructed, which can reflect model uncertainty (Meinshausen, 2006; Vaysse & Lagacherie, 2017).

1.3.4. Cross-validation

Cross-validation, sometimes called rotation estimation or out of sample testing, is a resampling method that uses different portions of the data to train and test a model in different iterations (Stone, 1977; Lendasse et al., 2003). Its goal is to estimate the model prediction ability for new data, in order to flag problems like overfitting or selection bias and to give insight into how a model will generalize to an independent data set (Yong & Liao, 2014).

1.3.5. Uncertainty propagation (Monte Carlo simulation)

Uncertainty propagation or propagation of error is the effect of variables' uncertainties or errors (more specifically, random errors) on the uncertainty of a model or function that uses these variables as input (Heuvelink, 1998b). When the variables are the values of experimental measurements, they may have uncertainties due to measurement limitations (e.g., instrument precision) which propagate through the function (Degiuli et al., 2007; Garamszegi, 2016). There are two methods to deal with uncertainty propagation: 1) the Taylor series method, which uses an analytical approach and mathematical formulae (Xu et al., 2006); and 2) Monte Carlo simulation (Jansen et al., 1994). The latter approach starts by sampling a large number of 'possible realities' from the probability distributions of the uncertain inputs, using a pseudo-random number generator. The model is run or the function applied to each of the simulated realities, and the variation in the model or function outputs reflect how input uncertainty has propagated. It is mainly used in three problem classes: optimization, numerical integration, and uncertainty analysis. It is simple, fast and highly-adaptable. It considers factors in a wide range of input values. In contrast, model risk will exist if it assumes the underlying risk factors (Bavajigari & Singh, 2019).

1.4. Problem definition (gaps)

NUE has large spatial and temporal variation in China, and its reasons need to be explored for improving NUE and keeping it at a sustainable range. According to the definition of PFP_N, yield influences NUE variation and it is also very important for food security. The space-time variation of crop yield and yield residual among different crops in China was not yet explained by MLR models in the literature.

NUE space-time variation could be analyzed through descriptive statistics, graphics and tables and so on; but it can also be done using statistical models, which could identify causal factors and their contribution to NUE variation. To the best of my knowledge, statistical modelling has not yet been applied for analyzing NUE variation in space and time in China at provincial and county scales.

No model is perfect, and model uncertainty deteriorates the quality of model predictions. Therefore, uncertainty of the model outcomes and the main sources of uncertainty need to be analyzed as well to evaluate the value of these outcomes and get insight into how model output quality could be improved.

1.5. Research objectives and questions

The overall objective of this thesis is to apply statistical methods to analyze and explain space-time patterns of crop yield and NUE in China at two spatial scales, i.e., provincial and county level, for supporting the development of effective strategies and policies. These policies could improve NUE in a sustainable way without impeding crop productivity. The models that are used for this analysis should take account of the trade-offs between economy, ecology and environment as well as climate change, land use change and N reallocation. I consider NUE to be optimal when it is sustainable and strikes a balance between agronomic productivity and the environment.

To achieve this objective, four sub-objectives (SO) and associated research questions were defined:

SO 1. Space-time statistical analysis and modelling of NUE indicators at provincial scale in China (**Chapter 2**).

- What are spatial and temporal patterns of NUE at provincial level?
- What prediction models for NUE are obtained using SMLR?
- What are the major explanatory variables of NUE as derived from the SMLR prediction model?
- How accurate are model predictions based on 10-fold cross-validation?
- Which recommendations can be drawn from the model to improve NUE in a sustainable way in China?

SO 2. Analysis of spatio-temporal variation of crop yield in China using stepwise multiple linear regression (**Chapter 3**).

- What are spatial and temporal patterns of yield at provincial level?
- What prediction models for crop yield are obtained using SMLR?
- What are the major explanatory variables of crop yield as derived from the prediction model?
- How accurate are model predictions based on 10-fold cross-validation?
- Which recommendations can be drawn from the model to improve yield in a sustainable way in China?

SO 3. Statistical analysis of nitrogen use efficiency in northeast China using multiple linear regression and random forest (**Chapter 4**).

- What are spatial and temporal patterns of NUE at county level?
- What prediction models for NUE are obtained using SMLR and RF?
- What are the major explanatory variables of the prediction models?
- How accurate are model predictions based on 10-fold cross-validation?
- How to derive recommendations from the models to improve NUE in a sustainable level in northeast China?
- Which model is more predictable for NUE at county level?

SO 4. Uncertainty quantification of NUE prediction in China using Monte Carlo simulation and quantile regression forests (**Chapter 5**).

- What are the uncertainty sources of NUE?
- How large is the uncertainty in measurements of NUE (through measurement uncertainty in yield, N fertilizer input and crop removal N)?
- How does uncertainty in these measurements propagate through RF?
- Which of the uncertain inputs are the main uncertainty sources of NUE?
- How to improve the NUE prediction models based on the results of the uncertainty analysis?

1.6. Study area

The study area of this thesis is China at two spatial scales, provincial and county scale. For the provincial scale analysis I included all provinces (31) excluding Hongkong, Macao and Taiwan; for the county scale analysis I selected all counties (183) in provinces Jilin, Liaoning and Heilongjiang in northeast China (Figure 1.3).

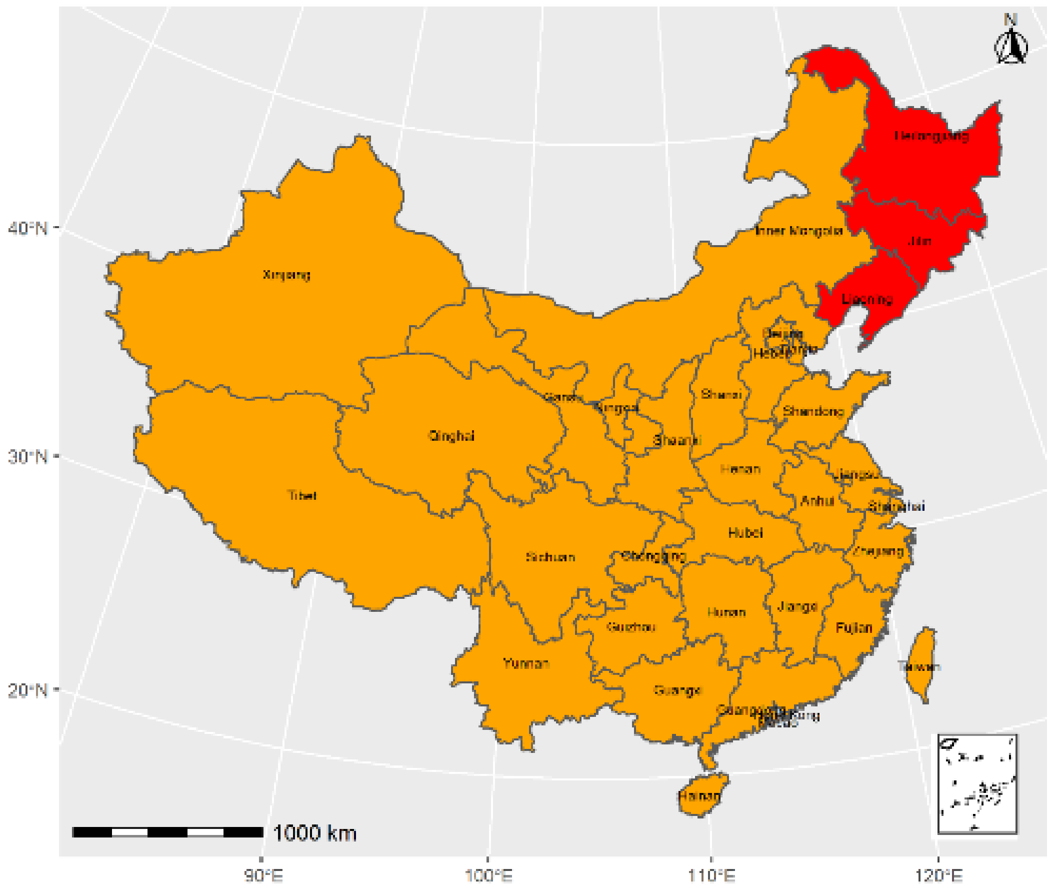


Figure 1.3: Study areas (China and northeast China, red color) of this thesis.

China has 167.5 million hectares of crop planting land, 17.5% of the total territory of China – only 8% of the global arable land (FAO, 2018), and it feeds 1.43 billion people surveyed in 2021 in China (National Bureau of Statistics of China, 2022) – 19% of the world population. The crop planting areas are distributed in all climatic zones from tropical to temperate zones spanning longitudes from 73°33′ to 135°05′E and latitudes from 3°51′ to 53°33′N. Planting systems range from one crop per year in north and west China, two crops per year or three crops within two years in northcentral and south China and three crops per year in south China. These crops grow in almost all soils and terrains. Mountains, plateaus and hills account for about 67% of the land area, and basins and plains account for about 33% of the land area (see Figure 1.4) (Feng et al., 2007). China's climate is dominated by dry seasons and wet monsoons, which lead to pronounced temperature differences between winter and summer. In winter, northern winds coming from high-latitude areas are cold and dry; in summer, southern winds from coastal areas at lower latitudes are warm and moist. The climate varies also from region to region due to highly complex topography. Figure 1.5 shows spatial patterns of annual temperature [1] and precipitation [2] in China.

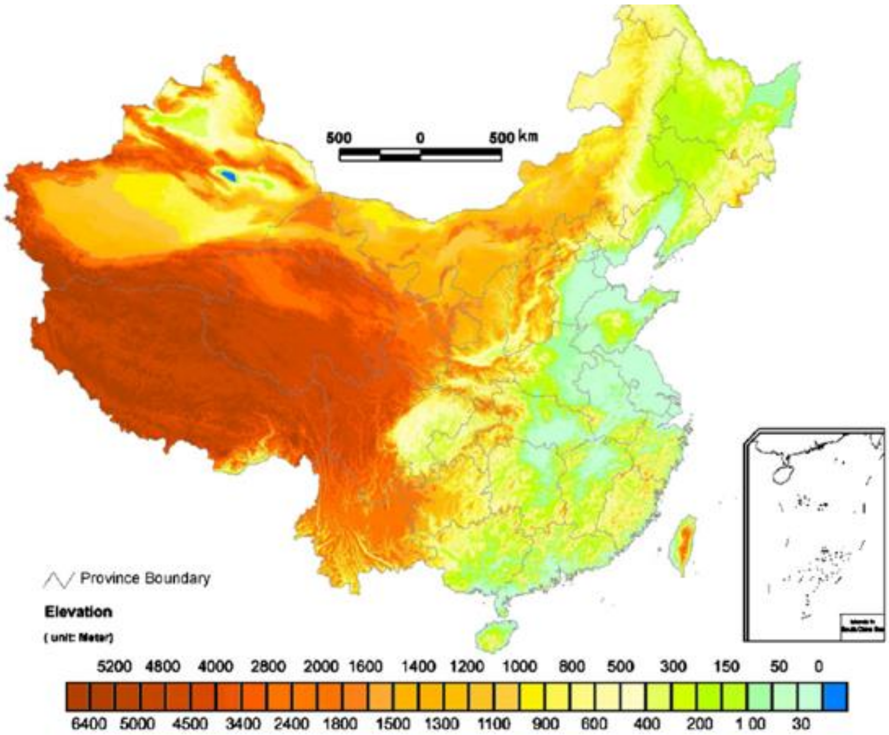


Figure 1.4: Digital elevation model (DEM) of China (Feng et al., 2007).

[1] https://www.chinamaps.org/china/china-temperature-map-annual.html.
[2] https://www.chinamaps.org/china/china-map-of-precipitation.html.

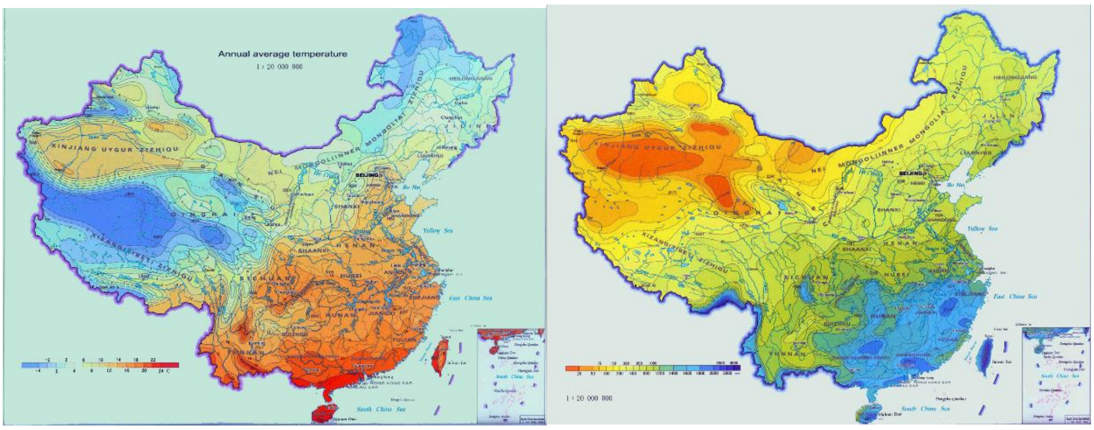


Figure 1.5: Annual temperature (a) and precipitation (b) in China [1,2].

For the county level analysis I chose Jilin, Liaoning and Heilongjiang in northeast China (Figure 1.3 marked in red) as this region is one of the three black soil areas in the world with high soil fertility (Ren et al., 2021) and the 'breadbasket' of China. The region is in the large Great Plains of China and has a high concentration of cultivated land, which is conducive to large-scale mechanized operation. The study area includes 183 counties (79 in Heilongjiang, 47 in Jilin and 57 in Liaoning) within latitude 38°46′-53°33′N and longitude 118°53′-135°05′E. The region is characterized as a temperate and cold temperate continental monsoon climate: it has long, cold winters and short, cool summers. Annual mean temperature ranges from −5 to 10°C, with mean winter temperature ranging from −28 to −2°C and mean summer temperature ranging from 15 to 25°C. Annual precipitation ranges from 400 mm in the northeast to 1000 mm in the southeast.

The population in the region is 98 million which is 7% of the total population in China. The region has 25.3 million ha of crop land, 15% of the total cultivated area of China, and grows one crop per year. Main crops are rice, maize, soybean, potato, and peanut.

1.7. Outline of thesis

This thesis comprises six chapters. **Chapter 1** contains this general introduction on the contribution of N fertilizer application to food security and environmental risk, NUE in the world and China, research objectives and questions, and a description of the study area and methods. **Chapter 2** presents an analysis of the spatial and temporal variation of NUE and its causal factors at provincial scale in China. **Chapter 3** investigates spatio-temporal variation of yields for different crop aggregations and influences of agricultural, environmental and economic factors on the space-time variation of yield at provincial level in China. **Chapter 4** further describes spatial and temporal variation of NUE and impacts of agricultural, environmental and economic factors on the variations at county scale in northeast China (Jilin, Liaoning and Heilongjiang). **Chapter 5** presents an uncertainty analysis and quantifies uncertainty sources of NUE modelling at provincial scale in China. **Chapter 6** summarizes the major conclusions of this thesis, implications, and recommendations for improvement of NUE in a sustainable range as well as an outlook for future study.

Chapter 2

Space-time statistical analysis and modelling of nitrogen use efficiency indicators at provincial scale in China

Nitrogen use efficiency (NUE) is crucial to establish efficient fertilizer application guidelines that balance crop yield, economic return and environmental sustainability. Although there are quite a few research about the spatial and temporal variation of NUE, little work has been done on modelling NUE through deriving empirical relationships with explanatory environmental variables and exploring their relative importance quantitatively. The space-time patterns of NUE indicators (i.e., the Partial Factor Productivity of nitrogen, PFP_N, and the Partial Nutrient Balance of nitrogen, PNB_N) at provincial scale in China were derived and related to environmental covariates using stepwise multiple linear regression. PFP_N was higher in east and south China than in central and west China and was smaller than 30 kg kg^{-1} yr^{-1} in most provinces, while PNB_N was moderate in most provinces (0.41-0.50 kg kg^{-1} yr^{-1}) and low (< 0.40 kg kg^{-1} yr^{-1}) in south China. The national PFP_N declined slightly from 32 kg kg^{-1} in 1978 to 27 kg kg^{-1} in 1995 and went up gradually to reach 38 kg kg^{-1} in 2015. The national PNB_N decreased from 0.53 to 0.36 kg kg^{-1} from 1978 to 2003, thereafter stabilizing at around 0.40 kg kg^{-1} yr^{-1} between 2004 and 2015. The multiple linear regression models explained 74% of the variance of PFP_N and PNB_N. The main explanatory variables of PFP_N were planting area index of sugar crop (32% of the R-square), followed by Arenosols (12%), planting area index of oil crop (8%), planting area index of vegetables (5%), silt content (5%) and total potassium (5%). For PNB_N, the variation was mainly attributed to mean annual daytime surface temperature (28% of the R-square), planting area index of crops (beans 20%, orchards 10% and vegetables 9%) and wet day frequency (5%). The results of this chapter indicate that crop types, temperature and soil properties are important variables that determine NUE. These should be considered by policy makers when agricultural land development decisions are made in order to balance NUE and productivity (i.e., agronomy and environment).

Based on:

Liu, Y., Heuvelink, G. B. M., Bai, Z., He, P., & Masiliūnas, D. (2020). Space-time statistical analysis and modelling of nitrogen use efficiency indicators at provincial scale in China. *European Journal of Agronomy*, 115, 126032.

Abbreviations: ALI, Alisols; AMP, agricultural management practice; AND, Andosols; ARE, Arenosols; BEAN, planting area index of beans; CHE, Chernozems; COA, coarse fragments volumetric; DIE, agricultural diesel engines; DTR, diurnal temperature range; EXH, exchangeable acidity; INC, per capita annual net income of rural households (farmer income); INF, inorganic fertilizer output; KAS, Kastanozems; MEL, planting area index of melons; NUE, nitrogen use efficiency; OIL, planting area index of oil crop; ORC, planting area index of orchards; PFPN, Partial Factor Productivity of nitrogen; PNBN, Partial Nutrient Balance of nitrogen; POP, rural population; REG, Regosols; ROC, absolute depth to bedrock; SILT, silt content; STA, Stagnosols; SUG, planting area index of sugar crop; TC, total carbon; THR, motorized threshing machines; TK, total potassium; TMP, mean annual surface temperature at daytime; TOB, planting area index of tobacco; TS, total sulfur; VEG, planting area index of vegetables; VIF, variance inflation factor; WET, wet day frequency.

2.1. Introduction

Cereal crop yields in China increased substantially from 1961 (1200 kg ha^{-1}) to 2017 (6000 kg ha^{-1}), mainly driven by the increasing use of chemical fertilizers, improved crop varieties and agronomic management (Mueller et al., 2012; Chen et al., 2014) (FAO, 2018). Nitrogen (N), as a major constituent of chemical fertilizer, is applied to agricultural fields to improve the growth and yield of crops (Sharma & Bali, 2018). For example, the chemical fertilizer application in agriculture was 58.6 million metric tons in mainland China in 2017, in which N fertilizer accounted for 50.8% of the total (National Bureau of Statistics of China, 2019). As the influence of N fertilizer on crop yield becomes increasingly prominent, the N consumption for agriculture in China has been increasing significantly since the 1980s. In 2016, China became the largest worldwide consumer with 27.8% of the global N consumption (FAO, 2018). However, the N application rate in China is high in comparison with other countries and regions, e.g., the annual chemical N application was 256 kg ha^{-1} in 2016, which was 3.3 times the global average N application rate, and 3.9 and 4.7 times the American and European average application rate, respectively (FAO, 2018).

The excess N application not only decreased the economic efficiency of fertilizer application (Zhu et al., 2018), but also resulted in serious environmental problems, such as waterbody eutrophication (Reidsma et al., 2012), greenhouse gas emission (Liu et al., 2015b) and soil acidification (Tian & Niu, 2015). Farmers in China need to use less N for environmental sustainability but at the same time must also make sure that the crop yields do not suffer from a lowered N application rate. Food security is of utmost importance for a large and continuously growing population in a country with limited agricultural land (China has 19% of the global population with an average annual increment of 9 million since 1990, while it has only 8% of the global arable land) (FAO, 2018). In other words, there is an urgent need to improve nitrogen use efficiency (NUE), since this would allow increasing yield and profits with minimal environmental impact (Kant et al., 2011). Two important NUE indicators are the Partial Factor Productivity of nitrogen (PFP_N) and the Partial Nutrient Balance of nitrogen (PNB_N) (defined in Section 2.2.1). The mean PFP_N was 59% for cereals in Western European countries from 1999 to 2002, 45% in North America, but only 32% in East Asia (Dobermann & Cassman, 2005). The overall observed global trend was a distinct decrease of PFP_N in the period 1961–1980 (from 68% to 45%), thereafter stabilizing to about 47% (Lassaletta et al., 2014).

China experienced a downward trend in PFP_N from 32% in 1980 to 26% in 2005 (Ma et al., 2012). The PNB_N in China varied largely among provinces in 2013 (12–33%) (Wang et al., 2018b). In recent years, the Chinese government has enacted several policies and regulations, e.g., 'Double High Agriculture', i.e., transforming from solely high yield to both high yield and high resource use efficiency, and 'Zero Growth in Chemical Fertilizer Use by 2020', aiming at improving fertilizer use efficiency and reducing pollution. However, it is difficult to evaluate the effectiveness of these policies without a clear understanding of the mechanisms behind and variation of NUE in space and time (Liu et al., 2018). In order to support decision making on allocation of fertilizer resources and define effective agricultural management practice (AMP), it is important to get more insight into how NUE indicators in China vary in space and time and how these are related to and respond to various environmental and socio-economic variables.

Although there are quite a few researches about the spatial and temporal variation of NUE, little work has been done to explore its explanatory variables and model NUE through deriving empirical relationships with explanatory variables (Ma et al., 2012; He et al., 2018). Explanatory variables of NUE, such as socio-economic variables (e.g., income), AMP (e.g., irrigated area, agricultural machinery) and environmental variables (e.g., soil, climate) are crucial for explaining the variation of NUE in space and time, developing strategies to balance crop yield, profitability and environmental sustainability, and achieving suitability-based highly-efficient agricultural management. However, most researches only concentrated on the influence of the nitrogen application rate (Xu et al., 2014a), crop variety (Barraclough et al., 2010) and soil type (Meena et al., 2016) on NUE by performing experiments for specific sites, which does not yield representative relationships between NUE and explanatory variables for the country as a whole. A few studies explored the correlation between total soil nitrogen and explanatory variables (Wang et al., 2018c) and the correlation between soil organic carbon and explanatory variables (Yang et al., 2016). (Ichami et al., 2019) applied a meta-analysis using linear regression to explore the explanatory variables of NUE in Africa. However, their model explained less than 30% of the variation of NUE and cannot easily be extrapolated to China.

The objectives of this chapter were to derive space-time patterns of two NUE indicators (PFP_N and PNB_N) at provincial scale in China for the 1978-2015 period, and construct NUE prediction models by stepwise multiple linear regression (SMLR), aiming to derive and interpret the major explanatory variables of the space-time variation in NUE.

2.2. Data and Methods

2.2.1. NUE indicators and covariates

The study covers 31 provinces of China (excluding Hong Kong, Macao and Taiwan) (Figure 2.1). Two popular NUE indicators (PFP_N and PNB_N, Table 2.1, Dobermann, 2007) were calculated using provincial data on the amount of chemical N fertilizer (single N fertilizer and compound N fertilizer), livestock numbers (cattle, horse, donkey, mule, camel, pig, sheep, poultry), rural population, crop yield and planting area of different crops (cereals, vegetables, melons and fruits) from 1978 to 2015, using data obtained from the National Bureau of Statistics of China (2019). Detailed equations are given in Appendix A and required parameters (N ratio of compound fertilizer, amount of N content in excretion and urine in humans and livestock, returning field rate of manure, cake ratio, amount of N content in different cake crops, returning field rate of cake fertilizer, ratio of straw to grain, amount of N content in different straw and returning field rate of straw) used to calculate the NUE were presented in He et al. (2018).

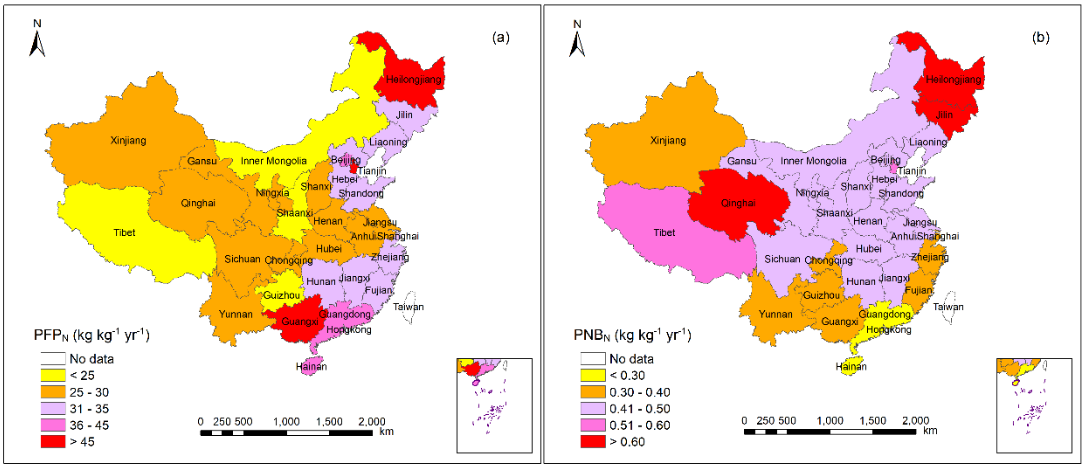


Figure 2.1: Spatial distribution of the average PFP_N (a) and PNB_N from 1978 to 2015 (b).

Table 2.1: NUE indicators and their application (Dobermann, 2007).

Term	Calculation	Unit	Question addressed	Typical use
Partial Factor Productivity	PFP_N = Yield/N_{input}	kg kg^{-1} yr^{-1}	How productive is this cropping system in comparison to its nitrogen input?	Long-term indicator of trends in NUE.
Partial Nutrient Balance	PNB_N = $N_{crop\_removal}$/N_{input}	kg kg^{-1} yr^{-1}	How much nitrogen is being taken out of the system in relation to how much is applied?	Long-term indicator of trends; most useful when combined with soil fertility information.

Note: Yield, the economic yield of crop with nitrogen applied; $N_{crop\_removal}$, total nitrogen removal in aboveground crop biomass with nitrogen applied; N_{input}, total amount of nitrogen applied to the field from different sources (chemical fertilizer, manure, cake fertilizer and straw).

For deriving provincial NUE indicators over four decades, the initial data needed to be pre-processed, mainly by filling missing data in some years, calculating indicators which were not present in the original database but that can be obtained using other available data or parameters, and original data verification. Gap filling was done by consulting statistical yearbook/published literature and using data from adjacent years if the absent data only occurred in a few years. Deriving data and parameters that were not in the original database was more difficult since we needed a large number of parameters. For example, for calculating PNB_N, the N of crop removal is required, but no information about it was available from any authorized source. In this case, we used economic yield data (dry matter of the main harvested product for cereal crop cultivation purposes; fresh weight for vegetables, melons and fruits) and parameters about the N requirement per unit of economic yield per crop to calculate the N of crop removal. The data verification step is important because it is directly related to the quality of the database and accuracy of the results. The database was too large to manually identify erroneous data. Therefore, a set of strict rules for data quality evaluation were applied to detect suspicious values. We flagged suspicious values if there was a large difference between data from adjacent years (i.e., outliers). For example, if a value was more than 10 times larger or smaller than that of an adjacent year, then this may be due to a typo and the value was flagged. Flagged data were closely examined and compared with other available data sources (e.g., total economic yield must equal the product of the economic yield per planting area and planting area of different crops). If a value was not accepted, it was replaced by the average value of adjacent years. This was the case for planting area of tubers in Tibet in

1978; phosphorus application rate in Shanghai and Beijing in 1996; total fertilizer application rate in Tianjin, Inner Mongolia, Liaoning, Jilin, Heilongjiang, Jiangsu, Hubei, Hunan, Yunnan, Guangxi, Shanxi, Gansu, Qinghai, Ningxia and Xinjiang in 1978.

Explanatory variables representing crop types, topography, soil properties (including soil types), climate, economy and AMP were used in the multiple linear regression models. Some covariates, such as related to biomass productivity (e.g., enhanced vegetation index, topography and soil properties) were considered static, while others were treated as dynamic, so that for these variables we prepared annual time series maps from 1978 to 2015. Explanatory variables were derived from SoilGrids (2018), Hengl et al. (2017), SoilGrids covariates, Climate Research Unit (2015), the National Bureau of Statistics of China (2019) and published articles (e.g., Shangguan et al., 2014), see Table A.1. All covariates were aggregated to the provincial level by taking their spatial average. Categorical variables crop type and soil type were represented as continuous-numerical variables by using proportions (planting area index of each crop in total provincial planting area, proportion of dominant soil types in each province). As SoilGrids distinguishes 118 soil types, with differences between them that are not all relevant for this chapter, we grouped the 118 soil types into 30 (i.e., the reference groups of the World Reference Base) by summing the proportions of soil types that have the same reference group. In summary, there are 1159 NUE observations (30 provinces from 1978 to 2015; Chongqing province was established in 1997) and 108 explanatory variables. The mean values of the main covariates for each province are summarized in Table A.2.

2.2.2. Stepwise multiple linear regression and model fitting

Multiple linear regression, a classical linear method, aims at explaining the variation of the dependent variable by a linear combination of explanatory variables (Ryan, 2008). In this chapter, the dependent variables were PFP_N and PNB_N at provincial scale and the explanatory variables were covariates that might influence the NUE indicators (see Section 2.2.1). However, the problem of multi-collinearity occurs when explanatory variables are significantly and strongly correlated. Stepwise regression analysis reduces multi-collinearity of explanatory variables (Khanal et al., 2018). It is an iterative process that continues to add or remove variables from the regression equation until there is no improvement (Dutta et al., 2015). SMLR is widely used in developing empirical models from large data sets (Peerlinck et al., 2018). R statistical computing software (R Core Team, 2018) was used for regression analysis, and the adjusted R^2 (coefficient of determination) and residual standard errors were recorded for each model. During the modelling process, transformation of the dependent variable was applied if the residuals of the regression model deviated substantially from the normal distribution. In this chapter, transformation of the dependent variables was not necessary since the residuals of the models were sufficiently normal (Figure A.1).

There are many criteria that can be used to select explanatory variables in stepwise regression, such as a sequence of F-tests or t-tests, the adjusted R^2, Akaike information criterion (Soergel et al.) and Bayesian information criterion (BIC) (Efroymson, 1960; Montgomery et al., 2021). AIC is often used but tends to keep a large number of explanatory variables, which makes the model more difficult to interpret. Therefore, in this chapter, we used the BIC for variable selection. It is implemented in the stepAIC function in the MASS package (Khanal et al., 2018). The stepAIC function begins with either a full or a null model, and methods for stepwise regression can be specified in the direction argument, with character values 'forward', 'backward' and 'both' (Zhang, 2016). In this chapter, we used the 'forward' approach, which starts with a null model without covariates and iteratively adds covariates based on the BIC criterion. As the choice of the explanatory variables during stepwise regression is carried out automatically, there is a risk of including spurious explanatory variables in the final model, leading to interpretation difficulties (Zuur et al., 2009). The remedy is to check the explanatory variables in the final model and substitute the variables not easily explained with other highly correlated variables, which are easier to interpret. However, in our study, variable substitution appeared to be not necessary.

Pearson's correlation was calculated using the chart.Correlation function in R to explore the direct linear relationships between NUE indicators and explanatory variables (de Carvalho Junior et al., 2014). The variance inflation factor (VIF) of all covariates in the model were monitored in the modelling process to check for multi-collinearity (multi-collinearity is considered a serious problem if the VIF value of any covariate is larger than 10) (Wang et al., 2018c). It should be noted that the model is usually only reliable within the range of covariate values exhibited by the training data. Thus, care was taken that 'extrapolation in feature space' did not occur, by excluding covariates that induce a strong extrapolation. The relative importance of explanatory variables of the SMLR models was obtained by the method of relative weights (Kabacoff, 2011). This method closely approximates the average increase in R-square obtained by adding an explanatory variable across all possible sub-models (LeBreton & Tonidandel, 2008).

2.2.3. Validation/Accuracy assessment

The performance of the SMLR model was evaluated with 10-fold cross-validation by comparing the predicted and observed NUE indicator values. Model validation metrics were the mean prediction error (ME), the root mean squared prediction error (RMSE), the normalized RMSE (nRMSE), R^2 and Lin's concordance correlation coefficient (LCCC) (Lin, 1989). These indices were computed as follows:

$$\mathrm{ME} = \frac{1}{n}\sum_{i=1}^{n}(P_i - O_i) \tag{2.1}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_i - O_i)^2} \tag{2.2}$$

$$\mathrm{nRMSE} = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_i - O_i)^2}}{\bar{O}} \tag{2.3}$$

$$R^2 = \frac{\sum_{i=1}^{n}(P_i - \bar{O})^2}{\sum_{i=1}^{n}(O_i - \bar{O})^2} \tag{2.4}$$

$$\mathrm{LCCC} = \frac{2r\sigma_o\sigma_p}{\sigma_o{}^2 + \sigma_p{}^2 + (\bar{O} - \bar{P})^2} \tag{2.5}$$

where n is the number of observations; P_i and O_i are the predicted and observed NUE values for observation i, respectively; $\bar{O}$ and $\bar{P}$ are the mean of predicted and observed values; σ_p and σ_o are the standard deviations of predicted and observed values; and r is the Pearson correlation coefficient between the predicted and observed values.

ME represents the prediction bias. RMSE reflects how far off the predicted values on average are from the observed values. R^2 and LCCC signify the degree to which the predicted and observed values are close to the 1:1 line (Yang et al., 2016). Predictions become increasingly better as ME is close to zero, RMSE and nRMSE are small, and R^2 and LCCC close to one.

2.3. Results

2.3.1. Spatial and temporal variations of NUE indicators

There is considerable spatial variation in the long-term average PFP_N (Figure 2.1a). It is higher in east and south China than in central and west China; 16 provinces have PFP_N less than 30 kg kg^{-1} yr^{-1}); three provinces (Guangxi, Tianjin and Heilongjiang in east China) have PFP_N higher than 45 kg kg^{-1} yr^{-1}, which is nearly twice as large as that in Tibet in west China, and Inner Mongolia, Shaanxi and Guizhou in central China (< 25 kg kg^{-1} yr^{-1}). There are also different temporal trends between provinces (Figure 2.2). In 1978, Tianjin and Heilongjiang had similar PFP_N (about 55 kg kg^{-1}), and both provinces showed a downward trend from

1978 to 2007, which increased again in Tianjin from 23 kg kg^{-1} in 2007 to 34 kg kg^{-1} in 2015, while in Heilongjiang it stayed stable (about 40 kg kg^{-1} yr^{-1}). PFP_N in Guangxi was fairly stable with 40 kg kg^{-1} yr^{-1} from 1978 to 1989, and then increased, becoming higher than that in Heilongjiang in 1991 and Tianjin in 1995. By 2015 the PFP_N in Guangxi had increased almost twofold. Although PFP_N in Tibet, Inner Mongolia and Guizhou fluctuated slightly over time, from 2006 onward it showed an increasing trend from 23, 20 and 24 kg kg^{-1} to 28, 26 and 33 kg kg^{-1}, respectively. PFP_N in Shaanxi decreased to a minimum value of 16 kg kg^{-1} in 1995 from 26 kg kg^{-1} in 1978, while after 1995 it experienced a dramatic increase to 33 kg kg^{-1} in 2015.

The temporal variations of PFP_N are shown in Figure 2.2a. The national PFP_N was 31 kg kg^{-1} yr^{-1} from 1978 to 2015, which declined slightly from 32 kg kg^{-1} in 1978 to 27 kg kg^{-1} in 1995, and then went up gradually to 38 kg kg^{-1} in 2015. There was a dramatic reduction of PFP_N (< 5 kg kg^{-1} yr^{-1}) in six provinces (Tianjin, Jilin, Heilongjiang, Anhui, Beijing and Fujian) in east China between 1978 and 2015. PFP_N in Jiangxi, Henan, Sichuan, Hunan, Hubei, Guangdong and Tibet remained stable varying from -5 to 5 kg kg^{-1} yr^{-1} for the 38-year time period, while PFP_N in the remaining provinces experienced an increasing trend (> 5 kg kg^{-1} yr^{-1}) over the entire time period.

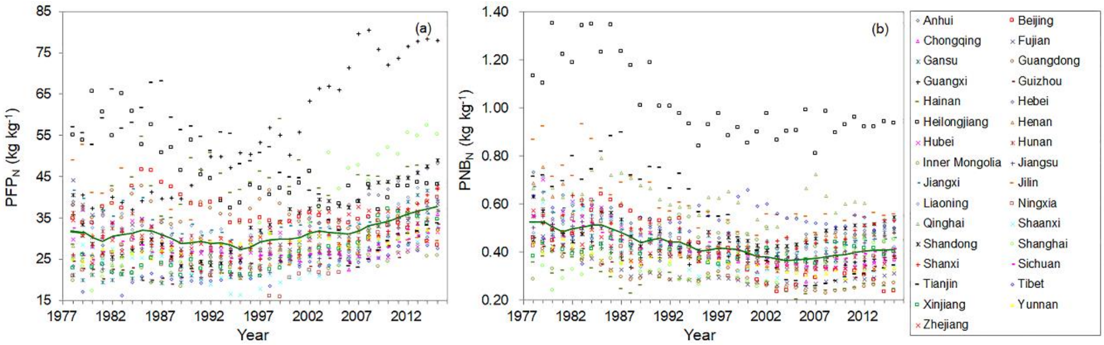


Figure 2.2: Temporal distribution of PFP_N (a) and PNB_N (b). The green line represents the annual national result computed over all provinces.

The spatial distribution of PNB_N was quite different from that of PFP_N. PNB_N was low in south China (< 0.40 kg kg^{-1} yr^{-1}) and moderate in the 15 provinces of central China (0.41-0.50 kg kg^{-1} yr^{-1}) (Figure 2.1b). PNB_N in Heilongjiang was the highest (1.03 kg kg^{-1} yr^{-1}) among all 31 provinces, followed by Jilin and Qinghai (0.60 kg kg^{-1} yr^{-1}), which doubled that in Hainan and Guangdong (0.30 kg kg^{-1} yr^{-1}), the lowest PNB_N. Interestingly, there was a similar temporal trend of PNB_N in Heilongjiang and Jilin (Figure 2.2b): a decline from 1978 to 1990 (from 1.13 to 0.84 kg kg^{-1} in Heilongjiang; from 0.87 to 0.50 kg kg^{-1} in Jilin), after which it became relatively stable (around 0.90 kg kg^{-1} yr^{-1} in Heilongjiang and 0.50 kg kg^{-1} yr^{-1} in Jilin) until 2015. From 1978 to 2015, PNB_N fluctuated in Qinghai (between 0.50 and 0.80 kg kg^{-1} yr^{-1}), while it gradually declined in Hainan (from 0.36 to 0.25 kg kg^{-1} yr^{-1}) and in Guangdong (from 0.45 to 0.27 kg kg^{-1} yr^{-1}).

The national PNB_N was around 0.43 kg kg^{-1} yr^{-1} from 1978 to 2015. It decreased from 0.53 kg kg^{-1} in 1978 to 0.36 kg kg^{-1} in 2003, thereafter stabilizing at around 0.40 kg kg^{-1} yr^{-1} over twelve years (Figure 2.2b). The PNB_N in most of the provinces (23) declined with time (PNB_N decreased more than 0.05 kg kg^{-1} yr^{-1} from 1978 to 2015), while in Jiangsu, Ningxia, Chongqing, Xinjiang and Gansu, it maintained steady (ranging between -0.05 and 0.05 kg kg^{-1} yr^{-1}). The PNB_N in Shanghai, Shanxi and Inner Mongolia showed an improvement (> 0.05 kg kg^{-1} yr^{-1}) for the 1978-2015 period.

2.3.2. Prediction model and performance evaluation

Table A.3 shows summary statistics of the NUE indicators and explanatory variables. PFP_N and PNB_N had positive skewness, with skewness coefficients of 1.75 and 2.27, respectively. Standard deviations (SD) of PFP_N and PNB_N were 10 and 0.15 kg kg^{-1} yr^{-1}, respectively, substantially smaller than the mean value of 32 kg kg^{-1} yr^{-1} (ranging from 16 to 81 kg kg^{-1} yr^{-1}) and 0.45 kg kg^{-1} yr^{-1} (ranging from 0.20 to 1.35 kg kg^{-1} yr^{-1}). To quantify the pairwise relationships between the dependent variables and explanatory variables, Pearson's correlation coefficients were calculated (Figure 2.3). PFP_N had a positive correlation with planting area index of sugar crop, planting area index of vegetables, planting area index of melons, farmer income, agricultural diesel engines, total carbon, and total sulphur. It was negative between PFP_N and planting area index of oil crop, Arenosols, total potassium, diurnal temperature range, coarse fragments volumetric, planting area index of tobacco and silt content. There was no significant correlation between PFP_N and Andosols, inorganic fertilizer output, rural population and Chernozems (Figure 2.3a). For PNB_N, most of the explanatory variables were significantly negatively correlated with PNB_N. There was a significant positive correlation between PNB_N and planting area index of beans, absolute depth to bedrock, Stagnosols, total carbon and Kastanozems, while no significant correlation was found with Arenosols, planting area index of oil crop, and coarse fragments volumetric (Figure 2.3b).

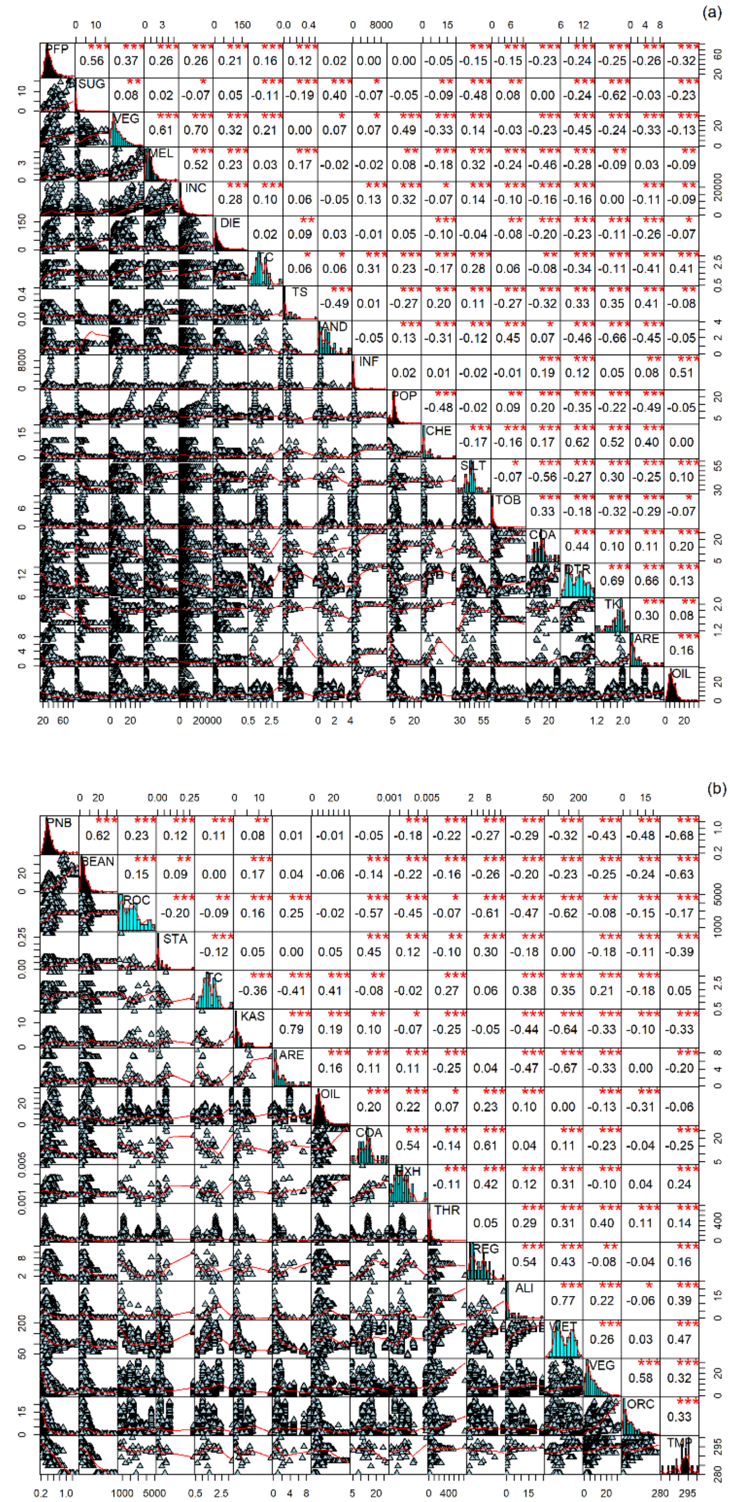


Figure 2.3: Pairwise comparison of all explanatory variables used in the PFP_N (a) and PNB_N regression models (b). The upper-right panel contains estimated pair-wise Pearson correlations. The diagonal panel shows histograms and the lower-left panel shows scatter plots with a LOESS smoother added to aid visual interpretation.

Note: * P < 0.05; ** P < 0.01; *** P < 0.001. See Abbreviations for explanation of the variable code names.

We established two SMLR models to predict PFP_N and PNB_N. The estimated regression coefficients of the models are shown in Table 2.2. All explanatory variables significantly influenced the NUE indicators. Half of the explanatory variables showed positive influence in the PFP_N model, while one fourth of the explanatory variables showed positive influence in the PNB_N model. The VIF of the explanatory variables were all smaller than 10, indicating that there was no multi-collinearity problem (Table 2.3).

Table 2.2: Regression coefficients of the PFP_N and PNB_N models.

Classification	Property	PFP_N model		PNB_N model	
		Regression coefficients	Std. error	Regression coefficients	Std. error
	(Intercept)	5.18E+01	3.06E+00	8.19E+00	3.91E-01
Crop	BEAN			2.88E-03	6.85E-04
	MEL	2.26E+00	2.56E-01		
	OIL	-3.10E-01	4.64E-02	-3.38E-03	6.54E-04
	ORC			-6.25E-03	7.31E-04
	SUG	2.53E+00	9.92E-02		
	TOB	-1.28E+00	1.46E-01		
	VEG	1.20E-01	4.64E-02	-5.57E-03	5.83E-04
Soil	COA	1.47E-01	5.72E-02	-1.26E-02	1.00E-03
	SILT	-1.48E-01	4.64E-02		
	ROC			-2.47E-05	3.56E-06
	ALI			-5.17E-03	8.79E-04
	AND	-3.67E+00	2.46E-01		
	ARE	-3.48E+00	1.50E-01	-1.21E-02	2.37E-03
	CHE	-2.10E-01	5.42E-02		
	KAS			-1.08E-02	1.49E-03
	REG			9.96E-03	1.86E-03
	STA			-3.82E-01	6.15E-02
	TC	1.90E+00	4.03E-01	4.37E-02	5.60E-03
	TK	-9.18E+00	2.00E+00		
	TS	1.69E+01	2.20E+00		
	EXH			4.00E+01	4.55E+00
Climate	WET			-1.04E-03	1.51E-04
	TMP			-2.53E-02	1.32E-03
	DTR	1.34E+00	1.85E-01		
Economic variables	INC	5.33E-04	6.38E-05		
AMP	POP	-1.71E+00	1.08E-01		
	INF	1.11E-03	2.84E-04		
	DIE	-3.19E-02	5.23E-03		
	THR			-1.43E-04	4.37E-05

Note: See Abbreviations for explanation of the variable code names. All variables were significant at the $p=0.01$ level.

Table 2.3: Variance inflation factors (VIF) for the multiple linear stepwise regression models.

Classification	Property	PFP_N model	PNB_N model
Crop	BEAN	-	2.6
	MEL	2.6	-
	OIL	2.4	1.8
	ORC	-	2.4
	SUG	2.5	-
	TOB	1.9	-
	VEG	3.9	2.4
Soil	COA	4.4	5.2
	SILT	4.4	-
	ROC	-	3.8
	ALI	-	4.7
	AND	3.1	
	ARE	5.5	5.3
	CHE	2.3	-
	KAS	-	3.8
	REG	-	4.7
	STA	-	2.0
	TC	2.6	1.9
	TK	9.6	-
	TS	2.6	-
	EXH	-	2.5
Climate	WET	-	7.9
	TMP	-	5.1
	DTR	8.6	-
Economic variables	INC	2.5	-
AMP	POP	2.8	-
	INF	1.5	-
	DIE	1.6	-
	THR	-	1.4

Note: See Abbreviations for explanation of the variable code names.

The predictive performance of the SMLR models was evaluated using 10-fold cross-validation, and from the results the ME, RMSE, nRMSE, R^2 and LCCC were computed (Table 2.4). The cross-validation statistics are only slightly worse than the model performance statistics, indicating that overfitting did not occur (Table 2.4). The models were highly predictive and explained 74% of the variance of PFP_N and PNB_N. The nRMSE of the two models were 0.15 and 0.17, respectively, indicating that the model performance was good. Other accuracy verification metrics also showed that the NUE models exhibited good performance with low ME and RMSE, and high LCCC (> 0.8). The relationships between observed and predicted data are shown in scatter plots in Figures 2.4a and b, while density plots are shown in Figures 2.4c and d. Both models tend to smooth the data and underrepresent extremes, although this effect is stronger for PFP_N than for PNB_N.

Table 2.4: Model performance and cross-validation metrics of stepwise multiple linear regression models.

Metrics	PFP_N	PFP_N CV	PNB_N	PNB_N CV
ME	0.00	-0.01	0.000	0.000
RMSE	4.86	5.02	0.079	0.080
nRMSE	0.15	0.16	0.17	0.18
R^2	0.74	0.73	0.74	0.73
LCCC	0.85	0.84	0.85	0.85

Note: CV means 10-fold cross-validated results.

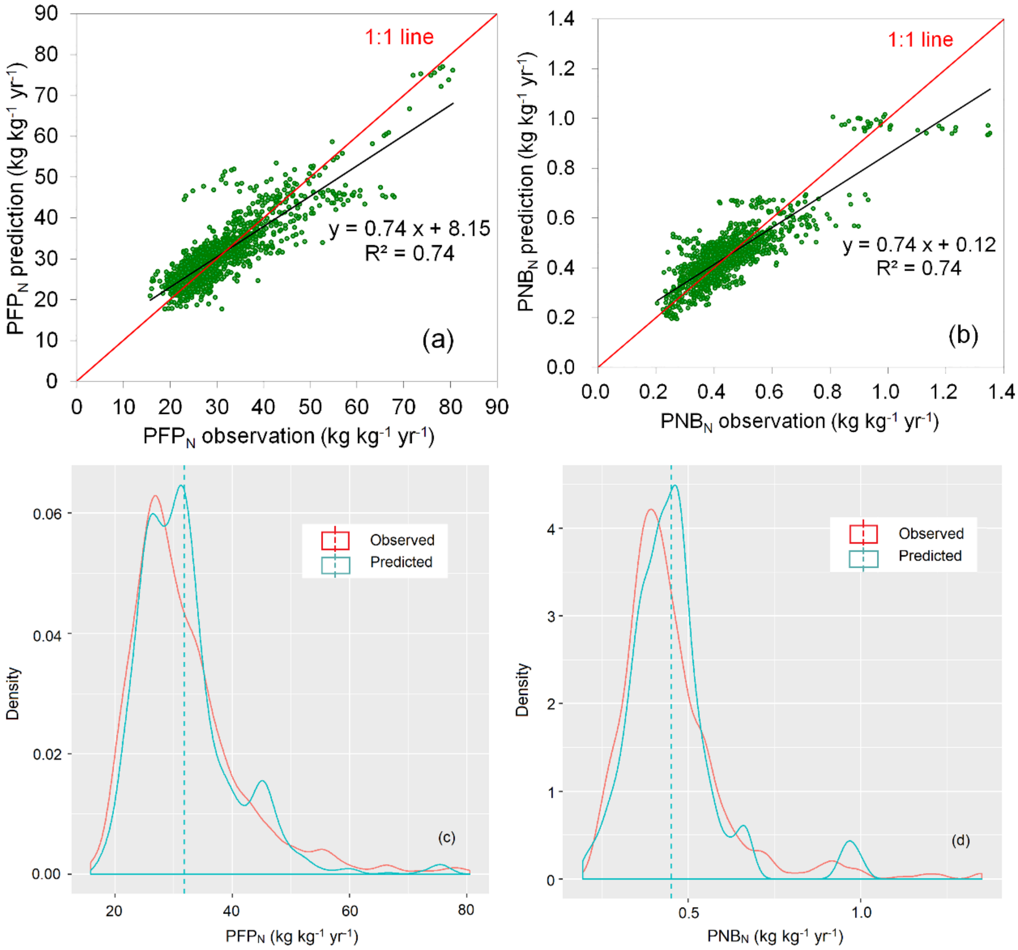


Figure 2.4: Scatter plots and Kernel density plots of observed and predicted PFP_N (a, c) and PNB_N (b, d).

2.3.3. Relative importance of explanatory variables

The relative importance of the explanatory variables for both NUE models are shown in detail in Figures 2.5a and b. The variance of PFP_N is primarily explained by planting area index of sugar crop (32% of the R-square), followed by Arenosols (12%), planting area index of oil crop (8%), planting area index of vegetables (5%), silt content (5%) and total potassium (5%). The variance of PNB_N is mainly attributed to mean annual surface temperature at daytime (28% of the R-square), planting area index of crops (beans 20%, orchards 10% and vegetables 9%) and wet day frequency (5%). The relative importance of other explanatory variables was below 5%. Overall, the crop types and soil properties accounted for 52% and 36% of the R-square, respectively; the AMP, economic factors and climate factors accounted for 6%, 4% and 3% of the R-square in the PFP_N model, respectively. For PNB_N, the most influential factors were crop types (45% of the R-square), climate (34%) and soil properties (19%). In contrast, the AMP only accounted for 2% of the R-square.

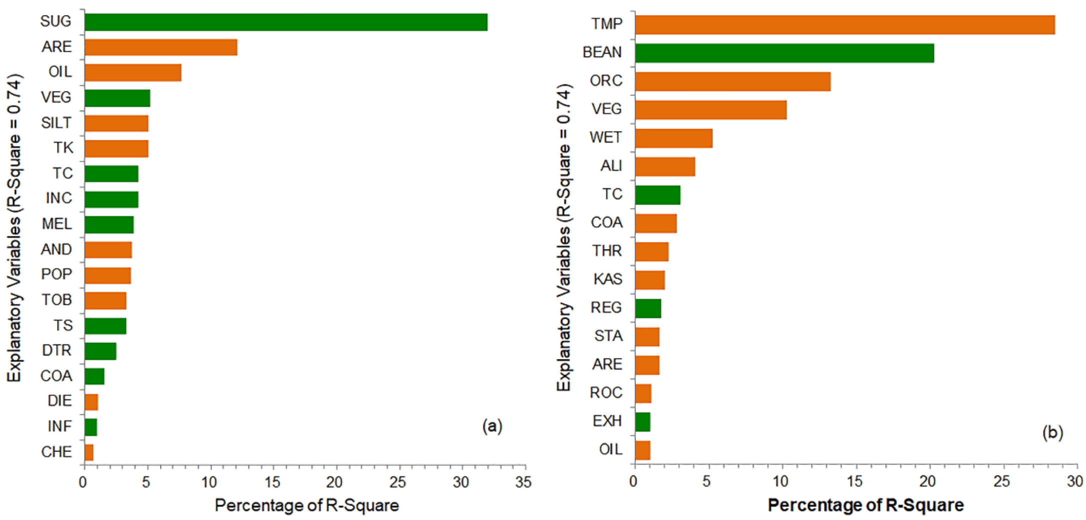


Figure 2.5: Relative importance of explanatory variables for the PFP_N model (a) and PNB_N model (b). Green shading means positive influence and orange shading means negative influence.

Note: See Abbreviations for explanation of the variable code names.

2.4. Discussion

2.4.1. Spatial and temporal variability of NUE indicators

The results showed large spatial variations in PFP_N among provinces (23-55 kg kg^{-1} yr^{-1}), which were caused by the great spatial variation in both yield and N application rate. Therefore, the policy of 'Zero Growth in Chemical Fertilizer Use by 2020' proposed by the Ministry of Agricultural and Rural Affairs should be considered specifically by provinces or regions. For example, the N rate may be further reduced in Jiangsu, where the N input was more than 250 kg kg^{-1} yr^{-1}, while the yield achieved was only moderate (Figure A.2.2b). PFP_N was higher in east and south China than in central and west China. Other researches obtained similar results, with PFP_N decreasing from 52 kg kg^{-1} yr^{-1} in northeast China (Qiu et al., 2015), 47 kg kg^{-1} yr^{-1} in south China (Pan et al., 2017), 23 kg kg^{-1} yr^{-1} in north-central China (Meng et al., 2012) to less than 20 kg kg^{-1} yr^{-1} in west China (Zhang, 2012; Wang et al., 2016a). Li et al. (2013) also obtained similar PFP_N results for grain crops, which was largest in north-east China (60 kg kg^{-1}), followed by south-west China (33 kg kg^{-1}), central China (29 kg kg^{-1}) and west China (27 kg kg^{-1}) in 2008. The only difference was that PFP_N in south-east China in their study was the lowest (20 kg kg^{-1}). The reason for the difference is that Li et al. (2013) only considered grain crops, while planting area of grain crops accounts for less than 40% of total crops in south-east China.

Analysis of the causes of the spatial differences could lead to opportunities for improvement of PFP_N in some provinces. PFP_N in Guangxi, Tianjin and Heilongjiang was at the highest level (> 45 kg kg^{-1} yr^{-1}), almost twice as large as that in Tibet, Inner Mongolia, Shaanxi and Guizhou (< 25 kg kg^{-1} yr^{-1}). This was caused by the higher yield level than N input in Guangxi, Tianjin and Heilongjiang; whereas Inner Mongolia and Guizhou had lower yield level than N input, and both yield and N input levels were low in Tibet and Shaanxi (Figures A.2a and b). Taking the influence of explanatory variables into account showed that there was a larger planting area of sugar crop (sugarcane) in Guangxi, which contributed largely to improving PFP_N (see Section 2.3.3). In Tianjin, there was more planting area of vegetables, total carbon and farmer income and less planting area of oil crop. In Heilongjiang, there was more planting area of sugar crop (sugar beet) and less planting area of oil crop, whereas in Tibet, Inner Mongolia, Shaanxi and Guizhou, there was less planting area of sugar crop, planting area of vegetables, farmer income, and more planting area of oil crop (Table A.2).

Although the yield almost doubled from 1978 to 1995, the PFP_N decreased from 32 kg kg^{-1} to 27 kg kg^{-1} during this period, because the N inputs increased even more. For the 1995-2015 period the PFP_N increased gradually in the first decade because the government introduced a balanced fertilization policy in 1995. From 2005 onward the PFP_N increased much faster due to the promotion and application of soil testing and fertilizer recommendations in 2005. Another plausible reason for the PFP_N increase between 1995 and 2015 is the improvement of AMP and crop varieties. For example, the PFP_N of irrigated rice from 2000 to 2013 with optimal treatment was 47 kg kg^{-1} yr^{-1} (Xu et al., 2016b). Qian et al. (2016) used 21 maize hybrids within four decades from the 1970s to 2000s, which were the most representative hybrids of the time, to do experiments in the main maize-growing area in north-east China. Their results indicated that PFP_N was increased by changing maize hybrids over time. The PFP_N for global cereals decreased from 245 kg kg^{-1} yr^{-1} in 1961-1965 to 52 kg kg^{-1} yr^{-1} in 1981-1985 and was 37 kg kg^{-1} in 2002 (Dobermann & Cassman, 2005). However, the PFP_N was lower than 30 kg kg^{-1} yr^{-1} in most of the provinces in this chapter. It was lower than the global level during the same period, which proves the need for the 'Double High Agriculture' policy in China. In addition, the national PFP_N was 31 kg kg^{-1} yr^{-1} from 1978 to 2015, which was much lower than that of well-managed systems (> 60 kg kg^{-1} yr^{-1}) (Dobermann, 2007).

The spatial distribution of PNB_N showed substantial spatial variation: PNB_N was low in south China (< 0.40 kg kg^{-1} yr^{-1}) and moderate in most provinces of central China (0.41-0.50 kg kg^{-1} yr^{-1}). This may be explained by the annual N amount of crop removal and input: south China had lower N levels of crop removal than input, while central China had similar N levels of crop removal and input (Figures A.2b and c). Liu et al. (2011) indicated that PNB_N was 1.70 kg kg^{-1} yr^{-1} in north-west China, larger than 1.10 kg kg^{-1} yr^{-1} in north-central China and 0.81 kg kg^{-1} yr^{-1} in central China and south of the Yangtze River. PNB_N in south China was the lowest (0.79 kg kg^{-1} yr^{-1}) (Wu et al., 2015), while it was highest in Heilongjiang in north-east China (1.03 kg kg^{-1} yr^{-1}), followed by Jilin in north-east China and Qinghai in west China (0.60 kg kg^{-1} yr^{-1}). This was twice the lowest PNB_N in Hainan and Guangdong in south China (0.30 kg kg^{-1} yr^{-1}). The explanation for these differences was that there were higher N levels of crop removal than input in Heilongjiang, Jilin and Qinghai, and lower N levels of crop removal than input in Hainan and Guangdong (Figures A.2b and c). The differences can also be interpreted by explanatory variables: lower surface temperature at daytime, more planting area of beans and less planting area of orchards and vegetables have large contributions to the higher PNB_N of Heilongjiang, Jilin and Qinghai. In contrast, Hainan and Guangdong have higher surface temperature at daytime, less planting area of beans and more planting area of orchards and vegetables (Table A.2).

Most published researches only focused on cereals and rarely included vegetables, melons and fruits, which may not reflect the agricultural PFP_N and PNB_N of the whole region (Dobermann & Cassman, 2005). In this chapter, we considered all crops in agricultural land and major fertilizer input resources (chemical N, manure and cake N, and straw returning N). As reported in (He et al., 2018), the PNB_N in China was 0.39, 0.36, 0.34 and 0.37 kg kg^{-1} yr^{-1} in the 1980s, 1990s, 2000s and 2010s, respectively. These values are much lower than the results presented in this chapter, i.e., 0.48, 0.43, 0.38 and 0.41 kg kg^{-1} yr^{-1}, respectively. This was because the former also considered other N input sources (fixation, symbiotic and non-symbiotic, dry and wet deposition from the atmosphere, irrigation water and crop seed). The differences suggest that other N input sources play a less important role over time. According to this chapter, the long-term national PNB_N was 0.43 kg kg^{-1} yr^{-1}, which reflected that the applied N was much more than the N of crop removal (PNB_N < 1). Applied N that is not taken up by the crop or immobilized in soil organic N pools, including both microbial biomass and soil organic matter, has a risk to be lost by volatilization, denitrification and leaching. Therefore, increasing uptake efficiency from applied N inputs is an effective method to achieve higher NUE, by reducing the amount of N loss from soil organic and inorganic N pools.

2.4.2. NUE models and main explanatory variables

The results of this chapter show that the available data on topography, crop type, soil property, climate, economic variables and AMP explain 81% of the variation of PFP_N. To simplify the models and make them easier to interpret, we selected 18 explanatory variables in the final model, which explained 74% of the variation of PFP_N, higher than that of Ichami et al. (2019), which could only explain 33% of the variation of NUE. For PNB_N, the full model explained 84% of the variation, while the final model with 16 explanatory variables still explained 74% of the variation. Thus, both models had the same R^2 with almost the same number of explanatory variables. The ME, RMSE, R^2 and LCCC metrics also indicated that the PNB_N model had the same performance as the PFP_N model.

PFP_N is a broad indicator of NUE, incorporating the contributions from indigenous soil N, N uptake efficiency, and the efficiency of converting N acquired by the plant into grain yield (Cassman et al., 2002). PNB_N is a simpler measure of N recovery efficiency, which is useful when combined with soil fertility information. Thus, differences of the NUE indicators depend on crop types, their attainable yield potential, soil quality, climate, amount and form of N application, and the overall timing and quality of other crop management operations (e.g., crop variety, planting density, tillage, seeding rate, sowing time, weeds, pest management, irrigation infrastructure) (Cassman et al., 2002; Zhang et al., 2016). However, these factors have different relative importance under different conditions. The main factors affecting NUE (agronomic efficiency of N) for Kenya were P-Olsen, silt content, soil pH, clay and rainfall, whereas only soil pH, exchangeable K and texture were important for Sub-Saharan Africa (Ichami et al., 2019). However, in China, crop type and soil property contributed most to the variance explained of both NUE indicators. A possible reason might be that fewer explanatory variables were used in the Kenya and Sub-Sahara Africa studies (Ichami et al., 2019), and that the NUE indicators were also different (agronomic efficiency of N, PFP_N, PNB_N). In this chapter, temperature was also an important factor for PNB_N, but not PFP_N. PFP_N is related to economic yield produced by grain or fruits, while PNB_N considers N uptake by the whole plant, including grain/fruits and straw/leaves (brunches). In addition, yield was largely affected by accumulated temperature, while N uptake benefited from temperature (Kaur et al., 2012; Fan et al., 2014). Therefore, for agricultural resource use efficiency balancing NUE, productivity, and environment, suitable crop types, temperature and soil properties should be considered by policy makers when taking decisions on developing agricultural land.

In this chapter, planting area index of sugar crop was the most important explanatory variable for PFP_N. This was likely because the economic yield per unit of sugar crop was the largest among different crops, since the yield of sugar crop was considered by fresh weight, including large water content. All vegetables, melons and fruits were considered fresh weight. This explains that planting area index of sugar crop, vegetables and melons all had positive influences on PFP_N. In view of exploring the approach for improving the NUE, it is advised that the biomass of crops is used instead of yield. However, the database that we used did not list all different kinds of vegetables, melons and fruits, which means that this correction could not be made. On the contrary, planting area index of oil crop had a negative effect on PFP_N, since most of the oil crops had lower yield in agriculture.

Interestingly, Arenosols was also an important negative factor for PFP_N at provincial level, based on the regression coefficients and relative importance analysis. This is because Arenosols have coarse soil texture and very low cation exchange capacity. Cation exchange capacity is an important indicator of soil fertility, because it indicates an abundance of essential nutrients which can be taken up by plants (Sharma et al., 2015). In addition, Arenosols have high permeability and a low water and nutrient storage capacity, which cause lower yield with the same N inputs, i.e., low PF_PN. Silt content and total potassium had a negative influence on PF_PN. Ichami et al. (2019) also noted that less silt led to higher agronomic nitrogen use efficiency. This might be because soils with more than 40% silt tend to have low permeability and intake rates (if the compaction of soil or bulk density are the same), which reduce the water and nutrient use efficiency (Diebold, 1954; Ishaq et al., 2001). In fact, moderately compact and compact soils containing less than 40% silt had three times the median percolation rate than soils with 40% or more silt (Diebold, 1954). The mean silt content in this chapter was 41% (Table A.3), which means that more than half of the soils had silt content over 40%. Another explanation for the lower nutrient use efficiency might be that higher silt content increases gross nitrification rates (i.e., nitrogen loss), since silt content is negatively correlated with the C/N ratio (Li et al., 2020c), while the gross nitrification rate is negatively related to the C/N ratio (Bengtsson et al., 2003). For total potassium, it is hard to explain why this had a negative influence on PF_PN (Figure 2.8a). Moreover, total carbon and total sulphur were positively associated with PF_PN, which is consistent with the result of Ichami et al. (2019). Many studies show that climate has a great influence on yield (e.g., Challinor et al., 2014). In this chapter, diurnal temperature range was revealed as an important positive factor for PFP_N at provincial scale. This might be because a higher diurnal temperature range could increase seed weight during the grain-fill stage (Fang et al., 2017). Although the diurnal temperature range had a negative impact on yield (Lobell, 2007a; Tao et al., 2008) and this chapter also showed a negative impact of the diurnal temperature range on PFP_N, the interaction of other factors in the model might also have caused this result. The data used in this chapter was the annual temperature, which differs from the mean growing season temperature, which might not sufficiently reflect the real relation between PFP_N and diurnal temperature range. However, it was difficult to consider the mean growing season temperature because many crops were included, which have different growing seasons.

The most important explanatory variable for PNB_N was mean annual surface temperature at daytime, which had a negative influence on PNB_N. Temperature and atmospheric CO_2 concentration affect soil microbial physiology, nutrient cycles and availability of nutrients for crop growth, but the net effect (positive or negative) is different in different crops and regions (Pilbeam, 2015). In this chapter, the negative influence of temperature on PNB_N in China might be because a higher temperature causes higher emission and leaching of nitrogen, especially in south China (Sänger et al., 2011). Similarly, wet day frequency is a negative factor for PNB_N in our study (the mean value was 123 days per year). It might be negative because more slight precipitation promotes ammonia volatilization and nitrogen leaching (Wang et al., 2019; Liu et al., 2020). However, these results are in conflict with those of (Peng et al., 2015). Therefore, further study is needed. The planting area index of beans positively influenced the PNB_N, according to the regression model. This might be caused by the fact that beans have higher nitrogen requirement per unit economic yield, which can promote nitrogen removal. On the contrary, planting area index of orchards and vegetables had negative impact on PNB_N, since these had lower nitrogen requirement per unit economic yield. Total carbon also had a positive influence on PNB_N, since this was positively correlated with important crop nutrients. However, improving soil quality and its indigenous N supply is a slow process and depends on other factors as well, such as reduced tillage and the return of crop residues (Ding et al., 2018). Instead, achieving higher NUE at provincial or national level will require policies that favor increases in NUE at the field scale, with emphasis on technologies that can achieve greater congruence between crop N demand and N supply from all sources, including chemical fertilizer, organic inputs, and indigenous soil N (Cassman et al., 2002).

2.4.3. Weakness and limitations

Some explanatory variables used in this chapter were dynamic and recorded as annual time series. These variables often display temporal autocorrelation. The data analysis and modelling might therefore benefit from a longitudinal analysis approach (Cook et al., 2002), which takes temporal correlation into account. Results might also improve using a geostatistical approach (e.g., regression kriging), in which spatial correlation is considered (Webster & Oliver, 2007). The limitations of the data (e.g., we had no fertilization rate for each crop and ignored temporal variation in soil and other properties) may also affect the accuracy of the results presented in this chapter. Uncertainties were also introduced by gap filling of missing values in certain years, notably in N, phosphorus and potassium ratio in compound fertilizers, manure and straw returning rate, and straw/grain ratio changing among varieties of crops.

It should be noted that this chapter did a statistical analysis on the provincial scale, which was much too coarse for deriving detailed recommendations for farmers. In order to derive more spatially refined fertilizer recommendations, a higher spatial resolution should be used, such as on county or site level. This could be done if NUE indicator observations are available at finer spatial scales. Note that most of the covariates used in this chapter were available at high spatial resolution and were aggregated to provincial level for the purposes of this chapter.

Finally, although the models could explain a large part of the spatial and temporal variation, they may be improved by expanding the covariate set with additional relevant variables. While SMLR is a useful method to quantify the effect of explanatory variables on dependent variables and allows interpretation because of the simplicity of the resulting models, it is important to note that it only reflects linear and additive relationships. These relationships are much more complex, which calls for the use of non-linear statistical models. Machine learning approaches (e.g., random forests) are a potential option to address this problem (James et al., 2013), and we will investigate this in future research.

2.5. Conclusions

This chapter developed and calibrated multiple linear regression models that predict nitrogen use efficiency indicators at provincial scale in China from explanatory variables (crop type, climate, topography, soil type and properties, economic variables and AMP). It also included an analysis of the spatial and temporal variations of NUE indicators at the provincial scale in China from 1978 to 2015. The results showed substantial temporal and spatial variation of PFP_N and PNB_N. PFP_N was larger in east and south China than in central and west China. It was also smaller than 30 kg kg^{-1} yr^{-1} in most provinces. PNB_N was low in south China (< 0.40 kg kg^{-1} yr^{-1}) and moderate in most provinces (0.41-0.50 kg kg^{-1} yr^{-1}). The PFP_N in China decreased from 32 kg kg^{-1} in 1978 to 27 kg kg^{-1} in 1995, after which it increased to 38 kg kg^{-1} in 2015. PNB_N varied from 0.53 in 1978 to 0.38 kg kg^{-1} in 2000, after which it remained constant until 2015.

Stepwise multiple linear regression is an effective and powerful modelling approach to model and predict NUE and can derive the major influencing factors of the dependent variables. The models derived in this chapter explained more than 70% of the variation of NUE. Crop types and various soil properties were influential factors of the PFPN model, while crop types, climate and soil properties accounted for most of the variation of PNBN. Although the models could explain a large part of the spatial and temporal variation, they may be improved by expanding the covariate set with additional relevant variables and by exploring the use of non-linear statistical models. Suitable crop types, temperature and soil properties should be considered by policy makers when taking decisions on developing agricultural land management in an agricultural resource use efficiency way, by balancing NUE, productivity, and the environment.

Supplementary materials

The supplementary materials of this chapter, Appendix A, can be downloaded from the journal version of this chapter: Liu, Y., Heuvelink, G. B. M., Bai, Z., He, P., & Masiliūnas, D. (2020). Space-time statistical analysis and modelling of nitrogen use efficiency indicators at provincial scale in China. European Journal of Agronomy, 115, 126032. https://doi.org/10.1016/j.eja.2020.126032.

Chapter 3

Analysis of spatio-temporal variation of crop yield in China using stepwise multiple linear regression

With increasing discrepancies between population growth and food production in China, the monitoring of crop yield is essential to support food security policies. However, current studies about spatio-temporal variation of yield mainly focus on the influence of climatic factors on grain crops, and do not explore the contributions of agricultural, environmental and economic factors on crop yield in China. In this study, a large yield dataset, covering 31 provinces and a 38-year period from 1978 to 2015, and related explanatory variables were collected for analyzing the spatio-temporal variation of different yield aggregations using stepwise multiple linear regression. At the national scale, the average aggregate yield increased from 3.04 Mg ha-1 in 1978 to 10.04 Mg ha^{-1} in 2015. Overall, the average aggregate yield increased in all provinces, but the average annual growth rates varied: it was smaller than 2.5% in Heilongjiang, Guizhou, Beijing, Qinghai and Jilin, more than 4.0% in Hainan, Guangxi, Ningxia, Hebei and Shaanxi, and between 2.5% and 4.0% in other provinces. The spatial patterns of the average yield from 1978 to 2015 were different for different crop aggregations. Most of the regression models explained more than 60% of the yield variance, except for rice, potato and cotton models. Agricultural management practices, soil and economic covariates were important explanatory variables in all models. Topography and climatic covariates were also important for some of the crop models. The regression model of the aggregate yield for all crop explained 95% of the yield variance, which was mainly explained by planting area index of vegetables (20%), followed by farmer income (14%), planting area index of other crops (orchards 11%, melons 8%, sugar 6%, cereals 6%), and density of agricultural diesel engines (5%). Although the regression residual of the aggregate yield model was zero on average, the trends were different in different provinces: most provinces demonstrated a small negative or positive residual; the yield was substantially lower (< -0.20 Mg ha^{-1} yr^{-1}) than predicted by the regression model in three provinces in central China (Hebei, Shaanxi and Anhui) and substantially higher (> 0.20 Mg ha^{-1} yr^{-1}) in four provinces (Shanxi, Shandong, Sichuan and Guangdong). These systematic over- and underpredictions may be caused by other factors, such as plagues, pests, natural hazards, market structures (such as competition for labor or impediments to market access) and farmer's management skills. With the increasing population and limited agricultural land resources, enhancing economic growth might be an adequate solution to meet the growing demand for food. It can also promote agricultural efficiency in China, certainly when combined with better management practices, crop composition, breeding and planting technologies.

Based on:

Liu, Y., Heuvelink, G. B. M., Bai, Z., He, P., Xu, X., Ding, W., & Huang, S. (2021). Analysis of spatio-temporal variation of crop yield in China using stepwise multiple linear regression. *Field Crops Research*, *264*, 108098.

3.1. Introduction

China must make strategic decisions on enhancing food production and ensuring food security for 1.4 billion people, and those decisions will have a large effect on agriculture and land use. Over the past 60 years, crop yield in China has increased dramatically, on average by 125 kg ha^{-1} yr^{-1} (FAO, 2018). For example, the average annual yield increases are 2.5, 1.7 and 3.1% for maize, rice and wheat, respectively (National Bureau of Statistics of China, 2019). However, there are large spatio-temporal variations of crop yields in China, among others due to climate and soil variation (Chen et al., 2011a). Overall, the grain production in eastern China was higher than that in the west of China (Wang et al., 2018a). Therefore, understanding the temporal and spatial variation of crop yield across China is crucial for national food availability and food security.

Many studies were conducted on analyzing the spatio-temporal variation of crop yields in China. High-yield maize was mainly produced in Heilongjiang, Jilin and Liaoning provinces in northeast China, with average yields of 1.07, 1.38 and 1.52 Mg ha^{-1} in 1961 and increasing rates of 54 kg ha^{-1} yr^{-1}, 81 kg ha^{-1} yr^{-1} and 67 kg ha^{-1} yr^{-1} over the past 60 years, respectively (Guo et al., 2017). Rice yield increased by 23 kg ha^{-1} yr^{-1} since 2000. The highest rice yield was produced in central China with 7.07 Mg ha^{-1} in 2015; while the lowest rice yield was in south China, with 5.85 Mg ha^{-1} in 2015 (Wang et al., 2018d). Wheat yield varied from 6.33 Mg ha^{-1} in north China to 14.80 Mg ha^{-1} in south China. It also increased from west to east along the same latitude (Lv et al., 2017). Sunflower yield was steady from 1985 to 2008 and increased slowly afterwards; in 2015, productive sunflower was mainly distributed in north and central China (Gansu 3.67 Mg ha^{-1}, Hebei 2.82 Mg ha^{-1}, Xinjiang 2.81 Mg ha^{-1}, Inner Mongolia 2.74 Mg ha^{-1}, Ningxia 2.56 Mg ha^{-1}) (Fu et al., 2019).

In order to improve crop yields, exploring the causes of spatio-temporal yield variation is helpful to design policies. Important factors of agricultural production and its variation include technology, genetics, climate, soil, field management practices and associated decisions such as fertilizer application, tillage and crop hybrid selection, irrigation management, row spacing, planting date and depth (Kukal & Irmak, 2018). On a global scale, over 21% of yield variation could be explained by agro-climatic variation (Iizumi & Ramankutty, 2016). Climatic factors, especially their annual variation, exhibit a stronger overall linkage to changes in late paddy rice yield of China than technological factors (Wang et al., 2016b). The warming trend increased rice yield in northeast China and soybean in north and northeast China; however, it decreased maize yield in seven provinces in central and northeast region and wheat yield in three provinces (central and northeast China) (Tao et al., 2008). Thus, the relationship between climatic factors and yield is scale-, location- and crop-dependent. Large-scale statistical data and regional climate datasets are important for investigating general response patterns of crop yields to climate change and variation.

Most studies on yield variation addressed either one specific crop (Deng et al., 2019) or main grain crops (Iizumi & Ramankutty, 2016). In such a case, the result does not reflect the integral productivity for all crops. Moreover, existing research mostly focused on the influence of climatic factors on yield (e.g., (Kukal & Irmak, 2018). Few studies aimed at explaining the yield variation using agricultural, environmental and economic factors, let alone explore the importance of these factors. As a consequence, a better understanding of the contributions of agricultural, environmental and economic factors on yield and yield variation is needed. A provincial-scale evaluation of the spatio-temporal variation for aggregate yield over a long-term period could provide a general overview and the required high-level information for decision makers.

A disadvantage of analysis of aggregate yield variation is that different crops may have very different yields. For instance, the average yield of tomatoes, rice and maize per unit area is quite different and only evaluating the aggregate yield would not recognize these differences. To account for this, it is also useful to run separate analyses for different crop categories and for individual crops. Such analyses do not suffer from the problem that yields from different crops are aggregated, but an important drawback is that the analysis must be done for many different crop categories, individual crops and perhaps even for different crop varieties. This may yield too detailed and too much information, thus obscuring general patterns. Decision and policy makers in particular need integrated information that shows general patterns and trends.

In this chapter, we collected agricultural, environmental and economic factors that may influence the spatio-temporal variation of yield and analyzed their correlation with the aggregate yield of all crops in 31 provinces in China, from 1978 to 2015. This study aims at: a) analyzing the space-time patterns of yield aggregations at multiple levels in China with large yield datasets; b) constructing empirical models of different yield aggregations using stepwise linear regression; c) exploring the major explanatory variables of spatio-temporal variation of different yield aggregations and their relative importance; d) analyzing the temporal and spatial patterns of the regression residual (difference between observations and predictions) for the models for yield aggregated over all crops, staples crops, cash crops and three individual crops (maize, rice and wheat).

3.2. Data and Methods

3.2.1. Yield and explanatory variables

We collected yield data from 1978 to 2015 of 31 provinces of China (excluding Hong Kong, Macao and Taiwan; Chongqing started from 1997). For potatoes, the data was from 1982 to 2015, while watermelon data was from 1996 to 2012. To assess the effect of fertilization policies on temporal variation of yield at provincial scale, we distinguished three fertilization periods: (a) high-yield fertilization (1978–1995), (b) balanced fertilization (1996–2005), (c) soil-test based fertilization (2006–2015).

The yield was calculated using data on crop production and planting area of various crops (cereals, beans, tubers, oil crop, sugar crop, fiber (fiber and cotton), tobacco, tea, vegetables, melons and fruits), originating from the National Bureau of Statistics of China (2019). Explanatory variables included topography, crop types, soil types and soil properties, climate, economy and agricultural management practices (AMP). The explanatory variables were obtained from the Climate Research Unit (1978-2015), (Harris et al., 2014), National Bureau of Statistics of China (2019) and published articles (Shangguan et al., 2014; Hengl et al., 2017; SoilGrids, 2018), see Table B.1. For convenience, we use abbreviations to represent these variables (see Table B.1). All explanatory variables available as raster maps were aggregated to provincial scale by taking the spatial average over all raster cells within a province. Most of the variables were dynamic, except for enhanced vegetation index, topography and soil covariates, which were considered static because these had negligible temporal variation during the period studied. Crop and soil types were transformed to continuous-numerical variables by computing area proportions. In this study, we used 30 reference soil groups of the World Reference Base (Hengl et al., 2017). Climatic variables were detrended according to the growing season of crops: we used seasonal averages for some crops (major staples crops: maize, rice, wheat, potato and soybean; major cash crops: cotton, peanut, rapeseed, watermelon, see Table B.2), and annual climate information for crops whose growing season is a full year or more (Ray et al., 2015). A total of 1,159 annual yield observations and 115 explanatory variables were employed to analyze and explain spatio-temporal variation in aggregate yield. Mean values of the explanatory variables used in the final aggregate yield model for each province are summarized in Table B.3.

3.2.2. Stepwise multiple linear regression and model fitting

The regression models used in this study can be divided into three levels according to crop yield aggregation: 1) all crops; 2) staples (aggregation of cereal, beans and tubers) and cash crop categories (aggregation of oil, sugar, fiber, tobacco, tea, vegetables, melons and fruits); 3) twelve individual crops: maize, rice, wheat, potato and soybean (major staples crops); cotton, sugarcane, peanut, rapeseed, apple, citrus, watermelon (major cash crops). If there is no further specification, then yield refers to the first level, that is the aggregate yield over all crops.

Regression analysis is widely used for prediction. Multiple linear regression models the relation between a dependent variable (i.e., crop yield) and explanatory variables by fitting a linear equation using observed data. A first step in multiple regression is to examine pairwise relationships among all variables, since this is helpful to understand the data. In some circumstances, the emergence and disappearance of relationships can indicate important findings that result from multiple regression models. The bivariate correlations and corresponding scatter plots were plotted in a single figure using the chart.Correlation function of the PerformanceAnalytics package, v2.0.4 of the R language for statistical computing (R Core Team, 2018) (See Figure 3.1).

Stepwise regression methods use criteria to select explanatory variables automatically, such as F-test, adjusted R^2, Akaike information criterion (Soergel et al.) and Bayesian information criterion (BIC). Here, we used BIC for variable selection and forward stepwise regression, as implemented in the stepAIC function in the MASS package, v7.3-53 of the R language for statistical computing (R Core Team, 2018). Furthermore, multicollinearity, which refers to a situation in which two or more explanatory variables in a multiple regression model are highly linearly related, can be detected using the variance inflation factor (VIF). A rule of thumb is that if VIF is more than 10 then multicollinearity is too high (Kutner et al., 2004). If this occurred, we deleted the explanatory variable with the highest VIF and refitted the model. We repeated this procedure until the condition that all VIFs are below 10 was satisfied. Next, we used the confint function in the stats package (v3.6.2) to check the statistical significance of the remaining explanatory variables. Insignificant variables (i.e., p-value larger than 0.05) were deleted in a stepwise procedure. A flowchart of the model selection process is given in Figure B.1.

In order to make the regression coefficients of the final model more interpretable, all explanatory variables were standardized to zero mean and unit standard deviation using the scale function in the base package of R. As a result, the regression coefficients describe the expected change in the dependent variable for a standard deviation change in an explanatory variable, holding the other explanatory variables constant. This is the simplest means to assess the relative importance of explanatory variables. But in this study, we used the method of relative weights to obtain the relative importance of explanatory variables (Kabacoff, 2011). This method closely approximates the average increase in R-square obtained by adding an explanatory variable across all possible sub-models. Multiple linear regression assumes that the regression residuals are independent and identically distributed Gaussian variables. The normality assumption was evaluated by computing the studentized residual, the histogram of which is generated in the 'residplot' function, which superimposes a normal curve, kernel density curve and rug plot, visually reflecting how close model errors are to a normal distribution (Kabacoff, 2011).

3.2.3. Validation/Accuracy assessment

Cross-validation is a resampling procedure used to evaluate prediction models on a limited data set. This method is commonly applied because it generally results in a more stable assessment of model performance than other methods, such as splitting a dataset into a calibration and validation set. In k-fold cross-validation, the original sample is randomly partitioned into k equally sized subsamples. Of the k subsamples, a single subsample is retained as validation data for testing the model, and the remaining k – 1 subsamples are used as training data. The cross-validation process is then repeated, with each of the k subsamples in the process used once as a validation set. The results from the k processes are merged to produce a single validation set. 10-fold cross-validation is widely used, also in this study, but other values of k can also be used.

The model validation metrics used were the mean error (ME), root mean squared error (RMSE), normalized RMSE (nRMSE), coefficient of determination (R^2) and Lin's concordance correlation coefficient (LCCC) (Lin, 1989). These are computed as follows:

$$\mathrm{ME} = \frac{1}{n}\sum_{i=1}^{n}(P_i - O_i) \tag{3.1}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_i - O_i)^2} \tag{3.2}$$

$$\mathrm{nRMSE} = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_i - O_i)^2}}{\bar{O}} \tag{3.3}$$

$$R^2 = \frac{\sum_{i=1}^{n}(P_i - \bar{O})^2}{\sum_{i=1}^{n}(O_i - \bar{O})^2} \tag{3.4}$$

$$\mathrm{LCCC} = \frac{2r\sigma_o\sigma_p}{\sigma_o{}^2+\sigma_p{}^2+(\overline{O}-\overline{P})^2} \tag{3.5}$$

where n is the number of observations; P_i and O_i are the predicted and observed yield values for observation i, respectively; $\overline{P}$ and $\overline{O}$ are the mean of predicted and observed values; σ_p and σ_o are the standard deviations of predicted and observed values; and r is the Pearson correlation coefficient between the predicted and observed values.

ME represents the prediction bias of the model (i.e., the systematic error). RMSE reflects how far off the predicted values on average are from the observed values. Smaller ME and RMSE values indicate better model performance. The nRMSE statistic (0 ≤ nRMSE ≤ 100%) is convenient for comparing the degree of agreement between predicted and observed values for variables measured in different units. It only applies to ratio variables (i.e., variables that have a natural zero). We consider nRMSE ≤ 15% as "good" agreement; 15–30% as "moderate" agreement; and ≥ 30% as "poor" agreement (Liu et al., 2013a; Liang et al., 2016). R^2 measures the closeness of the predicted and observed values to the 1:1 line. If it is over 0.90, this indicates a good agreement between observations and predictions (Aghdaei et al., 2017). LCCC is an alternative to R^2 and signifies the degree to which the predicted and observed values are close to the 1:1 line. An LCCC smaller than 0.90 is considered to be poor agreement, 0.90–0.95 as moderate agreement, 0.95–0.99 as substantial agreement, and larger than 0.99 as almost perfect agreement (de Beaufort et al., 2017).

In order to analyze the model performance for each province over 1978–2015, the temporal and spatial variation of the regression residual (i.e., observed yield minus predicted yield) was also calculated. Whenever the regression residual is positive (the observed yield is higher than predicted), the provinces in that year do better than expected based on explanatory variables. The opposite occurs in case of a negative regression residual.

3.3. Results

Because of space limitations, this section mainly focuses on the results of the first level modelling (aggregate yield over all crops) and presents fewer results for the level 2 (staples and cash crop categories) and level 3 (individual crops) analyses. More detailed results for levels 2 and 3 are provided in Appendix C.

3.3.1. Descriptive statistics of dependent and explanatory variables

Summary statistics of the yield and explanatory variables included in the final level 1 model are listed in Table 3.1. The yield exhibited a slightly skewed distribution, with skewness and kurtosis coefficients of 1.07 and 3.78, respectively, and a mean value of 6.59 Mg ha^{-1}, which was higher than its standard deviation (Std., 3.34 Mg ha^{-1}), indicating a fairly steady yield.

Table 3.1: Summary statistics of yield and explanatory variables included in the final model.

Classification	Property	Unit	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Std.	Skewness	Kurtosis
Dependent variable	Yield	Mg ha^{-1}	1.57	4.13	5.82	6.59	8.16	19.24	3.34	1.07	3.78
Topography	TPI	meter	-0.46	-0.06	0.02	0.03	0.08	0.82	0.23	1.16	6.27
	SLP	meter	0.0	0.1	0.2	0.2	0.2	0.4	0.1	0.07	1.96
Crop	VEG	%	0.0	3.2	5.9	7.8	10.9	32.4	6.1	1.35	4.75
	SUG	%	0.0	0.0	0.2	1.1	0.8	16.3	2.3	3.55	17.98
	BEAN	%	0.0	2.1	4.0	5.4	6.7	39.4	5.5	3.00	14.90
	ORC	%	0.1	1.1	3.3	5.0	6.9	23.9	4.9	1.39	4.35
	CER	%	25.9	47.7	58.1	56.5	65.6	93.0	12.7	-0.22	2.63
	MEL	%	0.0	0.4	0.7	1.0	1.2	5.6	0.9	2.18	9.64
	TEA	%	0.0	0.0	0.1	0.8	1.0	7.9	1.3	2.41	9.09
Soil	EXH	cmol kg^{-1}	1.05E-03	1.71E-03	2.18E-03	2.27E-03	2.75E-03	4.91E-03	8.07E-04	1.07	4.72
	PHO	ppm of weight	0.00	0.02	0.04	0.06	0.04	0.28	0.07	1.79	4.99
	TC	% of weight	0.57	1.06	1.37	1.49	1.90	3.33	0.58	0.94	4.35
	TP	% of weight	0.03	0.04	0.05	0.05	0.06	0.08	0.01	0.21	2.31
	SILT	%	27.6	36.1	41.3	41.4	44.0	60.9	6.5	0.64	4.25
	HP	%	0.0	0.3	0.5	0.7	0.9	2.8	0.6	1.75	6.64
	ALI	%	0	0	2	4	6	24	6	1.77	6.09
	LUV	%	4	7	12	13	18	31	8	0.62	2.38
	CRY	%	0	0	0	1	0	6	2	2.38	7.40
	CAM	%	8	13	18	20	22	58	11	2.02	7.15
	GLE	%	1	2	4	4	6	10	3	0.78	2.40
	SOL	%	0	0	1	2	4	9	2	1.22	3.50
Economic variables	INC	RMB $person^{-1}$	142	479	1697	2949	3950	23205	3537	2.09	8.20
AMP	TOW	$(1000\ ha)^{-1}$	0	3	8	16	20	333	26	4.70	39.73
	DIE	$(1000\ ha)^{-1}$	0	5	15	27	34	173	34	2.18	7.48
	TRA	$(1000\ ha)^{-1}$	0	3	6	12	13	419	23	10.24	150.93
	TRAS	$(1000\ ha)^{-1}$	1	21	43	64	70	561	78	3.53	18.52
	IRRI	ha ha^{-1}	0.1	0.2	0.3	0.3	0.4	1.0	0.2	1.54	5.87

Notes: Min.: minimum; 1st Qu.: first quartile; 3rd Qu.: third quartile; Max.: maximum; Std.: standard deviation; See Abbreviations for explanation of the variable code names.

Pearson's correlation coefficients were used to compare the pairwise relationships between yield and explanatory variables (Figure 3.1). There was a significant negative correlation between yield and local upslope curvature, exchangeable acidity, planting area index of beans, cryosols, total phosphorus, amount of water-soluble phosphorus, and planting area index of cereal, while no significant correlation between yield and planting area index of tea, topographic position index (TPI), solonchaks, histosol probability, alisols, cambisols, gleysols was found. Yield had a positive correlation with thirteen other explanatory variables. Most of the explanatory variables were significantly correlated with each other.

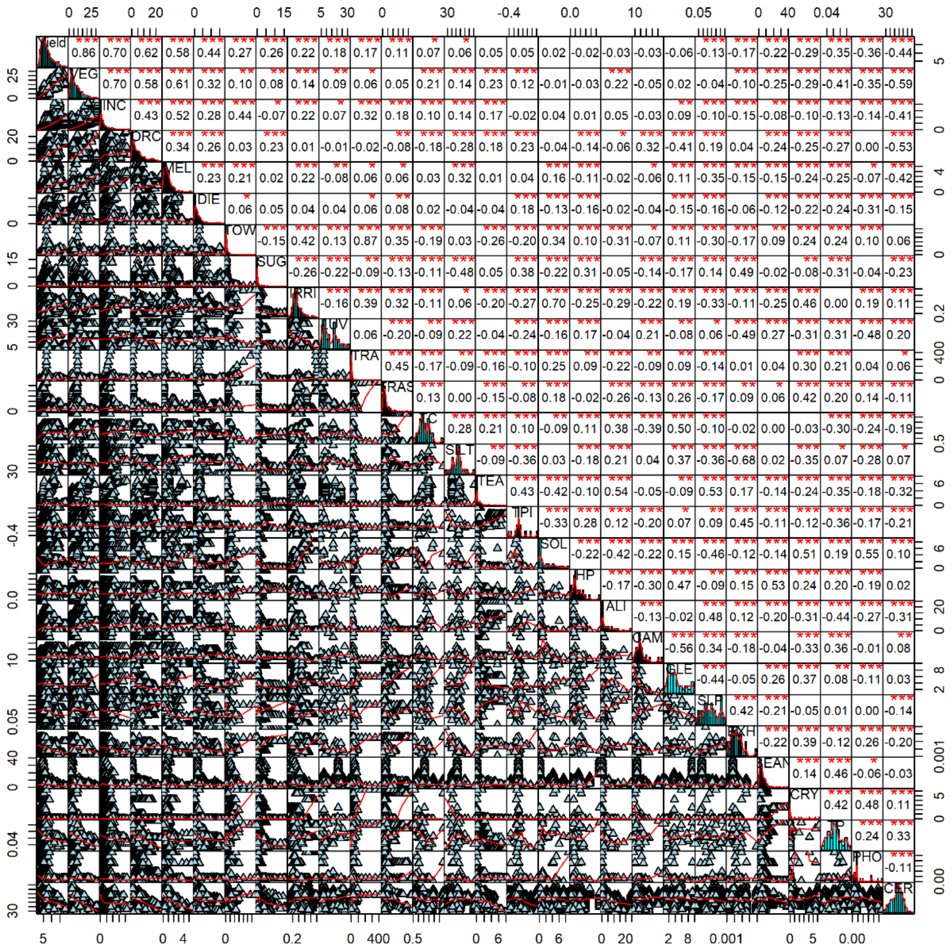


Figure 3.1: Pairwise comparison of yield and all explanatory variables used in the level 1 yield model. The upper-right panel contains predicted pair-wise Pearson's correlations. The diagonal panel shows histograms and the lower-left panel shows scatterplots with a LOESS smoother added to aid visual interpretation. Top row and left columns show bivariate relations between yield and explanatory variables.

Note: See Abbreviations for explanation of the variable code names. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

3.3.2. Temporal and spatial variations of yield from 1978 to 2015

A large temporal increase of aggregate yield from 1978 to 2015 was detected (Figure 3.2a). At the national scale, it increased from 3.04 Mg ha^{-1} in 1978 to 10.04 Mg ha^{-1} in 2015 (average yield of 6.37 Mg ha^{-1} yr^{-1}), with an average annual growth rate of 3.3%. The yield increased in all provinces but the average annual growth rates varied. It was smaller than 2.5% in Heilongjiang, Guizhou, Beijing, Qinghai and Jilin, more than 4.0% in Hainan, Guangxi, Ningxia, Hebei and Shaanxi, and between 2.5% and 4.0% in other provinces (Figure 3.2a).

Figure 3.2b shows results of the yield variation over time for crop categories and individual crops, on a national scale. The average yields of staples and cash crop were 4.17 and 12.79 Mg ha^{-1} yr^{-1}, respectively; while their average annual growth rates were 2.1 and 2.6%, respectively. The average yields of cotton, rapeseed and soybean were smaller than 2 Mg ha^{-1} yr^{-1} (0.98, 1.45 and 1.55 Mg ha^{-1} yr^{-1}, respectively) while they were higher than 10 Mg ha^{-1} yr^{-1} for watermelon and sugarcane (11.46 and 19.83 Mg ha^{-1} yr^{-1}, respectively). The lowest average annual growth rate was less than 2.5% (1.5, 1.5, 1.7, 1.7, 1.9 and 2.0% for soybean, rice, potato, sugarcane, watermelon and maize, respectively). The highest average annual growth rates were more than 4.0% (4.7 and 5.3% for apple and citrus, respectively) (Figure 3.2b).

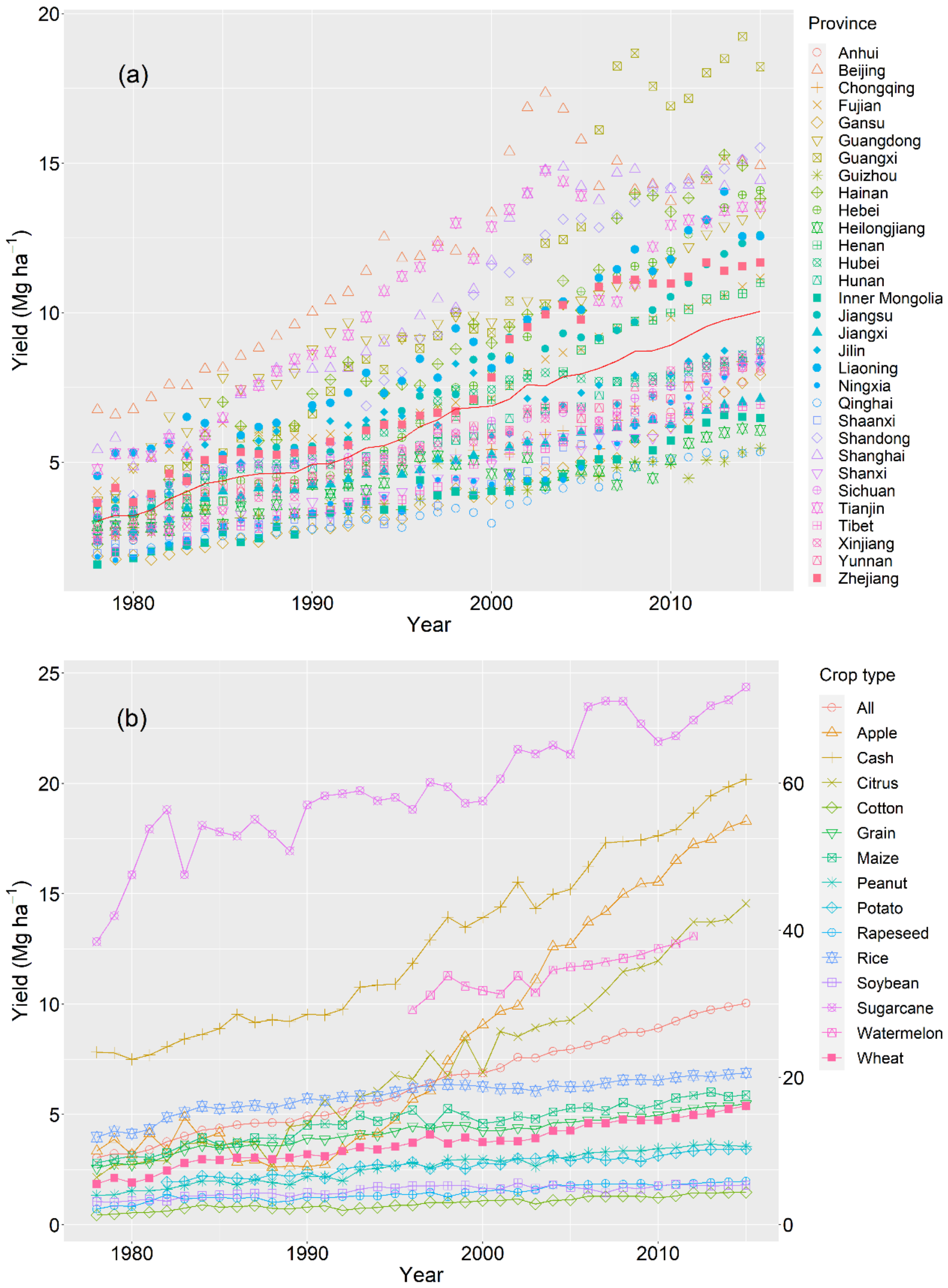


Figure 3.2: Temporal variation of crop yield from 1978 to 2015: (a) aggregate yield (points refer to provincial yield; red line is the national average); (b) national yield for different crop aggregations (sugarcane and watermelon units are plotted on the Y-axis on the right).

We grouped the aggregate yield into five grades: very low (< 4 Mg ha^{-1} yr^{-1}), low (4–6 Mg ha^{-1} yr^{-1}), moderate (6–8 Mg ha^{-1} yr^{-1}), high (8–10 Mg ha^{-1} yr^{-1}), and very high (> 10 Mg ha^{-1} yr^{-1}), and derived maps of the yield for the three fertilization periods (1978–1995, 1996–2005, 2006–2015) (Figure 3.3). Overall, the yields in the eastern coastal provinces were higher than those in the inland provinces; and higher in east and south China than in west and north China. This is most clear for the first period (1978–1995, Figure 3.3a). As the yield increased faster in west China, the yield discrepancies became smaller during the 2006–2015 period. Between 1978 and 1995 the yield was low (less than 6 Mg ha^{-1} yr^{-1}) in most of the provinces, moderate (6–8 Mg ha^{-1} yr^{-1}) in Liaoning, Shanghai, Guangdong and Tianjin, and high (8.99 Mg ha^{-1} yr^{-1}) in Beijing. From 1996 to 2005, there was not a single province with a yield at the very low level; most of the provinces had a yield at low and moderate level; four provinces (Henan, Fujian, Jiangsu and Zhejiang) had yield at the high level; and nine provinces (Hebei, Liaoning, Guangdong, Hainan, Shandong, Tianjin, Shanghai, Guangxi, and Beijing) had yield at the very high level. From 2006 to 2015, four provinces (Guizhou, Qinghai, Heilongjiang and Inner Mongolia) had yield at low level; twelve provinces had yield at moderate level; Hunan and Hubei had yield at high level and the other thirteen provinces had yield at the very high level.

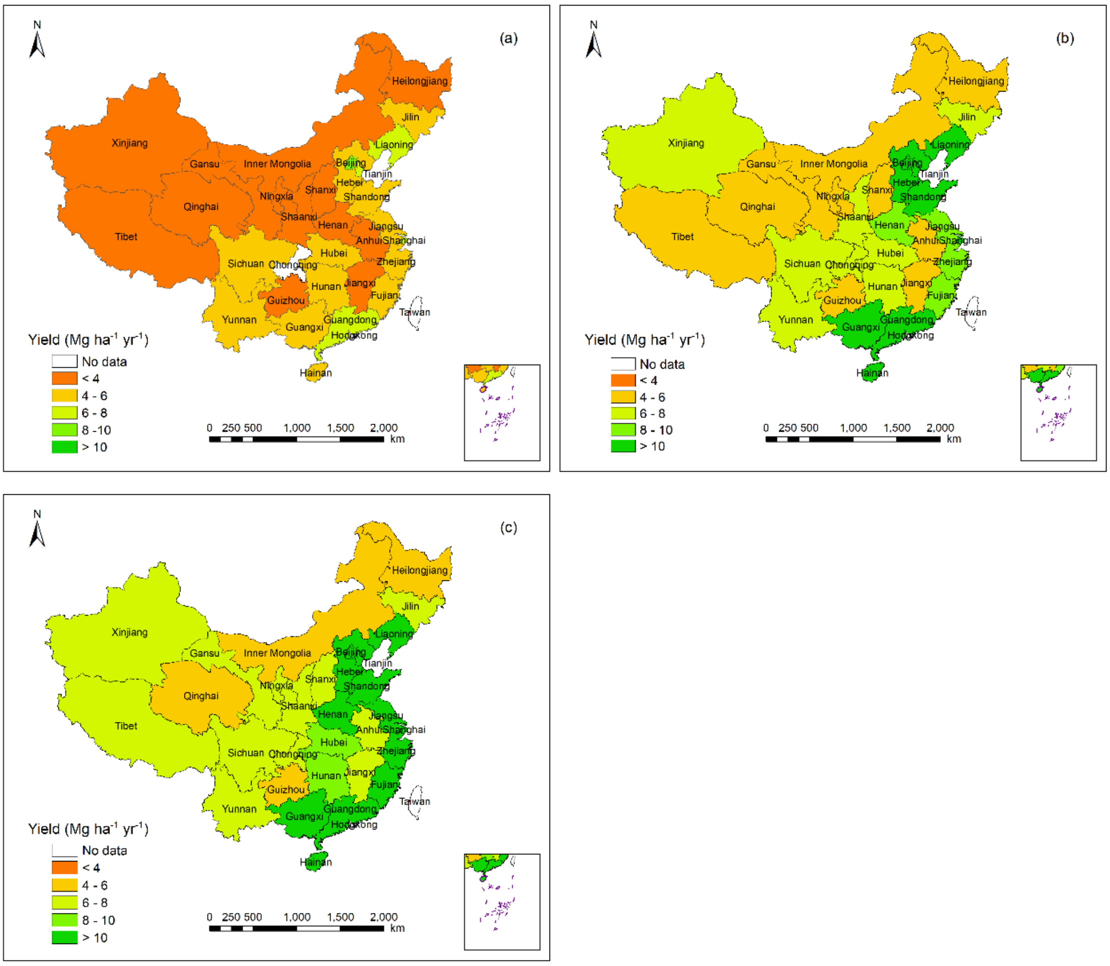


Figure 3.3: Spatial distribution of the average yield aggregated over all crops per province for three periods: (a) high yield fertilization (1978–1995), (b) balanced fertilization (1996–2005), (c) soil-test based fertilization (2006–2015).

The spatial patterns of the average yield from 1978 to 2015 were different for different crop yield aggregations (Figure 3.4). We present results for both crop categories (staples and cash) and for three individual crops (maize, rice and wheat). Results for all other individual crops are shown in Appendix C. For staples crops, the yield was less than 3.0 Mg ha^{-1} yr^{-1} in north China; yield between 3.0 and 4.0 Mg ha^{-1} yr^{-1} mostly located in southwest China, northeast China, northwest China and north central China. Yields higher than 4.0 Mg ha^{-1} yr^{-1} were mostly in east and south China, except for Xinjiang (Figure 3.4b). The spatial pattern of cash crops (Figure 3.4c) was similar to that of the aggregate yield (especially the yields in the eastern coastal provinces, which were much higher than those in the inland provinces). The maize yields (Figure 3.4d) were higher in north China than in south China. For rice, the yields in middle China were higher than those in north and south China. Rice yields in south China were lower than 5 Mg ha^{-1} yr^{-1}, while rice yield in northwest China (Ningxia) was the highest (8.11 Mg ha^{-1} yr^{-1}) (Figure 3.4e).

The spatial pattern of wheat yield was similar to that of rice, but only 8 provinces had a yield lower than 2.5 Mg ha^{-1} yr^{-1}. These were mostly provinces in north and south China. The wheat yield in south China (Guangxi and Jiangxi) was the lowest (< 1.5 Mg ha^{-1} yr^{-1}) while the highest yield (> 4.5 Mg ha^{-1} yr^{-1}) was situated in north-central China (Shandong and Beijing) and west China (Tibet) (Figure 3.4f).

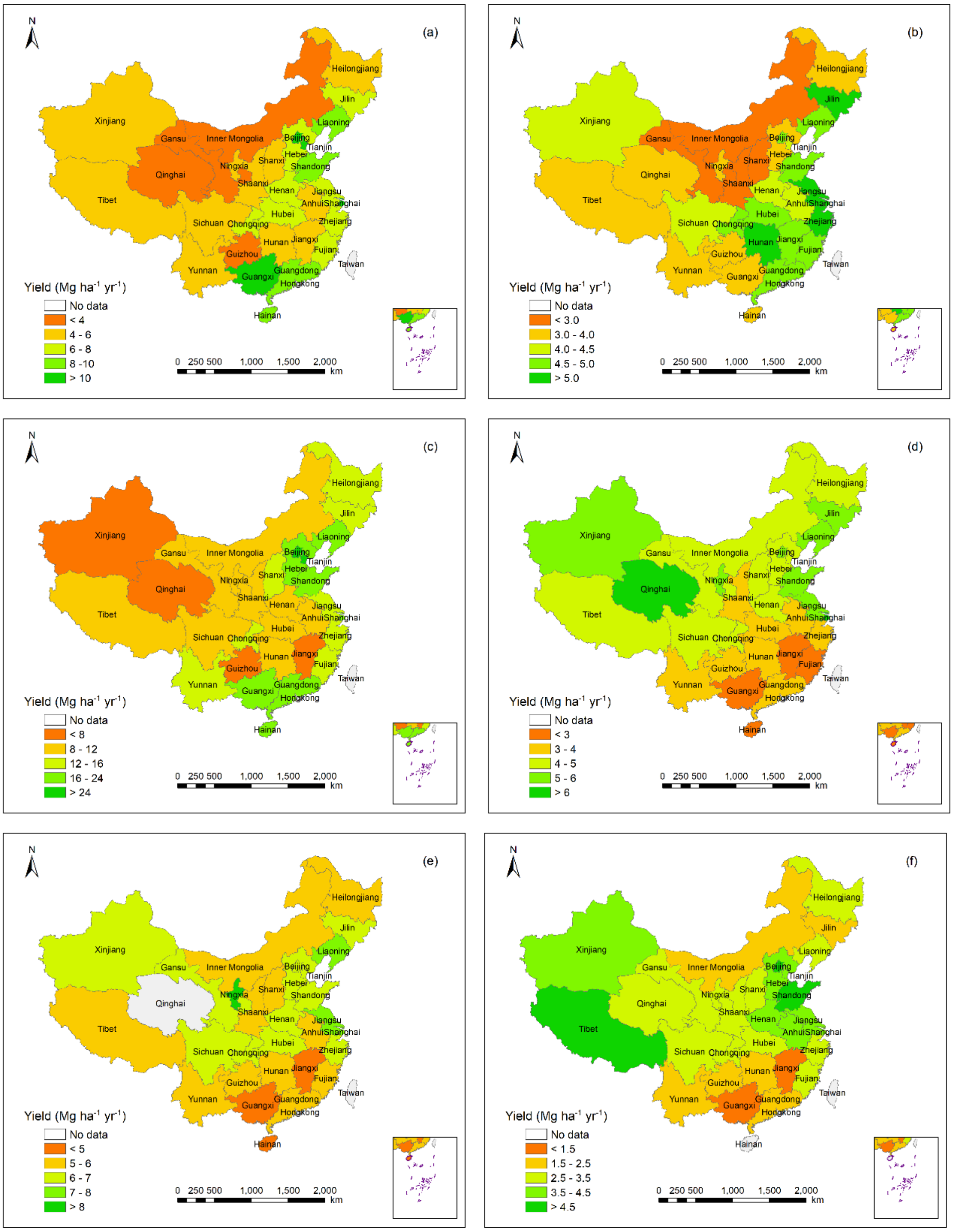


Figure 3.4: Spatial distribution of the average yield per province from 1978 to 2015: (a) all crops, (b) staples crops, (c) cash crops, (d) maize, (e) rice, (f) wheat.

3.3.3. Stepwise multiple linear regression model and performance evaluation

The initial stepwise regression model included 45 explanatory variables, which explained 97% of the variance of aggregate yield. Gross domestic product of agriculture (GDP), rural population, temperature, wet day frequency and other explanatory variables were dropped from the regression model in the VIF step to remove multicollinearity. This reduced the complexity of the model to 28 explanatory variables. Next agricultural electronic engines were removed because it was not statistically significant. The final model had 27 explanatory variables and explained 95% of the aggregate yield variance (Figure B.1).

The regression coefficients of the aggregate yield model are shown in Table 3.2. Note that these were obtained using standardized explanatory variables and hence can be mutually compared. The explanatory variables significantly affected the yield ($p < 0.01$). There were fifteen explanatory variables that had a positive impact on yield, while the other twelve explanatory variables had a negative contribution.

Table 3.2: Regression coefficients of the yield model.

Property	Regression coefficients	Std. error
(Intercept)	6.59	0.02
VEG	1.87	0.06
SUG	1.24	0.04
TOW	1.00	0.06
IRRI	0.60	0.05
ORC	0.55	0.05
SLP	0.39	0.06
TP	0.38	0.06
TRAS	0.36	0.04
INC	0.35	0.04
GLE	0.31	0.05
LUV	0.31	0.05
DIE	0.29	0.03
TC	0.22	0.04
CER	0.19	0.05
MEL	0.19	0.04
TPI	-0.11	0.04
ALI	-0.16	0.05
PHO	-0.20	0.06
BEAN	-0.20	0.05
SOL	-0.26	0.07
TEA	-0.33	0.04
CAM	-0.40	0.05
HP	-0.41	0.05
SILT	-0.42	0.07
CRY	-0.50	0.07
EXH	-0.59	0.06
TRA	-0.80	0.06

Note: See Abbreviations for explanation of the variable code names.

The regression diagnostics showed that the model studentized residuals follow a normal distribution quite well (Figure 3.5). None of the explanatory variables had a VIF greater than 10 (Table 3.3), which is as expected because we corrected multicollinearity during model selection.

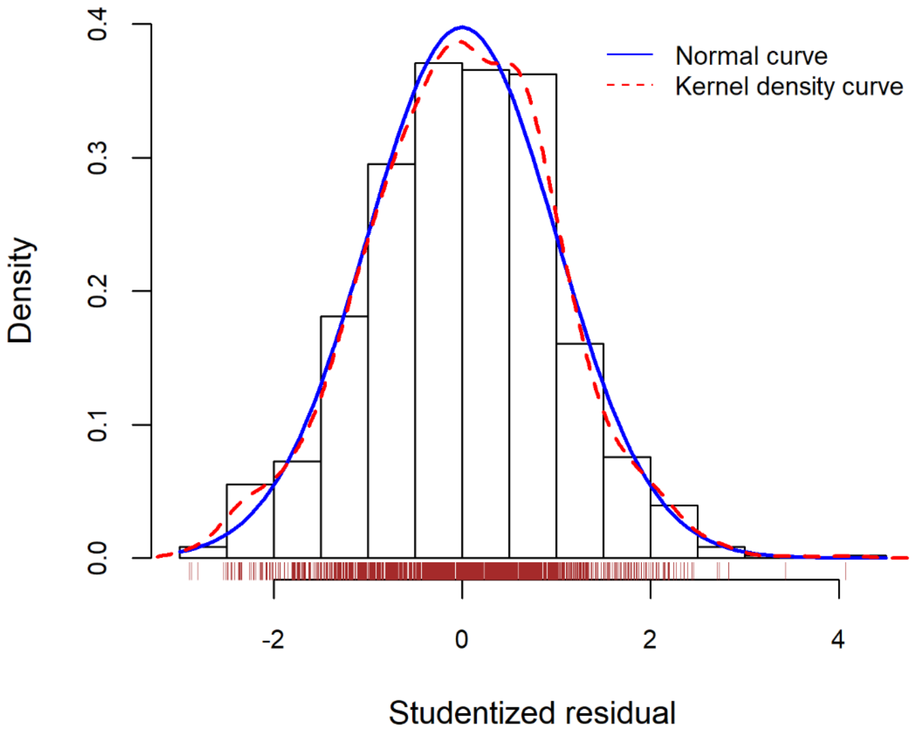


Figure 3.5: Distribution of the studentized residual of the aggregate yield model.

Table 3.3: Variance inflation factors (VIF) of the aggregate yield model.

Classification	Property	VIF
Topography	TPI	3.8
	SLP	8.0
Crop	VEG	7.9
	SUG	2.7
	BEAN	4.4
	ORC	4.3
	CER	5.1
	MEL	2.9
	TEA	3.8
Soil	EXH	7.2
	PHO	7.2
	TC	2.6
	TP	8.2
	SILT	9.0
	HP	5.0
	ALI	5.3
	LUV	4.2
	cRY	9.4
	CAM	4.3
	GLE	5.6
	SOL	9.3
Economic variables	INC	3.6
AMP	TOW	8.4
	DIE	1.8
	TRA	7.0
	TRAS	2.5
	IRRI	5.0

Note: See abbreviations (Table B 1) for explanation of the variable code names.

The 10-fold cross-validation results are shown in Table 3.4. The coefficient of determination defined in Equation 3.4 was 0.95 and practically equal to the adjusted R-square of the multiple regression model, indicating no over-fitting problem. Other accuracy verification metrics also showed good model performance, with low ME (0.00 Mg ha^{-1}), RMSE (0.75 Mg ha^{-1}), nRMSE (11%) and high LCCC (0.97). The scatter plot of observations against predictions and density plots of observations and predictions also showed that the model predicted yield quite well, even though it slightly smoothed the data and under-predicted extremes (Figure 3.6).

Table 3.4: Model performance and cross-validation (CV) metrics of the yield model.

Metrics	Yield	Yield CV
ME	0.00	-0.00
RMSE	0.75	0.77
nRMSE	11%	12%
R2	0.95	0.95
LCCC	0.97	0.97

Note: CV means 10-fold cross-validated results.

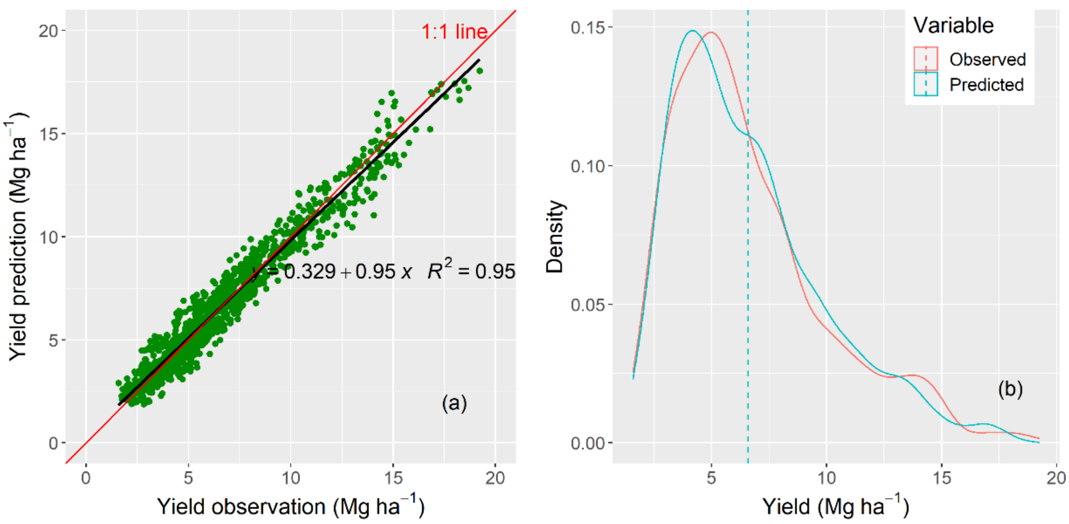


Figure 3.6: Scatter (a) and Kernel density (b) plots of observed and predicted aggregate yield.

Note: the dotted line means average value in Figure 3.6b; the average values of observed and predicted yield were the same.

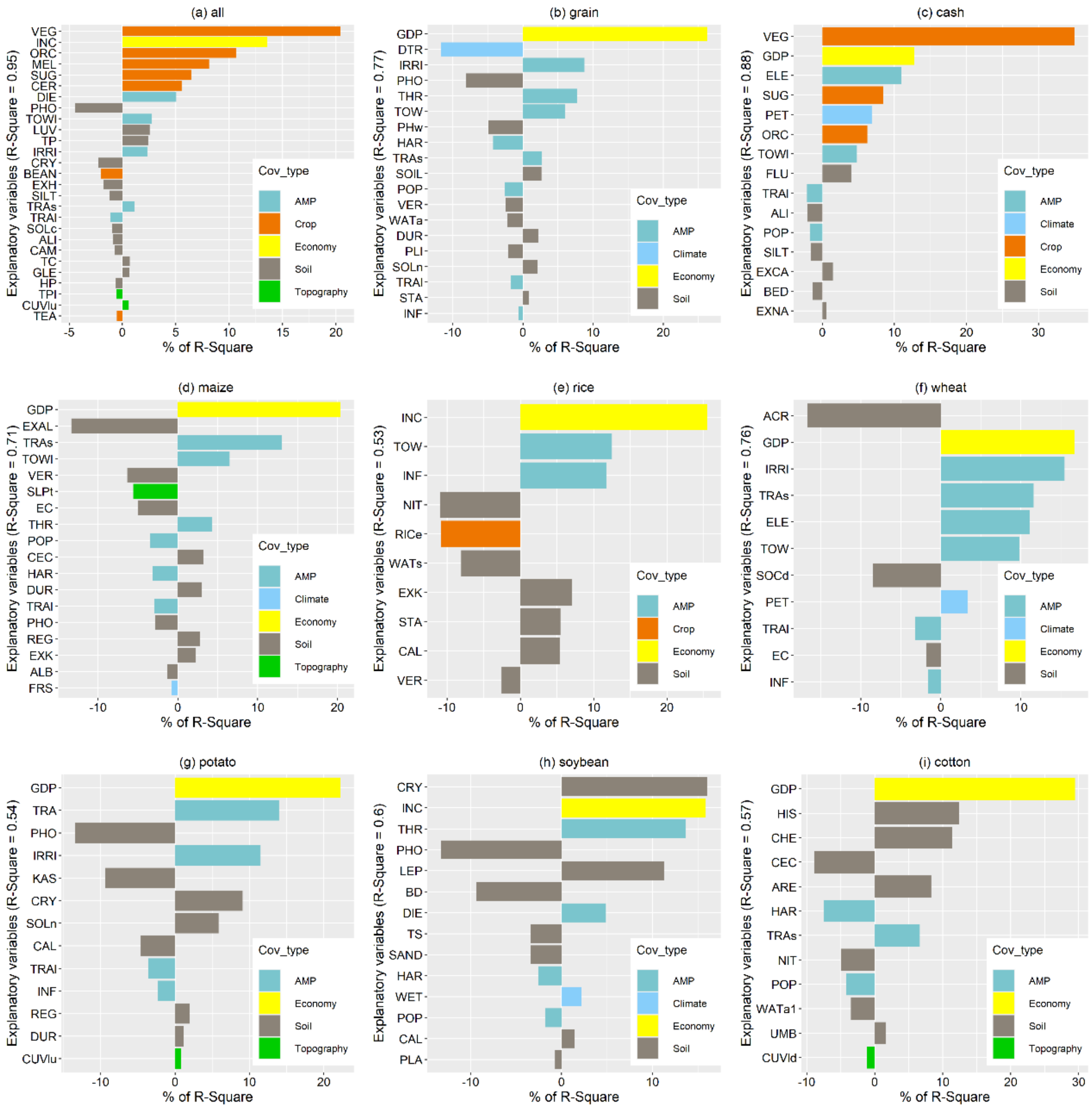
For other aggregation levels, the model performance between original and 10-fold cross-validation results was very similar, indicating no over-fitting problem (Appendix C). From the R-square (Figure 3.7 and Appendix C), we understand that the cash crops model (R-square of 0.88) explained more yield variation than the staples crops model (R-square of 0.77). Most of the individual cash crop models had an R-square higher than or around 0.7, except for the cotton model. On the contrary, for the individual staples crop models, only the maize and wheat models had a high R-square of 0.71 and 0.76, respectively, while the soybean, potato and rice models had an R-square lower or around 0.6. Note that the rice model had a low nRMSE (16%), suggesting that it was a moderate to good model.

3.3.4. Relative importance of explanatory variables

The relative importance of explanatory variables in yield models at different aggregation levels is shown in Figure 3.7. Most crop models had an R-square larger than 0.6, except for the models for rice, potato and cotton. In summary, AMP, soil and economic covariates were included in all models. Crop type covariates were important in the aggregate yield and cash crops models, but not for the staples crops model. Topography had influence in the aggregate yield model with low relative importance but had no impact on the staples and cash models. Instead, climatic covariates were relatively important for the staples and cash models, but not for the aggregate yield model. For individual crop yield models, the sugarcane model was the only model where topography and climate were both important and had a contribution larger than 5% of R-square. Topographic covariates were also important for the maize and rapeseed models (contribution > 5% of R-square), while climatic covariates were also important for the watermelon model.

For the aggregate yield model, the R-square was 0.95 (Figure 3.7a). The crop type had the highest importance (54% of the R-square, e.g., planting area index of vegetables 20%, orchards 11%, melons 8%, sugar 6%, cereals 6%), followed by soil covariates (19%, e.g., water soluble phosphorus 4%, luvisols 3% and total phosphorus 2%), economic variables (farmer income 14%), AMP (12%, e.g., agricultural diesel engines 5%, towing farm machinery of large and medium-sized agricultural tractors 3% and available irrigation area index 2%). Topography had the least influence, accounting for only 1% of the R-square (TPI 0.6% and local upslope curvature 0.6%). Most crops had a positive influence on yield (i.e., they had a positive partial regression coefficient), except for beans and tea (Figure 3.7a). AMP had similar characteristics: positive effects on yield for most explanatory variables but a detrimental impact of large and medium-sized agricultural tractors.

Farmer income was the only economic variable included and had, as expected, a positive influence on yield. Soil covariates had a more complex influence: many of them had a negative impact on yield, while luvisols, total phosphorus, total carbon and gleysols affected yield positively. The two explanatory variables in the topography group showed the same absolute value of relative importance: a negative impact of TPI and a positive effect of local upslope curvature.



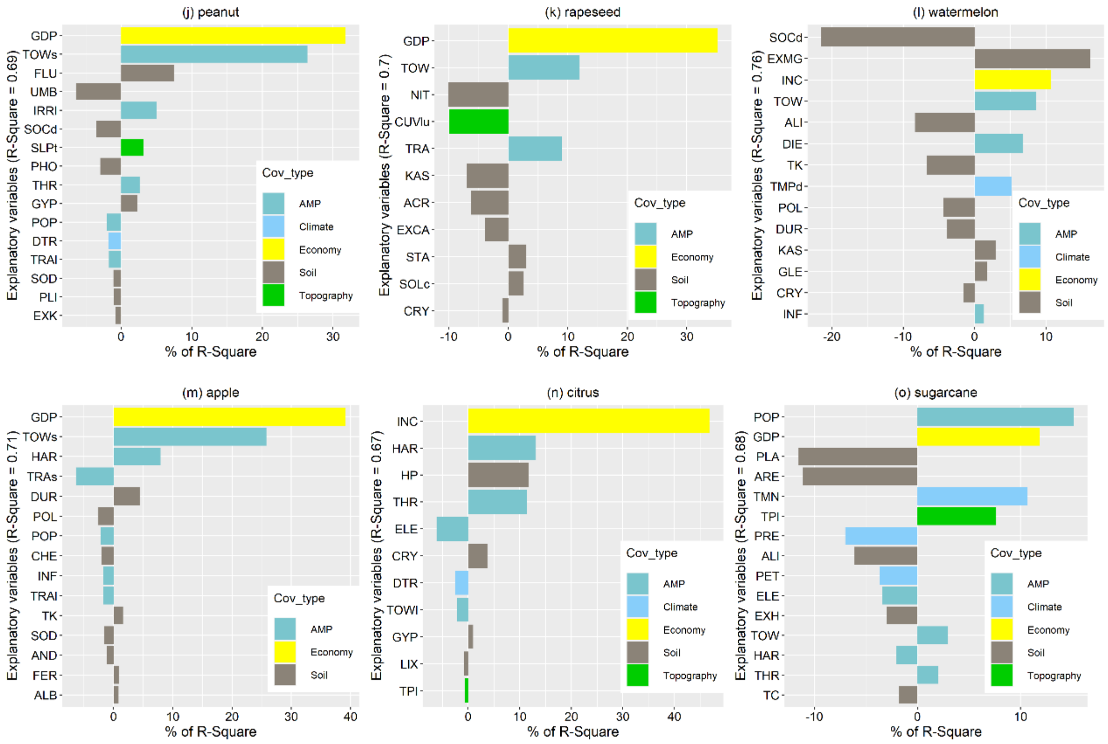


Figure 3.7: Relative importance of explanatory variables for the yield models for all three crop aggregation levels. Negative values mean that the explanatory variables have negative influence on yield; positive values, a positive influence.

3.3.5. Temporal and spatial patterns of regression residuals

Temporal variation trends of the regression residual varied between provinces: aggregate yield was over-predicted first and then under-predicted in northeast, northwest and northcentral China (six provinces) (Figure 3.8a; green color in Figure 3.9a). Conversely, yield was under-predicted first and over-predicted later in west and south China (eight provinces) (Figure 3.8b; red color in Figure 3.9a). Observed yield was higher than predicted yield in the mid-term and lower than predicted in the beginning and end of the 1978–2015 period in north, central and southwest China (six provinces) (Figure 3.8c; yellow color in Figure 3.9a). Unstructured variation in regression residual was observed in central, south and northeast China (eleven provinces) (Figure 3.8d; orange color in Figure 3.9a).

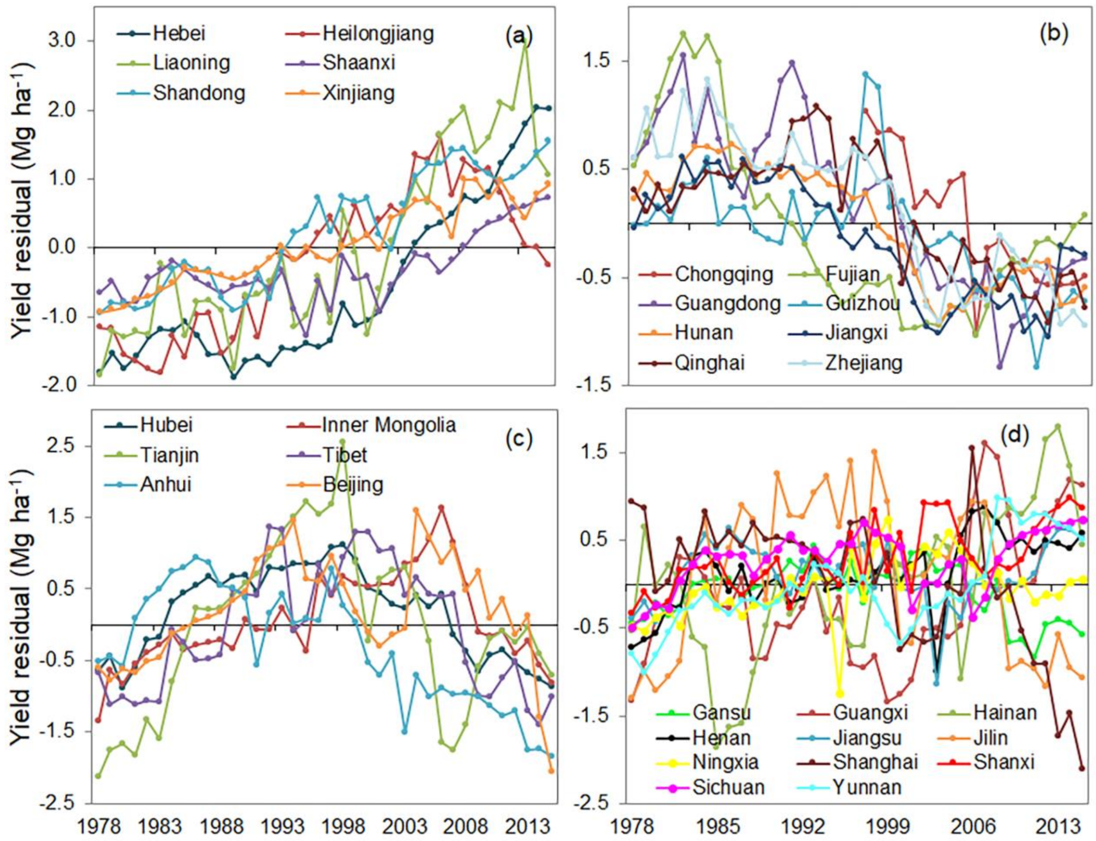


Figure 3.8: Temporal variation trend of regression residual for all provinces: (a) increasing trend, (b) decreasing trend, (c) arched trend, (d) complex fluctuation.

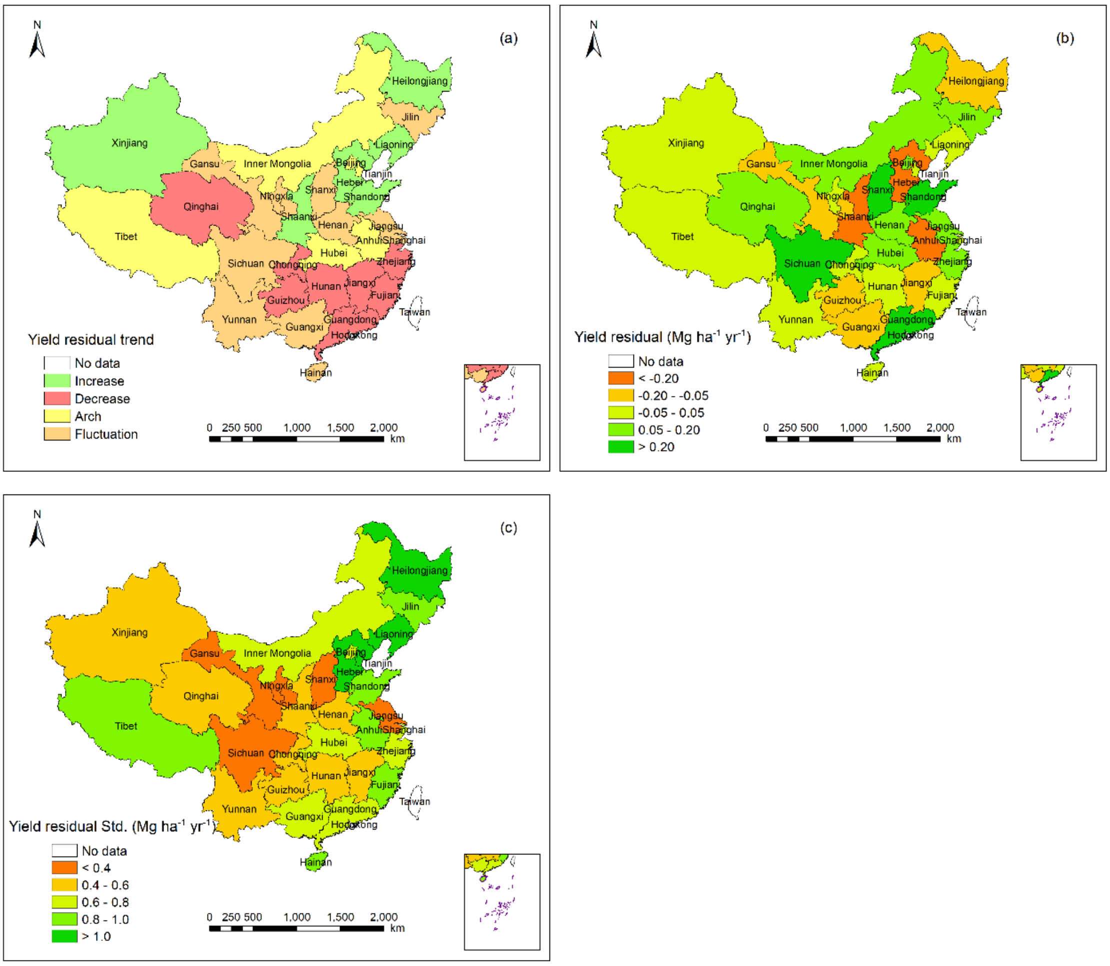


Figure 3.9: Spatial patterns of regression residual characteristics during 1978–2015: (a) temporal trend shape; (b) mean values; (c) standard deviation (Std.).

On average, most provinces demonstrated a near zero or positive regression residual (i.e., these provinces produced similar or higher yield when compared with predicted yield from explanatory variables). Eleven provinces had a small residual (-0.05 – 0.05 Mg ha^{-1} yr^{-1}) and were mostly situated in north and central China. Eight provinces in north and central China showed a large positive residual (0.05 – 0.20 Mg ha^{-1} yr^{-1}), while four provinces (Shanxi, Shandong, Sichuan and Guangdong in central and south China) had very large positive residual (> 0.20 Mg ha^{-1} yr^{-1}). Five provinces in south, west and northeast China indicated a large negative residual (-0.05 – -0.20 Mg ha^{-1} yr^{-1}); three provinces (Hebei, Shaanxi and Anhui in central China) had a very large negative residual (< -0.20 Mg ha^{-1} yr^{-1}) (Figure 3.9b). There was a considerable spatial variation in regression residual standard deviation (Figure 3.9c). It was higher in east China than in west China. Five provinces in central China had very small regression residual standard deviation (smaller than 0.4 Mg ha^{-1} yr^{-1}); nine provinces in central and west China had small regression residual standard deviation (0.4 –

0.6 Mg ha^{-1} yr^{-1}); seven provinces in east China had moderate regression residual standard deviation (0.6 – 0.8 Mg ha^{-1} yr^{-1}); six provinces in east and southwest had large regression residual standard deviation (0.8 – 1.0 Mg ha^{-1} yr^{-1}); four provinces in northeast and northcentral China had very large regression residual standard deviation (> 1.0 Mg ha^{-1} yr^{-1}).

For the staples crop, cash crop and three individual crop models, temporal variation trends of the regression residual varied between provinces (Appendix C). A few provinces showed an increasing trend of yield residual, while even fewer provinces showed a decreasing trend (i.e., changing from positive values to negative values) for regression residual in yield models of staples (Zhejiang and Qinghai), maize (Zhejiang), wheat (Zhejiang and Shanghai) and rice. For the cash yield model, nine provinces (e.g., Zhejiang) showed a decreasing trend. Therefore, the yield in Zhejiang was underestimated first and then overestimated in all crop models except for rice. Many provinces showed an arched trend of regression residual, had higher prediction in the mid-term and lower in the beginning and end of the 1978–2015 period. However, some provinces showed special characteristics: Shandong had higher prediction first and then lower prediction in staples yield (Figure C.3c); Gansu (in cash model, Figure C.7c) and Jiangxi (in rice and wheat model, Figures C.15b, 19c) had higher predicted than observed yield in most of the years; Qinghai and Liaoning had lower predicted than observed yield in the cash model during most of the years (Figure C.7c). Although the yield residual in Yunnan and Guizhou fluctuated, it had higher predicted yield than observed in the wheat model for most years (Figure C.19d). Overall, the cash yield model had higher mean values and standard deviation for the regression residual than other models. The standard deviation of yield residual was higher in north and west China for staples and three individual grain crops, while it was higher in east China for cash crops.

3.4. Discussion

3.4.1. Spatial and temporal variation of yield

The marked increase of aggregate yield in China between 1978 and 2015 (Figure 3.2a) is likely caused by improved crop varieties and increased fertilizer application. However, the average annual growth rates in Heilongjiang, Guizhou, Beijing, Qinghai and Jilin were smaller than those in other provinces. In Heilongjiang and Jilin, constraints from shallow topsoil resulting from long-term continuous cropping, severe black soil loss due to soil erosion and insufficient organic matter input led to decline in soil fertility and affected yield (Tang, 2016).

In Beijing, rapid urbanization and industrial development (rapid development of manufacturing, commercial service and animal husbandry) during this period led to fewer farmers and less input on agricultural land, compared with other provinces. The difficulty of large agricultural machines operation in rugged terrain and mountainous landscapes limited the yield increase in places like Guizhou and Qinghai (Wang et al., 2018a). For Guizhou in southwest China, insufficient solar radiation and uneven precipitation also limited yield increase (Xiao et al., 2004). High altitude, low temperature and shortage of precipitation limited yield improvement in Qinghai (Yang et al., 2019).

The average annual growth rates increased strongly in Hainan, Guangxi, Ningxia, Hebei and Shaanxi. Hainan and Guangxi were abundant in solar radiation and temperature, precipitation and capital investment, which created favorable conditions for production of fruit trees with high yield (Huang et al., 2018). Not surprisingly, promotion and application of advanced technology played a key role in increasing yield, such as introducing new variety breeding, soil-test based fertilizer recommendation and pest control (Zhao, 2010). Ningxia had favorable environmental conditions: high temperature in crop growing season, large day and night temperature difference, sufficient solar radiation with sufficient irrigation and good management practices made large improvements of yield possible (Yang, 2017). For Hebei, the cultivation of drought-tolerant varieties and crops with low water consumption alleviated the restriction of drought on yield (Zhang et al., 2020). The yield increased in Shaanxi province due to modern dry-land cultivars (Sun et al., 2014) and increasing solar radiation (Chen et al., 2017).

The average annual growth rates of cash crops were larger than those of staples crops during the study period, especially for apples and citrus. The most important reason is likely technological improvements has been stronger for cash crops since these crops have higher economic value than staples crops, which promotes the development of agricultural management and farmers' motivation (e.g., pruning, pest control, bagging, fertilization) and processing manufacturing (e.g., post-harvest processing technology, fruit juice processing technology).

The yield increase over the three effectual fertilization periods suggests that the adoption of these fertilization polices was effective. The spatial distribution pattern of yield conforms to the distribution characteristics of the Hu Huanyong Line from 1978 to 2005. This line is the divide between farming area and nomadic area in China. To the west of the line are mostly grasslands, deserts and snowy plateaus, which are not appropriate for high-yield crops. This area is sparsely populated and the economy is underdeveloped. Therefore, the yield is quite low. East of the line, especially south of the Qinling-Huaihe River, is a highly populated and economically developed area where agriculture is intensive with multiple cropping and crop yield is relatively high. The distribution pattern of staples production accords with the Hu Huanyong Line between 1998 to 2014 (Wang et al., 2018a). Although the crops considered in these studies, which only include staples crops, are different from those in our study, the distribution pattern is similar. Zhu et al. (2012) investigated crop straw production, such as grain, beans, tuber crops, cotton and cane in 2009, and found that the spatial distribution pattern was well in line with the 400 mm rainfall separation line, and close to the Hu Huanyong Line. Since 2006, the difference in yield characteristics on either side of the Hu Huanyong Line has become less distinctive. Instead, the yield tended overall to be strongly correlated with the GDP, except for Sichuan (high GDP with low yield), Hainan and Guangxi (low GDP with high yield). The increased correlation between yield and GDP since 2006 indicates that economic factors became more important for crop yield.

The spatial pattern of the aggregate yield was more similar to that of cash yield than that of staples yield. This may be explained from the much higher yield observed in cash crops than in staples crops. Surprisingly, the spatial pattern of staples yields was quite different among the three main grain crops. For example, maize in north China had much higher yield than in south China (Xu et al., 2017), while single season and middle rice in northeast and central China had higher yield than early and late rice in other regions (Ding et al., 2018). Winter wheat with much higher yield is mainly grown in central China, while spring wheat with shorter growing season and lower yields is mostly grown in other regions (Liu et al., 2011).

3.4.2. Yield models and main explanatory variables

The aggregate yield prediction model had a high accuracy (R-square 0.95) since the explanatory variables were comprehensive and included topography, crop types, soil covariates, economic variables and AMP. For the staples and cash yield models, the R-square was lower and it was even lower for the individual crop models. A possible explanation for the low R-square for rice could be that rice is a mix of many different varieties (early rice, middle rice, late rice, and single-season rice) (Xu et al., 2016b; Ding et al., 2020). The main reason is that aggregate yield strongly depends on crop type (see Figure 3.7a). For obvious reasons crop type explains much less of the variation in crop categories and none of the variation in individual crops.

Miao et al. (2006) indicated that soil, landscape and hybrid factors could explain 68% of the observed yield variation at field level. Li et al. (2016a) indicated that the main factors affecting temporal variation of food production are economic factors and agricultural technology application; while the main factors that influence spatial variation of crop production are climate, topography, water availability. In this chapter, economy, AMP and soil covariates were important for all crops. For economic covariates, either farmer income or GDP had a positive and important influence on models. There is no doubt that economic development is beneficial to agricultural production. Crop types were more important for cash crops than for staples crops, since the yield variation was larger among different cash crops. The contribution of topography was less than 10% of the R-square in all the models. This indicates that crop yield relies less on topography, because it affects crop yield indirectly by influencing soil properties (Jiang & Thelen, 2004).

Many studies indicated that for the main cropping regions of China, major crop yields are significantly related to climate in the growing season (Tao et al., 2008; Iizumi & Ramankutty, 2016). In this chapter, climate factors were not selected in the aggregate yield model and in several individual crop models. This is likely because climate is correlated with other explanatory variables and dropped out during model selection because other covariates represented the effect of climate on yield. This shows that the results of the multiple linear regression models should be interpreted with care, because there is no doubt that climate does influence yield. In the staples and cash yield models, climate was included but not that important. Again, this may be because the effect of climate was represented by other covariates.

Agricultural management practice had a positive effect on aggregate yield model of all crops, while large and medium-sized agricultural tractors had a negative contribution. This negative contribution is counter-intuitive and indeed the Pearson correlation between large and medium-sized agricultural tractors and yield is positive. The negative contribution in the model is explained by cross-correlations with other explanatory variables. Large machines are more popular in north China than in south China (Table B.3), while lower yield is produced in north China than south China, due to unfavorable other yield factors. This explains that the modelled influence of large and medium-sized agricultural tractors on yield was influenced by other explanatory variables in the model.

Crop type was the most important factor in the aggregate yield model. High-yield crops such as vegetables, fruits, melons and sugar crops had a positive influence on yield, while low-yield crops, such as beans and tea, had a negative influence. This may be explained by the fact that the yield of vegetables, fruits, melons and sugar crops was reported in fresh weight. This leads to larger model contributions of crops that have relatively high fresh weight, while those that have a low fresh weight than other crops may get a lower or even negative contribution. Farmer income was the second most important explanatory variable for yield variability. Rich farmers can invest more in management practices, and this pays off in a higher yield. For the least important factor topography, the local upslope curvature had a negative correlation with yield, but it had a positive regression coefficient. The negative correlation is likely because topography restricts the yield in mountainous regions (such as Sichuan, Yunnan and Guizhou province), while the positive partial regression coefficient may result from the fact that some mountain regions (such as Guangxi, Zhejiang and Fujian) also have more high-yield crops: their planting area of vegetables and fruit accounted for more than 30% of the total planting area (Table B.3).

The six soil types included in the aggregate yield model had an R-square contribution of 8%. On average, their area accounted for 44% of province area, ranging between 26 and 71% among provinces, which means that the selected soil type factors covered a large part of the provincial areas (Figure B.2). Luvisols was the most important soil type, taking up 3% of the R-square. It accounted for 13% of the total area, and had a positive influence on yield, as these are fertile soils with high clay content and cation exchange capacity. Although cryosols occupied only 1% of the total area, it was important (2% of R-square) and produced a negative impact on yield. These soils are influenced by freeze-thaw and lack of nutrients, especially calcium and potassium, which are easily leached above the permafrost. Other soil types (cambisols, alisols, gleysols and solonchaks) were less influential (1% of R-square), occupying 20, 4, 4 and 2% of the total area. The model showed a positive contribution of gleysols, where mostly rice was grown. On the contrary, cambisols, solonchaks and alisols exerted negative impacts on yield. Cambisols are soils with incipient soil formation and low weathering degrees. These soils could be managed for yield improvement with good development and improvement. Solonchaks contain high soluble salts and alisols are acidic. Both are unproductive.

Soil total phosphorus is a major determinant and indicator of soil fertility and quality, so it had a positive influence in aggregate yield model (2% of R-square). However, it showed a significant negative correlation with yield according to Pearson's correlation coefficient. This is possibly caused by the high correlation with the planting area index of high-yield crops (vegetables, melons and fruits). The negative influence of water-soluble phosphorus may be caused by the application of large amounts of inorganic fertilizers, which lead to soil acidification (Liang et al., 2013). This conformed by its significant positive correlation with exchangeable acidity (2% of R-square), because higher exchangeable acidity is detrimental to nutrient uptake and plant growth. Although silt content (1% of R-square) had a negative impact on yield in the model, it had a positive correlation with yield. This is likely to be caused by a significant correlation with total carbon. Obviously, total carbon (1% of R-square) influenced yield positively, because it improves soil fertility and water retention, and, ultimately, to maintaining and increasing crop production (Wang et al., 2015). A very interesting result is that increasing histosol probability is accompanied by lower crop yield. This result looks counter-intuitive, but it is sensible if we consider the formation of histosol. Histosol are found in places where organic matter accumulation is greater than mineralization, i.e., in poorly drained soils (Lucas, 1982). So the negative influence of histosol probability on yield was likely due to poor drainage (Masuda, 2016). In this chapter, we could only use static soil covariates as explanatory variables because maps of the space-time distribution of soil type and soil properties were not generally available. In future studies, the model performance might improve if space-time soil maps become available and soil temporal variation is included in the modeling.

Note that we aggregated covariates over all cells within a province. Alternatively, we might have aggregated only over cells that have agriculture. This could be investigated in future research, and the effect of it on the importance of covariates analyzed as well. Particularly for topography, this might be important, because within each province there may be terrain with different topography that is unsuitable for agriculture.

3.4.3. Temporal and spatial patterns in regression residuals

There were different temporal variation trends of regression residuals for the aggregate yield model in China: increasing trends mostly occurred in north China, while south China mostly had decreasing trends. That is to say, the yield in north China was lower than expected in the first time period and higher than expected towards the end of the 1978–2015 period. This might be because China’s open gate policy first happened in south China, which started to attract new technologies investment, including high yield fertilization earlier (He et al., 2018). This effect may only be partially explained by the explanatory variables used in the regression model. Another possible reason for the north-south differences could be the temperature increase due to climate change during the time period considered in this chapter, which obviously had a greater effect in north China. Some provinces had an ‘arched’ trend with lower yield than expected in the beginning and end of the periods and higher yield than expected in the middle period. While these patterns were clear it was difficult to postulate plausible explanations for these trends. Careful analysis and more detailed modelling of yield for individual provinces might provide more insight into the causes of these patterns, which could be explained by local covariates that were not included in the national model that we developed in this chapter.

The model simulated aggregate yield well for most of the provinces (i.e., in most cases the regression residual was within 0.05 Mg ha^{-1} yr^{-1}). It is interesting to note that some provinces had a systematic positive residual, which means that these provinces produced a higher yield than predicted by the model over the 38-year period. There were also provinces with a systematic negative residual. These systematic over- and under-predictions may be caused by factors that were not included in our model, such as plagues, pests, natural hazards, market structures (such as competition for labor or impediments to market access) and farmer's management skills. For example, some provinces suffer more from natural hazards than others. A more detailed study at provincial level would be required to reveal the magnitude of these effects and determine if these could be the causes of the systematic differences between provinces.

The standard deviation of regression residuals of the aggregate yield model was higher in most provinces in east China than in west China. This might be because yield in east China is influenced more by highly variable factors, such as precipitation and economy, but again a closer look is needed to investigate these possible causes. While we could not explain the spatio-temporal patterns of the regression residuals shown in Figures 3.8 and 3.9, these results identify valuable research gaps and highlight the need for further research.

3.4.4. Strengths and limitations

The analytical results at provincial scale may support the provincial government to compare their yields with other provinces and formulate more effective policies to increase crop yield, such as high-yield fertilization, balanced fertilization and soil-test based fertilization. Many studies on yield variation of grain crops have been carried out but since these studies rarely include cash crops, these do not fully reflect the entire productivity (Simmonds et al., 2013). In this chapter, we analyzed spatial patterns and temporal trends in the aggregate yield of all crops at provincial level in order to analyze the total productivity of each province, among others using planting area of different crop types as explanatory variables. However, the fact that different crops are grown in different parts of the country and that planting area of crops varies over time, will have affected the spatio-temporal patterns of aggregate yield. We therefore also analyzed the relative importance for two other yield aggregation levels: main crop categories (staples and cash crops) and individual crop types within each of these categories. For these categories and specific crops, the effect of crop types on spatio-temporal patterns of yields is much less or absent.

To make results economically tangible, dry weight of staples crops (for beans) and fresh weight of cash crops (except for beans) were considered in this chapter. From the perspective of exploring the influence of factors on crop yield, it would be better to use dry matter of cash crops instead of fresh weight. However, the available datasets reported fresh weight and these are not easily converted to dry weight. For instance, the moisture content is highly dependent on the cultivar of cash crops. It is a challenging task to apply conversions from fresh weight to dry weight for every variety of cash crops at national scale for a 38-year period, in a case where essential conversion parameters are lacking.

Some explanatory variables that had high correlation with yield were excluded from the final model to avoid multicollinearity. This reduced model complexity without sacrificing much of the variance explained. But the decision which variables to remove and which to retain has some degree of arbitrariness, while it may have a large impact on model interpretation. This should be kept in mind when interpreting regression coefficients and relative importance of explanatory variables. There are important advantages to taking an empirical modelling approach, such as the relative ease of model building, but one must be cautious that empirical models detect correlations instead of causalities (e.g., the negative influence of soil organic carbon density in the watermelon model). Also, results are by construction limited to the explanatory variables included in the model and by correlations between covariates. For instance, we noted before that while climate was not often included in the models, this does not mean that climate has no effect on yield. Further, due to lack of data the influence of natural abiotic stress, plant disease and insect pests could not be included in this chapter. The model could perhaps be extended using predictions of natural disaster and pest diagnosis analysis, using e.g., longitudinal analysis approach (Cook et al., 2002).

The multiple linear regression analysis in this chapter was also limited to linear relationships between crop parameters and explanatory variables, and results may be suboptimal when these relationships are nonlinear (Kitchen et al., 2003). The regression residual also had different patterns for each province, which indicated that our model could not reflect all spatial complexity in data relationships.

Geographically weighted regression (GWR) may be an attractive way to extend the modelling by providing an intuitive and technically accessible tool to explore where non-stationarity is taking place, by fitting regression coefficients locally rather than on a national scale (Fotheringham et al., 2003). Huang et al. (2010) made an extended GWR model of real estate market data, composed of geographically and temporally weighted regression, by incorporating temporal effects into a GWR model. This might be an interesting approach for space-time modelling and analysis of yield data.

3.5. Conclusions

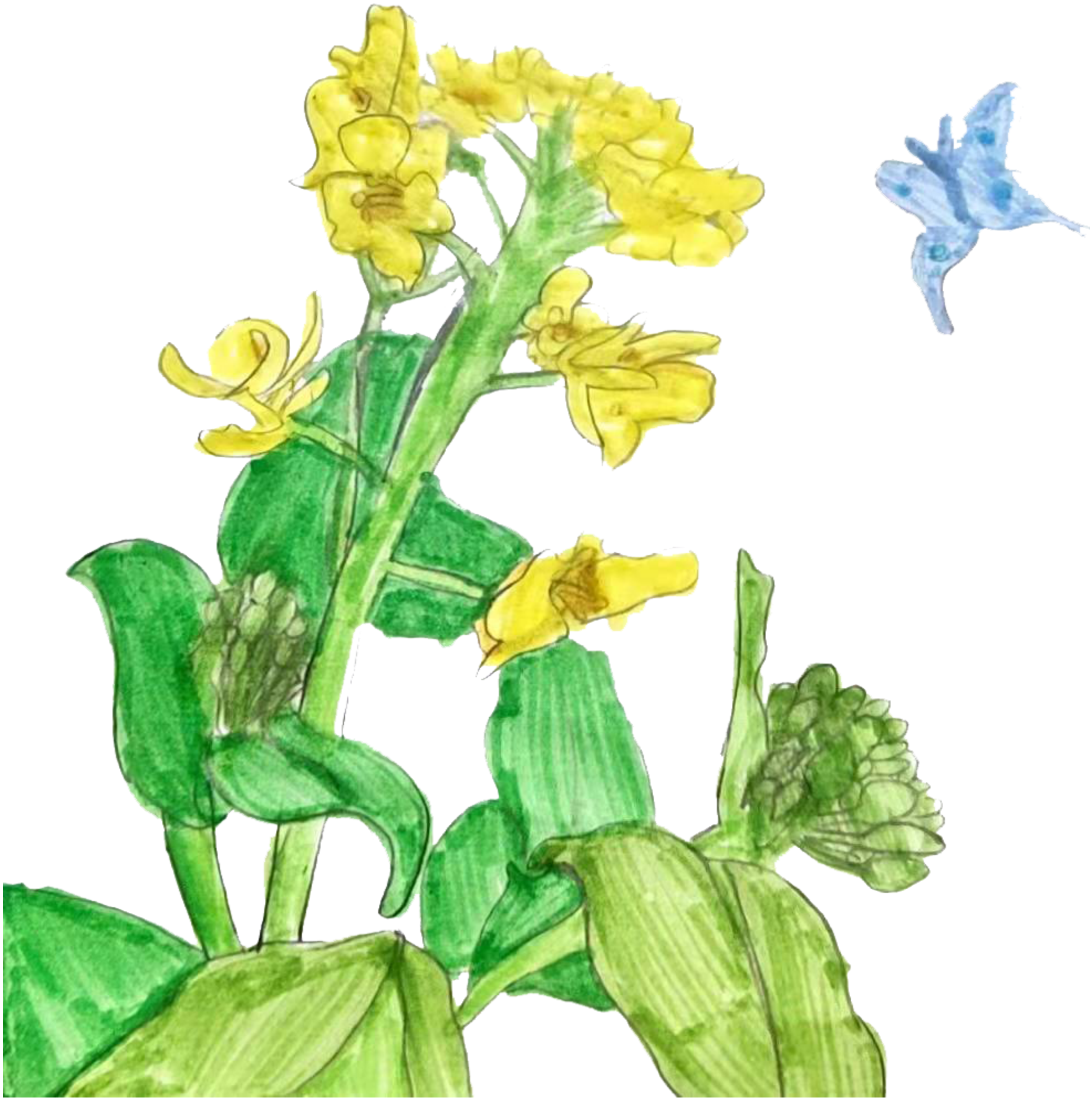
This chapter analyzed temporal and spatial variation in crop yield aggregations at provincial level in China from 1978 to 2015. Stepwise multiple linear regression was used to explore the relationships between crop yield and agricultural, environmental and economic explanatory variables. The temporal and spatial patterns of yields were different for different levels of crop aggregations. Most of the models had an R-square larger than 0.6, except for rice, potato and cotton. AMP, soil and economic covariates were the most important factors in all models. Topography had an influence on the aggregate yield model but was not included in the staples and cash model. Instead, climatic covariates were important for the staples and cash models, but not for the aggregate yield model. The model performance for the aggregate yield was different for each province in individual years and residuals of the regression model had distinct spatial and temporal patterns. Hence, a more detailed analysis of model performance and residual analysis is needed to explore the causes of these patterns. The models could not predict the impact of natural hazards, plant diseases and insect pests due to lack of data. This may be improved in future research using a combination of natural disaster prediction and pest diagnosis analysis. With the increasing food requirement and limited agricultural land resources, enhancing economic growth might be possible solutions for China to safeguard food security, in particular if this is combined with better management practices, breeding and planting technologies, and taking account of crop suitability (i.e., adaptability of crops to the local environment).

Supplementary materials

The supplementary materials, Appendices B and C, can be downloaded as Appendices A-B from the journal version of this chapter: Liu, Y., Heuvelink, G. B. M.,

Chapter 3 Analysis of spatio-temporal variation of crop yield in China using stepwise multiple linear regression

Bai, Z., He, P., Xu, X., Ding, W., & Huang, S. (2021). Analysis of spatio-temporal variation of crop yield in China using stepwise multiple linear regression. *Field Crops Research*, 264, 108098. https://doi.org/10.1016/j.fcr.2021.108098.



Chapter 4

Statistical analysis of nitrogen use efficiency in northeast China using multiple linear regression and random forest

Understanding the spatial-temporal dynamics of crop nitrogen (N) use efficiency (NUE) and the relationship with explanatory environmental variables can support land-use management and policymaking. Nevertheless, the application of statistical models for evaluating the explanatory variables of space-time variation in crop NUE is still under-researched. In this chapter, stepwise multiple linear regression (SMLR) and random forest (RF) were used to evaluate the spatial and temporal variation of NUE indicators (i.e., partial factor productivity of N (PFP_N); partial nutrient balance of N (PNB_N)) at county scale in northeast China (Heilongjiang, Liaoning and Jilin provinces) from 1990 to 2015. Explanatory variables included agricultural management practices, topography, climate, economy, soil and crop types. Results revealed that the PFP_N was higher in the northern parts and lower in the center of the northeast China and PNB_N increased from southern to northern parts during the 1990-2015 period. The NUE indicators decreased with time in most counties during the study period. The model efficiency coefficients of the SMLR and RF models were 0.44 and 0.84 for PFP_N, and 0.67 and 0.89 for PNB_N, respectively. The RF model had higher relative importance of soil and climatic covariates and lower relative importance of crop covariates compared to the SMLR model. The planting area index of vegetables and beans, soil clay content, saturated water content, enhanced vegetation index in November & December, soil bulk density, and annual minimum temperature were the main explanatory variables for both NUE indicators. This is the first study to show the quantitative relative importance of explanatory variables for NUE at a county level in northeast China using RF and SMLR. This novel study gives reference measurements to improve crops NUE which is one of the most effective means of managing N for sustainable development, ensuring food security, alleviating environmental degradation and increasing farmer's profitability.

Based on:

Liu, Y., Heuvelink, G. B. M., Bai, Z., He, P., Jiang, R., Huang, S. & Xu, X. (2022). Statistical analysis of nitrogen use efficiency in Northeast China using multiple linear regression and random forest. *Journal of Integrative Agriculture* (accepted).

4.1. Introduction

Estimation of nitrogen (N) use efficiency (NUE) is important for evaluating the performance of N use in agricultural systems and NUE is often used as a management tool for determining agronomic and environmental sustainability (Omara et al., 2019). There are different ways to express NUE, of which the recovery and agronomic efficiency of N use require field trial data while the partial factor productivity of N (PFP_N) and partial nutrient balance of N (PNB_N) can be obtained from survey data (Dobermann, 2007; Quan et al., 2021). In this chapter, we used PFP_N (yield divided by N input, unit: kg kg^{-1} yr^{-1}) and PNB_N (N removal divided by input, unit: kg kg^{-1} yr^{-1}) to analyze long-term NUE trends. However, NUE is also influenced by different N input sources. At a global scale, considering total N inputs to cropland (chemical fertilizer, manure, biologically fixed N, and N deposition), PFP_N decreased from 68 kg kg^{-1} in 1961 to 45 kg kg^{-1} in 1980, followed by a stabilization at around 47 kg kg^{-1} during the next 30 years (Lassaletta et al., 2014); while PNB_N remained relatively stable at 0.50-0.55 kg kg^{-1} between 1987 and 2006 when only chemical fertilizer was considered as an N input (Brentrup & Pallière, 2010), thereafter decreasing to 0.42 kg kg^{-1} in 2010 when total N inputs were considered (Zhang et al., 2015). However, China had much lower PFP_N (26 kg kg^{-1} in 2009, Lassaletta et al., 2014) and PNB_N (0.25 kg kg^{-1} in 2010, Zhang et al., 2015) values because it has been experiencing a massive N loss from agriculture to the environment, via ammonium volatilization, nitrate leaching and nitrification/denitrification. Such losses are not only an unnecessary wastage of natural resources (Galloway et al., 2008), but also pose serious potential ‘downstream’ pressures on the aquatic environment (Chien et al., 2009) and affect air quality (Xu et al., 2016a). Given this situation, improving the NUE of crops is one of the most effective means of managing N for sustainable development, ensuring food security, alleviating environmental degradation and increasing farmer’s profitability.

In China, determination of crop NUE is complex and has substantial spatial and temporal variation. At a national level, the PFP_N decreased from 32 kg kg^{-1} in 1978 to 27 kg kg^{-1} in 1995, after which it increased to 38 kg kg^{-1} in 2015 (**Chapter 2**). It was higher in eastern and southern China than in central and western China. PNB_N decreased from 0.53 in 1978 to 0.38 kg kg^{-1} in 2000, after which it remained constant until 2015. It was low in southern China and moderate in other regions (**Chapter 2**). Coarse-scale (Zhang et al., 2015) and short-term studies (Yousaf et al., 2016) provide only a limited understanding of the spatial-temporal dynamics of NUE. It would therefore be useful to quantify NUE using a fine spatial resolution and long-term data set. However, NUE studies at farm or field scale (Liang et al., 2018) may be too detailed and too limited to obtain regional overviews that would allow formulation of region-wide agricultural policies. Comprehensive county scale analyses may be feasible for large regions and provide insights that are obscured by provincial scale analyses (Lu et al., 2019a).

Northeast China has sufficient irrigation water (Pan et al., 2020) and higher NUE than other regions (**Chapter 2**). Thus, examining space-time trends and explanatory variables of crop NUE in this region can help explain when, where and why crops have reached a NUE peak, identify remaining potentials of NUE improvement, and avoid deterioration caused by environmental pollution (Lu et al., 2019a). The region is vast with various soil characteristics, different crops, complex topography and diverse climate. Therefore, many explanatory variables can markedly influence NUE in the region. For example, the maize PFP_N varies with soil type: it was about 60 kg kg^{-1} yr^{-1} in chernozem soils (i.e., black soils) of Heilongjiang and Liaoning provinces, and roughly 45 kg kg^{-1} yr^{-1} and 30 kg kg^{-1} yr^{-1} in cambisol soils (i.e., cinnamon and fluvo-aquic soils) of Shanxi and Hebei provinces during 2010-2012 (Xu et al., 2014b). The PNB_N variation is large among crops: at the global level, it is more than 0.80 kg kg^{-1} for soybean, and less than 0.14 kg kg^{-1} for fruits and vegetables and other crops (Zhang et al., 2015). NUE variation also likely depends on a broad spectrum of soil fertility and different capabilities of N uptake and yield among crop varieties (Lu et al., 2019a). Crop genotypes vary in removing N (Fageria & Baligar, 2005), which can result in NUE variation. Many causes of NUE variation, including socioeconomic variables (e.g., farmer income and crop price), agricultural management practices (AMP) (e.g., irrigation area, nutrient management measures, agricultural machinery) and environmental variables (e.g., soil, climate, topography) may explain NUE variation across counties and over time.

Most methods or models only evaluate NUE from an agronomic perspective, by optimizing nutrient management strategies with crop yield models (Xu et al., 2013), precision agriculture (Diacono et al., 2012), site-specific nutrient management (Dobermann et al., 2003), 4R nutrient stewardship (i.e., right source, right rate, right time and right place) (Johnston & Bruulsema, 2014), Nutrient Expert Systems (Xu et al., 2014b) and soil testing (He et al., 2009). More advanced statistical methods need to be applied when exploring the factors influencing NUE. Stepwise multiple linear regression (SMLR) and random forest (RF) are both intuitive, meaningful, and informative methods for exploring explanatory variables and quantifying their relative importance in explaining NUE variability from a large data set. SMLR produces explicit equations which detail the importance of every explanatory variable through standardized regression coefficients. In contrast, RF, as a black-box model, does not lend itself to specific mathematical equations, but usually has a higher accuracy than SMLR (Sakamoto, 2020). SMLR and RF models have been applied in agriculture, such as for predicting sugarcane yield (Everingham et al., 2016), rice aboveground biomass (Cen et al., 2019), and maize yield and NUE (Li et al., 2020a). Nevertheless, SMLR and RF modelling for evaluating the explanatory variables of space-time variation in crop NUE is still under-researched.

The objectives of this chapter were to: (1) explore and interpret space-time patterns of NUE indicators (PFP_N and PNB_N) at county scale in northeast China from 1990 to 2015; (2) construct NUE prediction models at county scale using SMLR and RF models and evaluate their performance; (3) compare the relative importance of explanatory variables in SMLR and RF models.

4.2. Materials and methods

4.2.1. Study area and time

In this chapter, we analyzed NUE at a county scale in Heilongjiang, Jilin, and Liaoning provinces in northeast China. We chose this region because it has highly fertile black soil (Ren et al., 2011) and it is known as the 'breadbasket' of China. The region is located in the large Great Plains of China and has a high concentration of cultivated land, which is conducive to large-scale mechanized operation. The study area includes 183 counties (79 in Heilongjiang, 47 in Jilin and 57 in Liaoning) (38°46′-53°33′N, 118°53′-135°05′E). More county information is provided in the Appendix D. The study area is characterized as a temperate and cold temperate continental monsoon climate: it has long, cold winters and short, cool summers. Annual mean temperature ranges from −5 to 10°C, with mean winter temperatures ranging from −28 to −2°C and mean summer temperatures ranging from 15 to 25°C. Annual precipitation ranges from 400 mm in the northeast to 1000 mm in the southeast. Monthly precipitation, temperature and land cover in 2015 are shown in Figure 4.1 (Climatic Research Unit, 2015; ESA, 2015). This region grows one crop per year. To explore time-related changes in spatial patterns, the study period (1990-2015) was divided into five sub-periods of five years each.

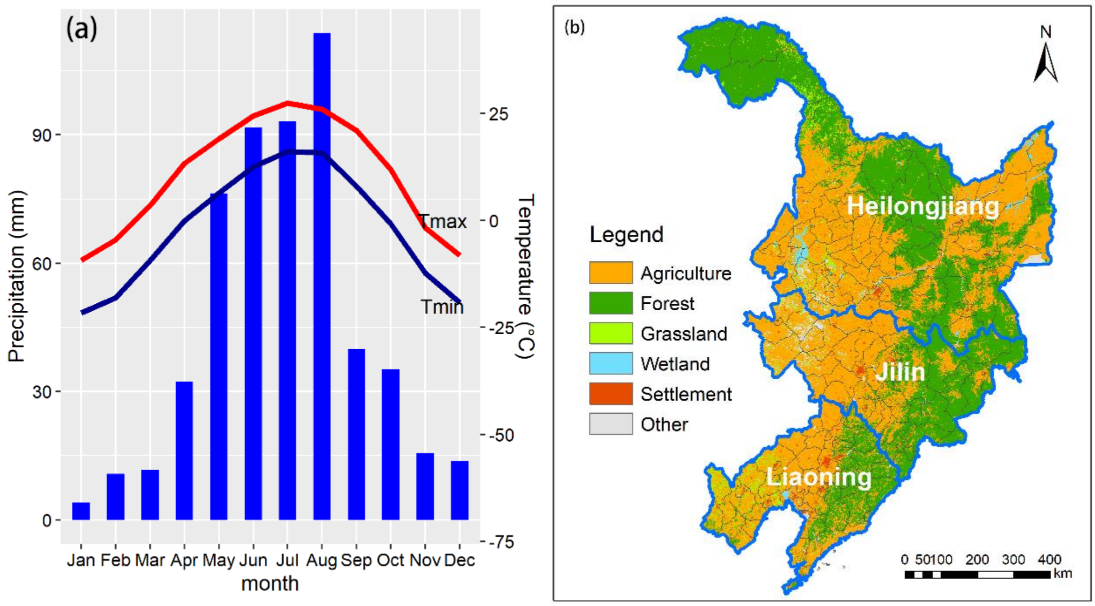


Figure 4.1: Monthly precipitation, average daily maximum (Tmax) and minimum temperature (Tmin) (a) and land cover (b) in northeast China in 2015.

4.2.2. NUE data and covariates

Two indicators of NUE long-term trends were used: PFP_N and PNB_N (Dobermann, 2007).

$$PFP_N = \frac{\text{Yield}}{N_{\text{input}}} \tag{4.1}$$

$$PNB_N = \frac{N_{\text{crop\_removal}}}{N_{\text{input}}} \tag{4.2}$$

Here, Yield refers to the economic yield of crop with applied; $N_{crop\_removal}$ is the total removal in aboveground crop biomass with applied and N_{input} is the total amount of applied to the field from different sources (chemical fertilizer, manure, cake fertilizer and straw).

The total number of observations for each indicator was 4687 (i.e., the product of years and counties). Economic yield and planting area were obtained for ten crops: cereals, beans, tubers, oil crop, sugar crop, fiber (including cotton), tobacco, vegetables, melons and fruits. Crop yields of the main harvested products were measured as fresh weight for vegetables, melons and fruits, dry matter for others. Chemical N fertilizer sources were single N fertilizer and compound N fertilizer. Manure fertilizers came from cattle, pigs, sheep, poultry, horse, donkey, mule and humans. More details about computing NUE indicators are given in He et al. (2018). The county data for NUE calculation were obtained from the National Bureau of Statistics of China (1991-2016) of northeast China.

Explanatory variables (i.e., covariates) contained topography (6 covariates), crop covariates (enhanced vegetation index (EVI, 6 covariates), crop types (10 covariates, as listed above)), soil types (17 covariates, which were aggregated from 118 soil sub-types based on reference groups of the World Reference Base), soil properties (34 covariates), climate (12 covariates), economy (1 covariates) and AMP (2 covariates). These were obtained from the Climatic Research Unit (1990-2015), Harris et al. (2014), the National Bureau of Statistics of China (1991-2016) and published articles (Shangguan et al., 2014; Hengl et al., 2017; Poggio et al., 2020). More details about the explanatory variables including abbreviations can be found in Table D.1. All explanatory variables available in raster maps were aggregated to county scale by taking the spatial average over agricultural land (obtained by land cover maps during 1992-2015 from the European Space Agency Climate Change Initiative Land Cover (ESA, 2015)). Most covariates were dynamic and varied between years, except for the enhanced vegetation index (for which we took long-term averages for each of the two-month time periods within a year, thus capturing seasonal variation), topography and soil covariates, which were considered static because these covariates had negligible temporal variation during the period studied. Crop and soil types were transformed to continuous-numerical variables by computing area proportions. Climatic variables (excluding night and daytime temperature, which were only available as annual values) were computed for the crop growing season (March-October) (Ray et al., 2015). A total of 88 explanatory variables were employed to analyze and explain the spatial-temporal variation of PFP_N and PNB_N.

The pre-processing of the initial data set before computing NUE indicators and building models was as follows: (a) filling missing data and covariates in absent years, either by consulting statistical yearbook/published literature, supplementing missing values by mean values of adjacent years (if the time gap was relatively small), or by using data from neighboring counties. For covariates, we used the poly2nb function of the spdep package (v1.1-5) to fill gaps; (b) we flagged suspicious values (i.e., outliers that are 10 times larger or smaller than that of adjacent years) and double-checked the reference data and compared them with other available data.

4.2.3. SMLR and RF models

The SMLR and RF models were used to help to summarize and understand spatial and temporal patterns of NUE. Although they cannot detect causal influences (only correlations instead), they provide insight into the causes of the variation by quantifying the importance of explanatory variables.

SMLR model

Multiple linear regression models the linear relation between a dependent variable (i.e., NUE indicator) and explanatory variables. Stepwise regression was used to select explanatory variables automatically and simplify the initial model that used all covariates. We used the Bayesian information criterion (BIC) and "both" direction for variables selection, as implemented in the stepAIC function in the MASS package, v7.3-53 of the R language for statistical computing (R Core Team, 2021). A common problem of SMLR is multicollinearity (highly linearly correlated explanatory variables), which can be detected using the variance inflation factor (VIF). A rule of thumb is that the VIFs of all covariates used in the model should be smaller than 10 to avoid multicollinearity (Kutner et al., 2004). If the highest VIF was greater than 10 we deleted the corresponding explanatory variable and refitted the model. This procedure was repeated until all VIFs were below 10. Next, we used the confint function in the stats package (v3.6.2) to obtain confidence intervals of the remaining explanatory variables coefficients. Insignificant variables (i.e., *P-value* > 0.05) were deleted from the model, again in a stepwise manner. More details about the model building process are given in Figure D.1.

Bivariate correlations between covariates and dependent variables were also computed and examined since this was helpful to understand the correlation of the data and interpret results. Corresponding scatter plots were plotted in a single figure using the chart.Correlation function of the PerformanceAnalytics package, v2.0.4. Multiple linear regression assumes that the regression residuals are independent and identically distributed Gaussian variables. The normality assumption was evaluated by computing the studentized residual, the histogram of which is generated in the residplot function (Kabacoff, 2011), which superimposes a normal curve, kernel density curve and rug plot, visually reflecting how close model errors are to a normal distribution. To make the regression coefficients of the final model more interpretable, all explanatory variables were standardized to zero mean and unit standard deviation prior to model calibration, using the scale function in the base package of R. As a result, the regression coefficients describe the expected change in the dependent variable for a standard deviation change in an explanatory variable, holding the other explanatory variables constant. This is the simplest method to obtain the relative importance of explanatory variables. In this chapter, we used the method of relative weights (abbreviated as relweights) to obtain the relative importance of explanatory variables (relweights function, Kabacoff, 2011). This method closely approximates the average increase in the model efficiency coefficient (MEC) (see Equation 4.4 in Section 4.2.4) obtained by adding an explanatory variable across all possible sub-models. Another popular method is the averaging over orderings proposed by Lindeman et al. (1980), abbreviated as LMG, which was calculated using the calc.relimp function of R package relaimpo v2.2-3 (Grömping, 2007).

RF model

The RF algorithm, developed by Breiman (2001), tends to have higher accuracy compared with other approaches (Hengl et al., 2015). It belongs to the family of ensemble machine learning algorithms that predicts a dependent variable (in this case a NUE indicator) from a set of explanatory variables (training data) selected randomly using a bootstrapping technique (sampling with replacement). In this way, it generates multiple trees in the training procedure without pruning (reducing the number of trees) and aggregates the results. Each tree in the forest is less correlated with other trees because only random subsets of covariates are included in each tree, which increases accuracy (Gislason et al., 2006). For each tree, there are two datasets: the "in-bag" data for model training that contained two thirds of the features data set and the "out-of-bag (OOB)" data to verify model error that contained the remaining one third of the features data (Fraiwan et al., 2012). Three user-defined parameters are required in the RF model: the number of trees (*ntree,* 500, following a rule-of-thumb), the number of variables as explanatory variables to grow in each tree (*mtry,* 32, 1/3 of explanatory variables), and the minimum size of terminal nodes (*node size, 5*) (Liu et al., 2015a). Default values of these parameters were used in this chapter. The *mtry* parameter determines the strength of each tree and the correlation among trees in the model, meaning that the strengths of the individual trees and the correlations among all trees increase as *mtry* increases (Prasad et al., 2006; Peters et al., 2008).

The complexity of the RF model means that it has a 'black box' character, although variable importance ranking is possible with RF and provides valuable information about what drives the model (Breiman, 2001). Two variable importance methods in the randomForest package v4.6-14 are the mean increase in mean square error (unscaled %IncMSE) and total decrease in node impurities (IncNodePurity). For convenience of presentation, we normalized the variable importance of the most important explanatory variables of the RF models (using an equal number of variables as used in the corresponding SMLR model) to sum to 100%. Computations of the RF model can be costly but this has been greatly improved with the development of the “RANdom forest GEneRator" (ranger v0.12.1) package, which has good scaling properties with the number of features, samples, trees, and features tried for splitting (Wright & Ziegler, 2017). The ranger package uses unscaled permutation and node impurity methods for ranking variable importance (Nicodemus et al., 2010).

4.2.4. Validation/Accuracy assessment

Cross-validation is a popular approach for assessing and selecting predictive models. In this chapter, we used 10-fold cross-validation by the shrinkage function (Kabacoff, 2011) for SMLR models and the RFcv function in the spm package v1.2.0 for RF models. In 10-fold cross-validation, the sample is divided into 10 subsamples. Each of the 10 subsamples serves as a hold-out group and the combined observations from the remaining 9 subsamples serves as the training group. The performance for the 10 prediction equations applied to the 10 hold-out samples are recorded and then averaged (Kabacoff, 2011). The model validation metrics were the mean error (ME), root mean squared error (RMSE), normalized RMSE (nRMSE), model efficiency coefficient (MEC) and Lin's concordance correlation coefficient (LCCC) (Lin, 1989). These are computed as follows:

$$\mathrm{ME} = \frac{1}{n}\sum_{i=1}^{n}(P_i - O_i) \tag{4.3}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_i - O_i)^2} \tag{4.4}$$

$$\mathrm{nRMSE} = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_i - O_i)^2}}{\bar{O}} \tag{4.5}$$

$$\mathrm{MEC} = 1 - \frac{\sum_{i=1}^{n}(P_i - O_i)^2}{\sum_{i=1}^{n}(O_i - \bar{O})^2} \tag{4.6}$$

$$\mathrm{LCCC} = \frac{2\sum_{i=1}^{n}(O_i - \bar{O})(P_i - \bar{P})}{\sum_{i=1}^{n}(O_i - \bar{O})^2 + \sum_{i=1}^{n}(P_i - \bar{P})^2 + n\cdot(\bar{O} - \bar{P})^2} \tag{4.7}$$

where n is the number of observations; P_i and O_i are the predicted and observed NUE values for observation i, respectively, and $\bar{P}$ and $\bar{O}$ are the means of the predicted and observed NUE values.

We consider nRMSE≤15% as “good” agreement; 15–30% as “moderate” agreement; and ≥30% as “poor” agreement (Liu et al., 2013b). If MEC is over 0.90, this indicates good agreement between observations and predictions (Aghdaei et al., 2017). According to Ichami et al. (2019) a model with a MEC higher than 0.3 could still be a useful model in this domain of science. An LCCC smaller than 0.90 is considered as poor agreement, 0.90–0.95 as moderate agreement, 0.95–0.99

as substantial agreement, and larger than 0.99 as almost perfect agreement (de Beaufort et al., 2017).

4.3. Results

4.3.1. Spatial and temporal variations of NUE indicators from 1990 to 2015

Spatial and temporal variations of PFP_N

In order to depict spatial patterns of county-level PFP_N in northeast China from 1990 to 2015, the PFP_N were grouped into five categories: very low (< 25 kg kg^{-1} yr^{-1}), low (25-35 kg kg^{-1} yr^{-1}), moderate (35-45 kg kg^{-1} yr^{-1}), high (45-60 kg kg^{-1} yr^{-1}), and very high (> 60 kg kg^{-1} yr^{-1}) (Figure 4.2). Overall, most counties in Heilongjiang province had higher PFP_N and in Jilin province had lower PFP_N values. Over the first four periods, more counties had very low PFP_N (increasing from 32 to 43 counties) and moderate PFP_N values (increasing from 37 to 52 counties), and fewer counties had high PFP_N (decreasing from 34 to 26 counties) and very high PFP_N values (decreasing from 32 counties to 14 counties). In the fifth period, fewer counties (23 counties; within Jilin province, southeast and northwest of Liaoning province) had PFP_N values at a very low level, and more counties (41 counties; within Heilongjiang province, middle and southwest of Liaoning province) had PFP_N values at a high level.

The temporal variation of PFP_N varied between counties in northeast China (Figures 4.3a and 4.4a). On a county scale, we selected the five biggest-area counties in each province for display in Figure 4.4. Their temporal variation trends from 1990 to 2015 fluctuated more than the provincial trend. Overall, they can be divided into two types (Figure 4.4a): (1) decrease in most counties (the average annual growth rate ranged from -6.1 to -0.1%); (2) increase in Huma county of Heilongjiang province, Beipiao city of Liaoning province and Antu county of Jilin province (the average annual growth rate ranged from 0.2 to 2.3%).

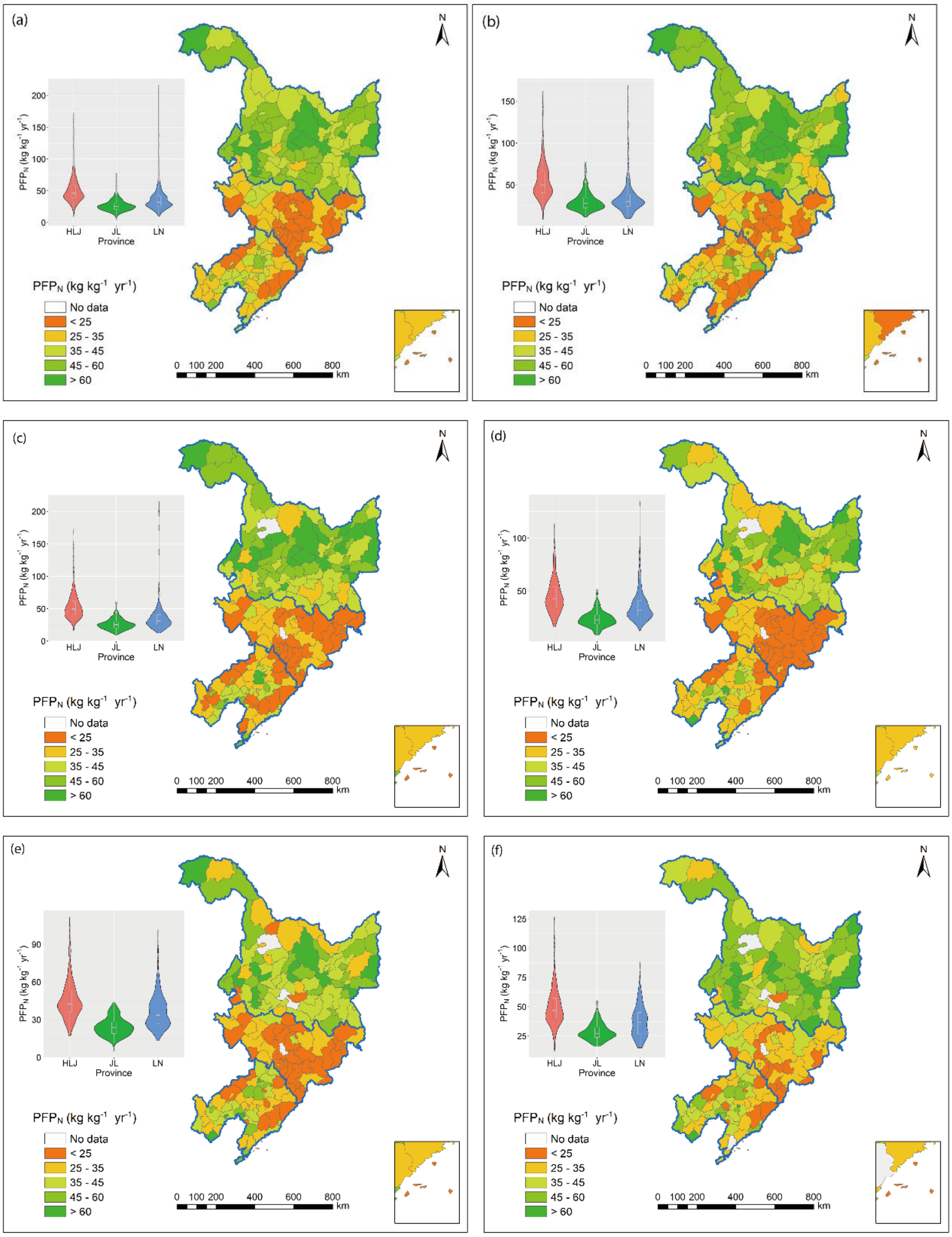


Figure 4.2: Spatial distribution of county-average PFPN in northeast China for six time periods: (a) 1990-2015, (b) 1990-1995, (c) 1996-2000, (d) 2001-2005, (e) 2006-2010, (f) 2011-2015. Violin plots show the distribution of county data for each province: HLJ, Heilongjiang; JL, Jilin; LN, Liaoning.

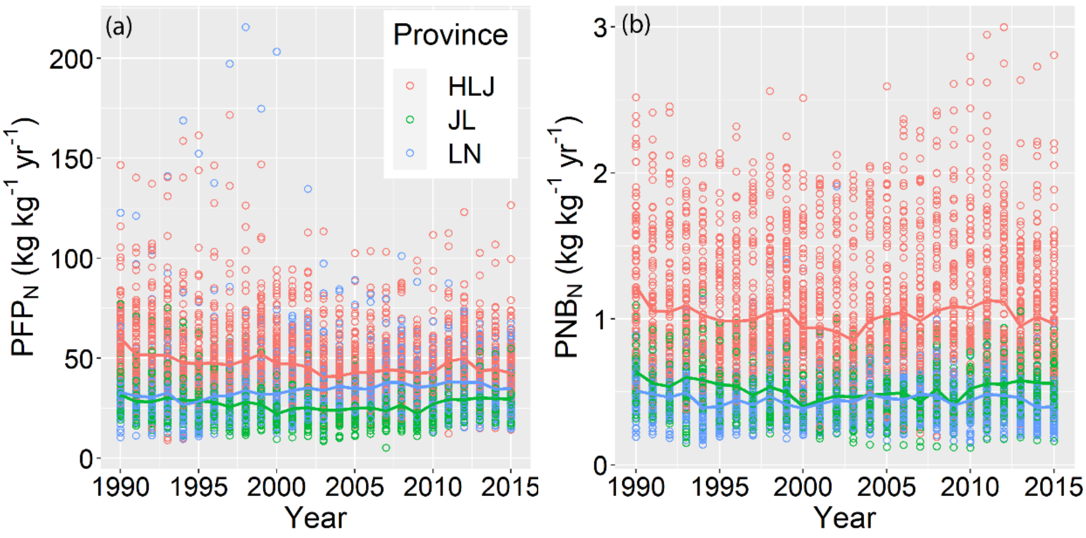


Figure 4.3: Temporal variation of PFP_N (a) and PNB_N (b) in northeast China from 1990 to 2015. The solid red, green and blue lines represent the aggregated provincial NUE indicators of Heilongjiang (HLJ), Jilin (JL) and Liaoning (LN) provinces, respectively, which was computed by averaging over counties (sum of yields/N of crops removal in each county divided by the sum of N input). County NUE indicators are shown by open circles with color indicating the province in which the county lies.

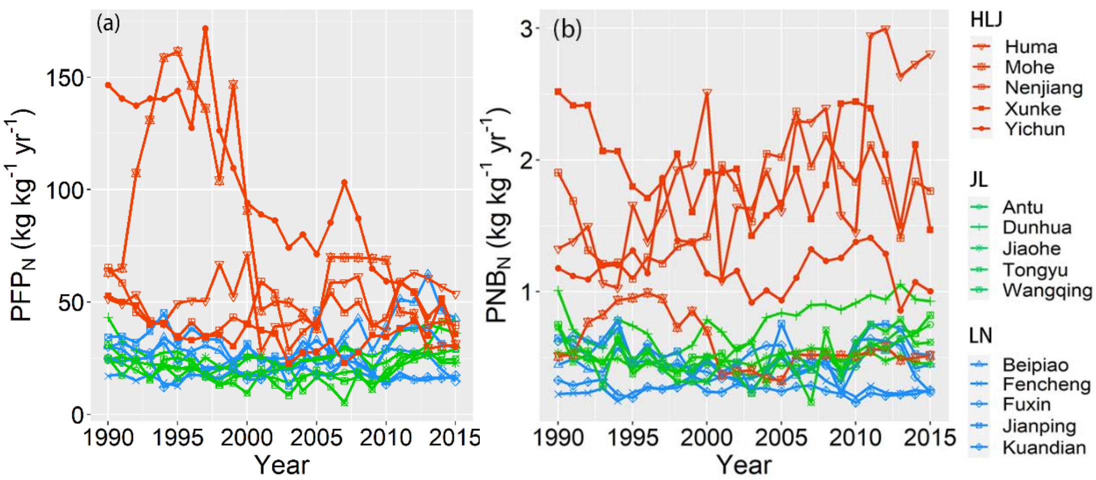


Figure 4.4: Temporal variation of PFP_N (a) and PNB_N (b) in fifteen counties in northeast China from 1990 to 2015. Note: Mohe city included Mohe county, Xinlin district and Huzhong district; Fencheng, Fuxin and Kuandian are Manchu Autonomous counties; the red, green and blue color represents the province of Heilongjiang (HLJ), Jilin (JL) and Liaoning (LN), respectively; the counties in the scatter plot were the five biggest-area counties of each province.

Spatial and temporal variations of PNB_N

Overall, the PNB_N increased from south to north in the study area in the period 1990-2015 (Figure 4.5a). This might be related to the spatial patterns of crop removal N and N input: these were lower in the north of northeast China (Figures D.2b and c. PNB_N was divided into five classes: very low (< 0.45 kg kg^{-1} yr^{-1}), low (0.45-0.65 kg kg^{-1} yr^{-1}), moderate (0.65-1.00 kg kg^{-1} yr^{-1}), high (1.00-1.50 kg kg^{-1} yr^{-1}), and very high (> 1.50 kg kg^{-1} yr^{-1}). More counties had very low and low PNB_N, and fewer counties had high and very high PNB_N from 1990 to 2005. In contrast, from 2006 to 2015, fewer counties had PNB_N at a very low level, and more counties had PNB_N at a moderate and high level.

The trend of PNB_N variation from 1990 to 2015 was different among the counties (the five biggest-area counties from each province were selected and plotted in Figure 4.4b): PNB_N decreased in most counties (the average annual growth rate ranged from -3.8 to -0.3%); it had no notable change in Beipiao city of Liaoning province and Mohe city of Heilongjiang province (~0.0% of the average annual growth rate); and increased in Fengcheng county of Liaoning province, Wangqing and Antu counties of Jilin province, and Huma county of Heilongjiang province (the average annual growth rate ranged from 0.3 to 3.0%).Prediction models and performance evaluation

SMLR model

Summary statistics of the NUE indicators and explanatory variables included in the final SMLR model are listed in Table 4.1. The PFP_N and PNB_N values exhibited a slightly skewed distribution, with skewness and kurtosis coefficients of 2.01 and 11.52 for PFP_N, and 1.38 and 4.65 for PNB_N, respectively; and mean values of 40 and 0.75 kg kg^{-1} yr^{-1} (ranging from 5 to 216 kg kg^{-1} yr^{-1} for PFP_N, and from 0.12 to 3.00 kg kg^{-1} yr^{-1} for PNB_N), respectively, which were higher than their standard deviations (Std., 19 and 0.46 kg kg^{-1} yr^{-1} respectively).

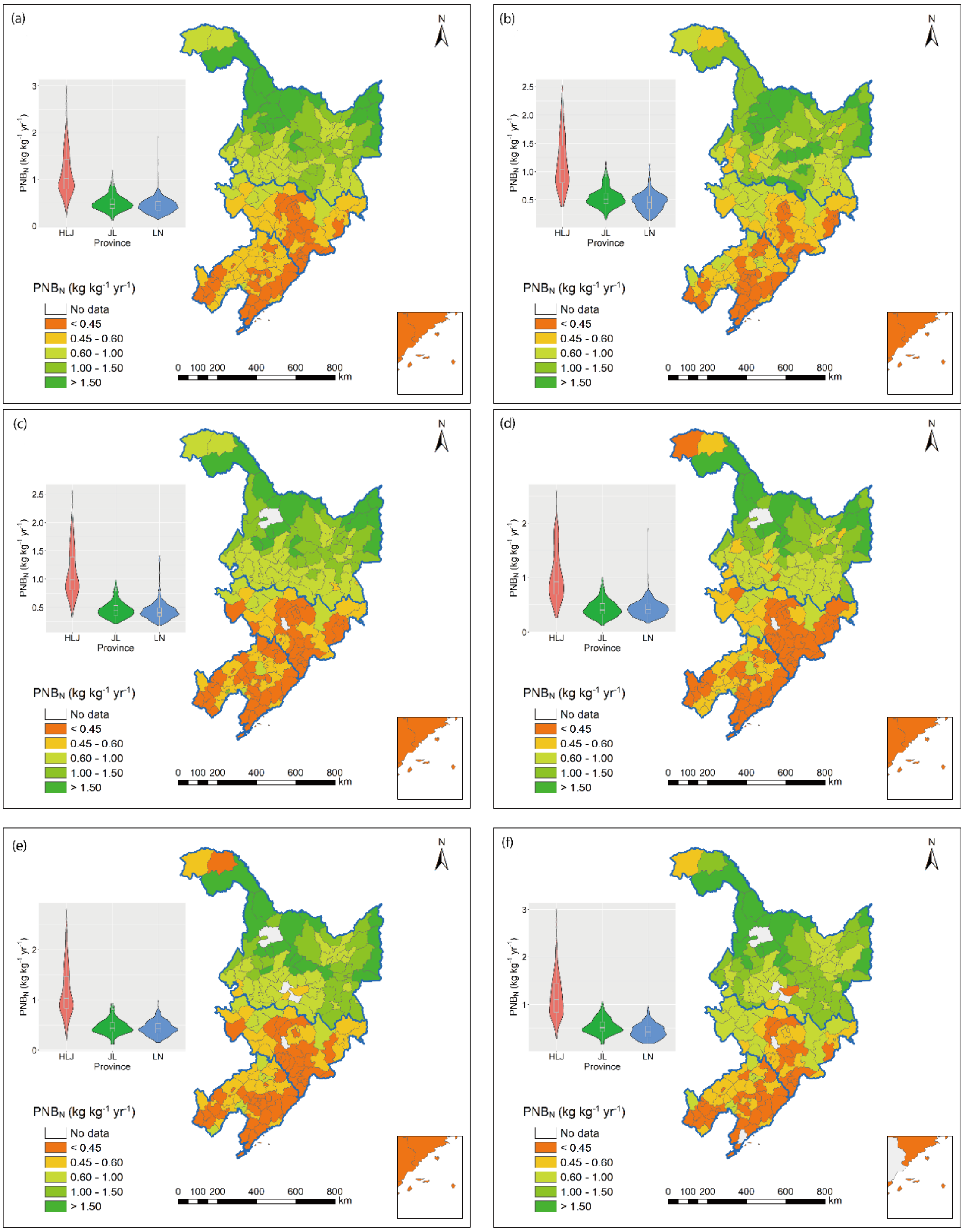


Figure 4.5: Spatial distribution of county-average average PNB$_N$ in northeast China for six time periods: (a) 1990-2015, (b) 1990-1995, (c) 1996-2000, (d) 2001-2005, (e) 2006-2010, (f) 2011-2015. Violin plots show the distribution of county data for each province: HLJ, Heilongjiang; JL, Jilin; LN, Liaoning.

Table 4.1: Descriptive statistics of nitrogen use efficiency indicators and explanatory variables in the stepwise multiple linear regression model.

Classification	Property[1)]	Min. [2)]	1st.Qu. [3)]	Median	Mean	3rd.Qu. [4)]	Max. [5)]	SD[6)]	Skewness	Kurtosis
Dependent variable	PFP_N	5	26	36	40	48	216	19	2.01	11.52
	PNB_N	0.12	0.42	0.59	0.75	0.94	3.00	0.46	1.38	4.65
Climate	TMN	-4.18	5.99	7.52	7.54	9.10	13.74	2.43	-0.41	4.34
Crop	CER	0.01	0.48	0.62	0.60	0.74	0.98	0.20	-0.52	2.87
	BEAN	0.00	0.05	0.13	0.20	0.30	0.98	0.19	1.40	4.80
	TUB	0.00	0.01	0.02	0.03	0.03	0.48	0.03	4.05	32.57
	SUG	0.00	0.00	0.00	0.01	0.01	0.35	0.02	4.19	36.24
	VEG	0.00	0.02	0.05	0.07	0.09	0.68	0.08	2.70	13.05
	ORC	0.00	0.00	0.01	0.04	0.06	0.69	0.07	2.45	10.81
	MEL	0.00	0.00	0.01	0.01	0.01	0.24	0.01	6.95	82.63
	EVI_JanFeb	0.03	0.08	0.10	0.09	0.11	0.19	0.02	-0.54	3.73
	EVI_SepOct	0.16	0.22	0.24	0.24	0.25	0.31	0.03	-0.16	3.12
	EVI_NovDec	0.05	0.08	0.10	0.10	0.11	0.19	0.02	0.12	3.48
Economy	GDP	0	5197	9359	24520	16859	4954333	148026	24.56	697.52
Soil	CAL	0.00	0.00	0.00	11.54	2.08	99.97	26.12	2.31	7.07
	GLE	0.00	0.00	0.00	0.83	0.00	39.96	4.26	6.39	45.95
	GYP	0.00	0.00	0.00	0.59	0.00	41.01	3.65	7.85	69.22
	SOCc	10.94	21.26	32.85	33.38	44.07	76.07	13.70	0.42	2.61
	HP	0.00	0.15	0.46	0.94	1.54	10.45	1.19	3.38	21.17
	PHk	4.73	5.45	5.74	5.73	6.02	6.69	0.41	-0.07	2.63
	ACID	0.00	0.00	0.57	0.51	0.97	1.01	0.44	-0.08	1.18
	COA	3.23	9.54	11.64	12.38	15.07	33.59	5.26	1.35	6.37

	CLAY	16.27	21.04	23.17	23.81	26.81	34.31	3.94	0.32	2.42
	SILT	27.31	38.58	42.77	41.50	44.92	50.00	4.82	-0.81	3.08
	TC	0.18	1.35	1.83	1.79	2.12	6.74	0.88	2.21	13.51
	TN	0.05	0.12	0.15	0.16	0.19	0.51	0.05	1.67	10.38
	TP	0.03	0.05	0.06	0.06	0.07	0.11	0.01	-0.01	3.02
	EXCA	0.02	0.09	0.11	0.11	0.13	0.18	0.03	0.25	2.74
	EXNA	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.44	4.04
	WATa2	12	13	14	13	14	15	1	-0.17	2.53
Topography	CUVlu	-0.01	0.02	0.04	0.05	0.07	0.19	0.03	1.35	5.13

Note: 1) PFP_N, partial factor productivity of nitrogen; PNB_N, partial nutrient balance of nitrogen; TMN, daily mean temperature; CER, planting area index of cereal; BEAN, planting area index of beans; TUB, planting area of tuber; SUG, planting area index of sugar; VEG, planting area index of vegetables; ORC, planting area index of orchards; MEL, planting area index of melons; EVI_JanFeb, enhanced vegetation index in January and February; EVI_SepOct, enhanced vegetation index in September and October; EVI_NovDec, enhanced vegetation index in November and December; GDP, gross output value of agriculture; CAL, calcisols; GLE, gleysols; Gyp, gypsisols; SOCc, soil organic carbon content; HP, histosols cumulative probability; PHk, pH in KCl; ACID, acid sub-soils grade; COA, coarse fragments volumetric; CLAY, clay content; SILT, silt content; TC, total carbon; TN, total nitrogen; TP, total phosphorus; EXCA, exchangeable calcium; EXNA, exchangeable sodium; WATa2, available soil water capacity with FC=pF 2.3; CUVlu, local upslope curvature. More explanations of the variable code names can be found in Table D.1.

2) Min., minimum value.

3) 1st. Qu., the first quantile value.

4) 3rd. Qu., the third quantile value.

5) Max., maximum value.

6) SD, standard deviation.

The regression coefficients and VIF of the SMLR models are shown in Table 4.2. Note that we standardized the explanatory variables to compare regression coefficients. All explanatory variables in the final model affected the NUE indicators significantly (*P*<0.01) and had a VIF less than 10: among them, seven explanatory variables had a negative impact and ten showed a positive influence on PFP_N; for PNB_N there were seven explanatory variables with a negative and eight with a positive influence. The studentized residuals were not strictly Gaussian distributed, which suggests that the normality assumption was not perfectly satisfied (Figure 4.6).

Table 4.2: Regression coefficients and variance inflation factor (VIF) of the stepwise multiple linear regression models.

Classification	Property[1)]	PFP_N model			PNB_N model		
		Regression coefficients	Std. Error[2)]	VIF[3)]	Regression coefficients	Std. error	VIF
	(Intercept)	39.64	0.21		0.75	0.00	
Climate	TMN	-5.15	0.48	5.4			
Crop	CER	2.63	0.30	2.1	0.07	0.01	4.2
	BEAN				0.16	0.01	6.0
	TUB				-0.02	0.00	1.2
	SUG	2.53	0.25	1.4			
	VEG	7.82	0.24	1.3			
	ORC	2.08	0.32	2.3			
	MEL				-0.02	0.00	1.1
	EVI_JanFeb	3.83	0.44	4.5			
	EVI_SepOct	-5.12	0.39	3.5			
	EVI_NovDec				-0.16	0.01	3.2
Economy	GDP	-1.59	0.22	1.1	-0.03	0.00	1.0
Soil	CAL				0.02	0.00	1.1
	GLE	4.41	0.24	1.3			
	GYP	5.32	0.24	1.3			
	SOCc				-0.05	0.01	7.4
	HP	-6.40	0.50	5.7			
	PHk				0.08	0.01	4.1
	ACID	3.66	0.49	5.5	0.12	0.01	7.4

	COA				-0.10	0.01	2.1
	CLAY	10.72	0.57	7.5			
	SILT				-0.11	0.01	3.6
	TC				0.02	0.01	1.8
	TN	4.96	0.52	6.3			
	TP				0.07	0.01	2.4
	EXCA	-1.90	0.42	4.1			
	EXNA	-3.17	0.37	3.2			
	WATa2				0.10	0.01	2.0
Topography	CUVlu	-3.56	0.32	2.4			

Note : [1] TMN, daily mean temperature; CER, planting area index of cereal; BEAN, planting area index of beans; TUB, planting area of tuber; SUG, planting area index of sugar; VEG, planting area index of vegetables; ORC, planting area index of orchards; MEL, planting area index of melons; EVI_JanFeb, enhanced vegetation index in January and February; EVI_SepOct, enhanced vegetation index in September and October; EVI_NovDec, enhanced vegetation index in November and December; GDP, gross output value of agriculture; CAL, calcisols; GLE, gleysols; Gyp, gypsisols; SOCc, soil organic carbon content; HP, histosols cumulative probability; PHk, pH in KCl; ACID, acid sub-soils grade; COA, coarse fragments volumetric; CLAY, clay content; SILT, silt content; TC, total carbon; TN, total nitrogen; TP, total phosphorus; EXCA, exchangeable calcium; EXNA, exchangeable sodium; WATa2, available soil water capacity with FC=pF 2.3; CUVlu, local upslope curvature. More explanations of the variable code names can be found in Table D.1.

[2] Std. error, standard error.

[3] VIF, variance inflation factor.

All variables were significant at the *P-value*=0.01 level.

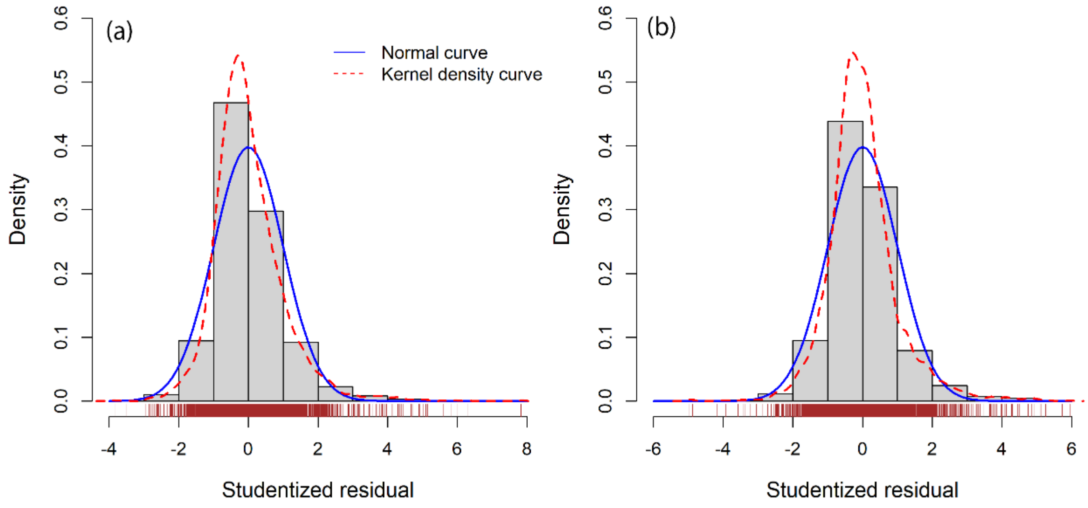


Figure 4.6: Distribution of residual standard errors for PFP_N model (a) and PNB_N model (b).

RF model

It took a while for the RF model to stabilize (Figure D.3). The MSE, RMSE and MEC converged to 57 (kg kg^{-1} yr^{-1})2, 7.5 kg kg^{-1} yr^{-1}, 0.84 for PFP_N, and 0.02 (kg kg^{-1} yr^{-1})^2, 0.15 kg kg^{-1} yr^{-1}, 0.89 for PNB_N, respectively, using 500 trees. This showed that a forest with 500 trees was large enough to get stable results. In this chapter, we therefore followed the rule-of-thumb that there should be at least 500 trees in a forest, even though a forest with 50 trees provided good results (Figure D.3).

Accuracy comparison between SMLR and RF models

RF used all 88 variables in a 'black box' approach while the SMLR models used fewer covariates. Thus, the latter were easier to interpret. The model performance metrics for each model were similar between 10-fold cross-validation and the internal model (original model built using the complete dataset) (Table 4.3), indicating there was no over-fitting problem. SMLR and RF models explained 44% and 84% of the PFP_N variance, and 67% and 89% of the PNB_N variance, respectively. Therefore, the RF model had much higher accuracy. The ME was close to zero and revealed unbiased predictions for all models. Regarding other model performance metrics, the RF models showed moderate performance, while SMLR models had poor performance, especially for PFP_N. The scatter and density plots of observations and predictions also showed that the RF models had smaller prediction errors than SMLR models (Figures 4.7f and h). As with most statistical prediction methods, both models smooth the reality and underpredict high extremes and overpredict low extremes.

Table 4.3: Model performance and cross-validation metrics of stepwise multiple linear regression and random forest models. Note: INT refers to internal model performance, CV to 10-fold cross-validation results.

Metrics	PFP_N				PNB_N			
	MLR		RF		MLR		RF	
	INT	CV	INT	CV	INT	CV	INT	CV
ME	-0.00	0.00	0.11	-0.11	-0.00	-0.00	0.00	0.00
RMSE	14.17	14.29	7.55	7.71	0.26	0.26	0.15	0.15
nRMSE	0.36	0.36	0.19	0.19	0.35	0.35	0.20	0.20
MEC	0.45	0.44	0.84	0.84	0.67	0.67	0.89	0.89
LCCC	0.62	0.61	0.91	0.91	0.80	0.80	0.94	0.94

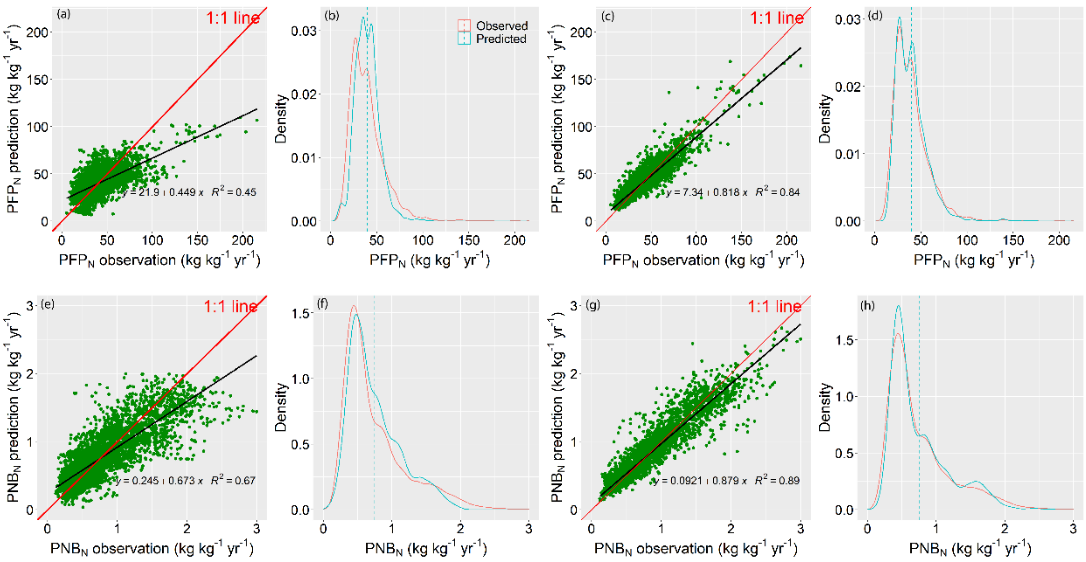


Figure 4.7: Scatter plots and Kernel density plots of observed and predicted PFP_N (a, b for SMLR model and c, d for RF model) and PNB_N (e, f for SMLR model and g, h for RF model). Note: the dotted line means average value in Kernel density plots; the average values of observed and predicted variables were the same.

4.3.2. Relative importance of explanatory variables

Here, we only show the variable importance results for the relweights method in the SMLR model, and the %IncMSE and IncNodePurity methods in the RF model (Figure 4.8). Results of the other variable importance methods (LMG and ranger) were similar to the above methods, so they are provided only in the supplementary material (Figure D.4).

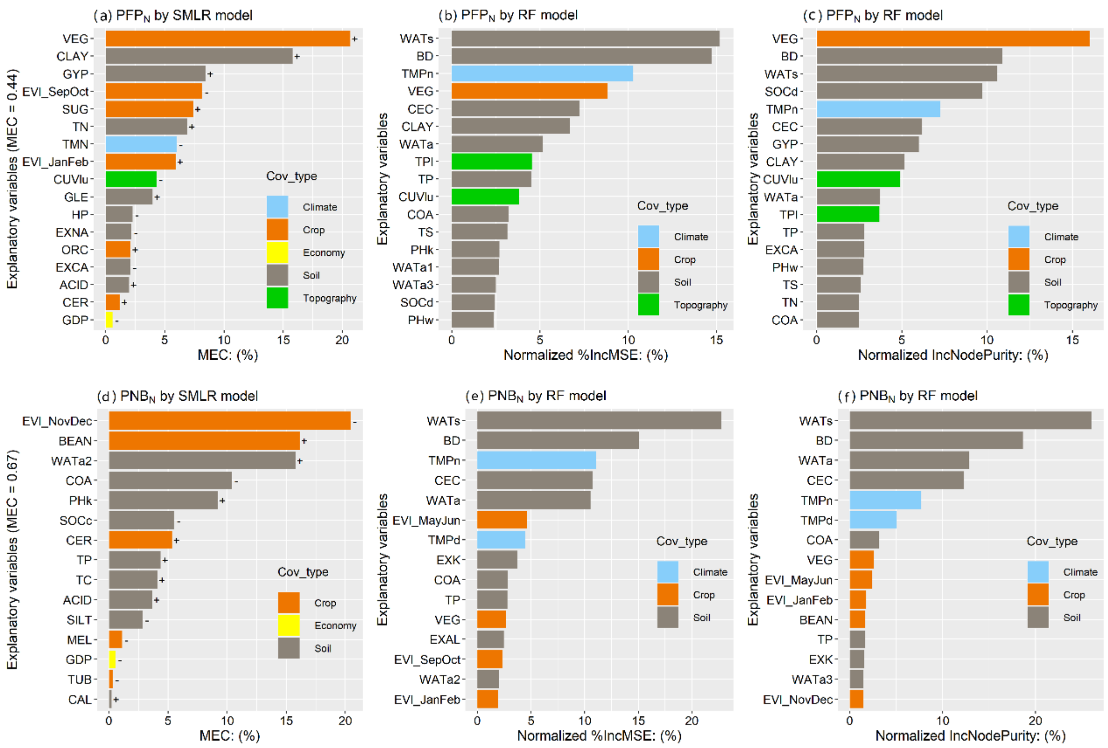


Figure 4.8: Relative importance of explanatory variables for the PFP_N: a) relweights method in SMLR model; b) %IncMSE method and c) IncNodePurity methods for RF model; and for PNB_N: d) relweights method in SMLR model; e) %IncMSE method and f) IncNodePurity methods for RF model. Note: ACID, acid sub-soils grade; BD, bulk density; BEAN, planting area index of beans; CAL, calcisols; CEC, cation exchange capacity; CER, planting area index of cereal; CLAY, clay content; COA, coarse fragments volumetric; CUVlu, local upslope curvature; EVI_JanFeb, enhanced vegetation index in January and February; EVI_MayJun, enhanced vegetation index in May and June; EVI_NovDec, enhanced vegetation index in November and December; EVI_SepOct, enhanced vegetation index in September and October; EXAL, exchangeable aluminum; EXCA, exchangeable calcium; EXK, exchangeable potassium; EXNA, exchangeable sodium; GDP, gross output value of agriculture; GLE, gleysols; Gyp, gypsisols; HP, histosols cumulative probability; MEL, planting area index of melons; ORC, planting area index of orchards; PHk, pH in KCl; PHw, pH in H2O; SILT, silt content; SOCc, soil organic carbon content; SOCd, soil organic carbon density; SUG, planting area index of sugar; TC, total carbon; TMN, daily mean temperature; TMNd, annual average surface temperature (daytime); TMNn, annual average surface temperature (nighttime); TN, total nitrogen; TP, total phosphorus; TPI, topographic position index; TS, total sulfur; TUB, planting area of tuber; VEG, planting area index of vegetables; WATa, available soil water capacity (volumetric fraction) until wilting point; WATa1/WATa2/WATa3, available soil water capacity with FC=pF 2.0/2.3/2.5; WATs, saturated water content (volumetric fraction). More explanations of the variable code names can be found in Table D.1.

Relative importance of explanatory variables for PFP_N

For PFP_N, the results of the relweights method suggested that crop and soil covariates had similar importance in the SMLR model accounting for 46% and 44% of the MEC, respectively (Figure 4.8a). The main crop covariates were planting area index of vegetables (21%), sugar (7%), EVI in September & October (8%), EVI in January & February (6%). Most of the covariates had a positive influence on PFP_N (EVI in September & October was an exception). Major soil covariates with positive regression coefficients were clay content (16%), gypsisols (8%) and total (7%). Soil covariates had more positive explanatory variables than negative variables (i.e., histosols cumulative, exchangeable sodium and exchangeable calcium). Climatic (i.e., annual average daily minimum temperature 6%), topographic (i.e., local upslope curvature 4%) and economic variables (i.e., agricultural GDP 0.6%) had negative impacts on PFP_N.

The %IncMSE and IncNodePurity methods showed similar results for variable relative importance (Figures 4.8b and c). Soil covariates accounted for the highest importance (73% for the normalized %IncMSE and 68% for the normalized IncNodePurity methods: e.g., bulk density 15 and 11%, saturated water content 15 and 11%, soil organic carbon density 2 and 10%, CEC 7 and 6%, clay content 7 and 5% for %IncMSE and IncNodePurity, respectively). Crop covariates (planting area index of vegetables) were less important than soil covariates, accounting for 9 and 16% for the %IncMSE and IncNodePurity methods, respectively. Topography ranked third (i.e., local upslope curvature 5%, topographic position index (TPI) 4% for %IncMSE, and the same for IncNodePurity) and climate factors fourth (i.e., annual average night land surface temperature 10 and 7% for %IncMSE and IncNodePurity, respectively).

There was a large difference in the variable importance between the SMLR and RF models. The relative importance of soil covariates in the RF model was much higher than in the SMLR model, while that of crop covariates was much lower. For individual variables, planting area index of vegetables and soil clay content had higher relative importance in the SMLR than in the RF model. In contrast, the relative importance of other soil covariates (e.g., bulk density, saturated water content, soil organic carbon density and CEC) and minimum temperature was higher in the RF than in the SMLR model.

Relative importance of explanatory variables for PNB_N

For PNB_N, the relweights method showed that soil covariates had the highest importance in the SMLR model (accounting for 56% of the MEC, e.g., available soil water capacity (with FC=pF 2.3) 16%, coarse fragments (>2 mm) 10%, pH (in KCl) 9% and soil organic carbon content 6%), followed by crop covariates (43%, e.g., EVI in November & December 20%, planting area index of beans 16%, cereal 5%) (Figure 4.8d). Economic covariates had the least influence, accounting for 0.5%

(i.e., agricultural GDP) of the MEC. Most soil covariates had a positive influence on PNB_N in the SMLR model, however coarse fragments, soil organic carbon content and silt content had negative effects (Figure 4.8d). EVI in November & December and planting area index of melons and tubers also decreased PNB_N, while planting area index of beans and cereals led to an increase. Agricultural GDP had, unexpectedly, a slight negative impact on PNB_N.

The variable relative importance of the %IncMSE and IncNodePurity methods for PNB_N were similar (Figures 4.8e and f). Soil covariates accounted for the highest importance (71% of the normalized %IncMSE and 77% of the normalized IncNodePurity methods, e.g., saturated water content 23 and 27%, bulk density 14 and 15%, available soil water capacity until wilting point 11 and 14% and CEC 11 and 13% for the %IncMSE and IncNodePurity methods, respectively), followed by climate factors (17% and 14% for the %IncMSE and IncNodePurity methods, respectively, i.e., annual average night land surface temperature 12 and 8% and annual average surface temperature at daytime 5 and 6% for the %IncMSE and IncNodePurity methods, respectively) and crop covariates (12 and 10% for %IncMSE and IncNodePurity, respectively, e.g., EVI in May & June 4 and 2% for the %IncMSE and IncNodePurity methods, respectively) (Figure 4.8f).

The main difference in the variable relative importance between the SMLR and RF models for PNB_N was that the RF model had higher relative importance of soil and climatic covariates, and lower relative importance of crop covariates. For individual variables, the relative importance of EVI in November & December and planting area index of beans were higher in the SMLR than in the RF model, while the relative importance of saturated water content, bulk density, CEC and annual average surface temperature was higher in the RF than in the SMLR model.

4.4. Discussion

4.4.1. Spatial and temporal variations of NUE indicators

Most counties in Jilin and Liaoning provinces had higher crop yields than most counties in Heilongjiang province where N application was lower (Figure D.2). Some counties such as Yichun and Qiqihaer cities in Heilongjiang province performed well since they produced higher yields with lower N inputs. These counties may have had high soil fertility, and good water and nutrient management. Management practices in these counties could be shared to support farmers in other counties, although it should be noted that in the long run soil depletion may still occur. During the 1990-2010 period there was an overall reduction in PFP_N: more counties had very low and moderate PFP_N levels, while less counties had high and very high PFP_N levels. After that period, the situation improved: less counties had PFP_N at very low levels (e.g. northeast of Jilin province), and more counties had PFP_N at a high level (e.g. middle and southwest of Liaoning province). Zhu et al. (2018) also found that PFP_N decreased from the 1980s to the 2000s. Therefore, with the NUE increasing in Jilin and Liaoning provinces, the spatial patterns of PFP_N might change in the future.

He et al. (2018) reported that PNB_N in northeast China was 0.50, 0.45 and 0.50 kg kg^{-1} yr^{-1} in the 1990s, 2000s and 2010s, respectively. The present study revealed a similar trend, but the values were higher, i.e., 0.76, 0.72 and 0.78 kg kg^{-1} yr^{-1}, respectively. The differences are likely caused by the fact that He et al. (2018) considered additional N input sources (symbiotic and non-symbiotic fixation, dry and wet atmospheric deposition, irrigation water and crop seed). Fixen et al. (2015) suggested that a PNB_N of 0.7-0.9 kg kg^{-1} is a typical level (with chemical and organic fertilizers as N inputs). In our study, while it is a good sign that more counties reached a moderate PNB_N level during the final time period, other counties still had lower levels. These counties require assistance as advanced management could improve efficiency or increase soil fertility. The EU Nitrogen Expert Panel (2015) observed that a PNB_N of 0.7-0.9 kg kg^{-1} indicates a balanced N fertilization, and suggested target values between 0.5-0.9 kg kg^{-1}. In our study, 64 counties were in this target interval in 2015. Note that the EU target is based on total N inputs, while we only included chemical and organic fertilizers. The average PNB_N in Heilongjiang province was higher than 1, which indicates means a soil mining situation which is not sustainable. PNB_N was lower than 0.55 kg kg^{-1} in Jilin and Liaoning provinces, suggesting avoidable N losses were occurring and the need for improved NUE (Snyder & Bruulsema, 2007).

4.4.2. Performance comparison between SMLR and RF models

In the SMLR models, regression coefficients are dependent on the unit of each explanatory variable. The SMLR models used 17/15 explanatory variables for PFP_N/PNB_N, but the RF model used all the variables in a 'black box' (no closed-form expression for the prediction equation). In this perspective, the SMLR model was simpler than the RF model and easier to explain. However, the RF model has a higher accuracy and the ability to model complex interactions between variables (Grömping, 2009; Richardson et al., 2017). Limitations of the SMLR and RF models are that these models are usually only effective within the range of covariate values exhibited by the training data, that relations found are empirical rather than causal and that overfitting may occur in instances where noisy data are being modelled (Hengl et al., 2015). Therefor it should be kept in mind that our prediction model should not be applied out of range. For model performance metrics (except ME), RF models were more predictive than SMLR models, especially for PFP_N. This suggests that including non-linear relationships between dependent and explanatory variables was more important for PFP_N than for PNB_N. The statistical model for predicting NUE at farm level had an MEC of 0.77 (Ramírez & Reheul, 2009). **Chapter 2** obtained an NUE model at provincial scale with an MEC of 0.74. They were less predictable than the RF model in this article (MEC of 0.84/0.89). In principle, our RF model is especially good at predicting NUE at county level in northeast China which is consistent with training data. Predictions outside the area of applicability should be handled with care or be left out from further consideration because the environmental properties differ too strongly from those observed in the training data (Meyer & Pebesma, 2021).

4.4.3. Relative importance comparison of explanatory variables

Relative importance of variables for PFP_N

For PFP_N, the crop covariates (e.g., planting area index of vegetables) were the most important factors for the relweights and %IncMSE methods, while they ranked second in the IncNodePurity method. In particular, planting area index of vegetables had a highly positive influence on PFP_N in the SMLR model. The yield per unit of vegetables (measured as fresh weight) was higher than that of other crops (measured as dry matter). This is not surprising given that fresh weight is typically larger than dry matter weight. However, this does not mean that more vegetables should be grown, because this would neither increase the nutritious value nor promote food security. Besides that, the EVI was an important factor in the relweights method. Although they are all negatively correlated with PFP_N, but the equation coefficients showed that EVI in September and October had a negative influence while EVI in January and February had a positive influence on PFP_N in the SMLR model.

Soil covariates were the most important variables for the IncNodePurity method, while they were the second most important for the relweights and %IncMSE methods. The representative covariates in the SMLR model were clay content, gypsisols and total N. Hamoud et al. (2019) indicated that an increase in soil clay content improved rice root morphology and NUE. In addition, more clay means less leaching because leaching mainly takes place in sandy soils. Cambisols, soil organic carbon density and the amount of phosphorus (Olsen method) were the representative variables in the %IncMSE method of the RF model. The IncNodePurity method suggested that bulk density, saturated water content, soil organic carbon density and CEC were the main soil covariates. These covariates are important soil properties, which have a direct influence on N uptake by crops and soil N mineralization (Ishaq et al., 2002). Saturated water content reflects the pore status of the soil and the maximum water holding capacity, which is beneficial to NUE (Iqbal et al., 2019). The high saturated water content (43-56%) also reflects high soil organic matter. CEC gives insight into the fertility and nutrient retention capacity of the soil, because exchangeable cations are the most important source of immediately available K, Ca and Mg (Mukhopadhyay et al., 2019). Soil covariates had more positive explanatory variables than negative variables (i.e., histosols cumulative, exchangeable sodium and exchangeable calcium). **Chapter 2** also indicated that histosols cumulative had a negative influence on yield, because histosols typically are poorly drained soils, where organic matter accumulation is greater than mineralization.

The relative importance of climate factors was similar in the relweights and IncNodePurity methods. In the relweights method of the SMLR model, annual average daily minimum temperature had negative impacts on $PFP_{N,}$ and in the RF model, annual average surface temperature at night was also important. Temperature is among the key driving factors of crop yield variability (Zscheischler et al., 2017). The temperature in northeast China is lower than in other regions of China and the minimum night land surface temperature restricts the PFP_N by limiting crop growth and N absorption (Jin et al., 2017). (Chen et al., 2011b) found that daily minimum temperature was the dominant factor in corn production, especially in May and September.

The economic variable (i.e., agricultural GDP) was more important in the IncMSE method of the RF model than in the relweights method of the SMLR model. Agricultural GDP refers to the total amount of all agricultural products expressed in monetary currency, reflecting the general trend of agricultural production in a certain period, which is related to crop types, crop yield and market price. This influences NUE at a macroeconomic level. In this chapter, agricultural GDP per unit of planting area, rather than per capita, was used to better reveal the agricultural productivity and explore the relationship with NUE. Higher agricultural GDP promotes the improvement of mechanization, and this pays off in a higher yield and NUE. Conversely, once the agricultural GDP reaches a certain level, excessive fertilization might also cause reduced NUE. The non-linear relationship between agricultural GDP and PFP_N explained why agricultural GDP had a low importance in the SMLR model.

Topography was more important in the RF model than the SMLR model. Local upslope curvature and TPI were the main topographic covariates. Local upslope curvature is detrimental to PFP_N. TPI measures the relative topographic position of a point as the difference between the elevation at this point and the mean elevation within a predetermined neighborhood (De Reu et al., 2013). TPI could influence climate variability (temperature and precipitation) (Tang et al., 2017), and N loss (Singh et al., 2019).

The relative importance of soil covariates in the RF model was much higher than in the SMLR model while that of crop covariates was much lower. This may be because crop covariates tend to be more linearly correlated with PFP_N, while soil covariates are more complex and tend to be nonlinearly correlated. For individual variables, planting area index of vegetables had higher relative importance in the SMLR than in the RF model, because vegetables had higher PFP_N than other crops (> 80 kg kg^{-1} in Li et al. (2017). The soil clay content is also more important in the SMLR model. This might be because the range of soil clay content (16-34%) in this chapter had a linear relation with PFP_N, so its influence is well-represented in the SMLR model. The relative importance of minimum temperature was higher in the RF than in the SMLR model. This might be because the relationship between temperature and NUE is nonlinear. Peng et al. (2004) showed that increased night temperature decreased rice yields, which was associated with global warming. In contrast, Jin et al. (2017) indicated that higher NUE and rice yield could be obtained with a combination of higher effective temperature and fewer sunshine hours during the entire growth period in northeast China.

Relative importance of variables for PNB_N

For PNB_N, three methods showed that soil covariates were the most important variables (>50%). In the relweights method, available soil water capacity and pH had a positive influence while coarse fragments and soil organic carbon content had a negative impact on PNB_N. Pan et al. (2020) showed that greater water holding capacity improved maize growth and thus enhanced N absorption and accumulation by plants. Table 4.1 shows that the minimum pH was 4.73, so soil acidification of croplands decreased NUE (Pan et al., 2020), which can be overcome by the use of a combination of chemical and organic fertilizers (Miao et al., 2010). Coarse fragments increase the risk of N loss and ammonia volatilization, and thus reduce crop NUE (Peng et al., 2015). The saturated water content was very important in the %IncMSE and NodePurity methods. Saturated water content refers to the water content when all pores in the soil are filled with water. It is one of the soil moisture constants, reflecting the pore status of the soil and the maximum water holding capacity. Improved soil porosity is beneficial to NUE (Iqbal et al., 2019). The bulk density, available soil water capacity until wilting point and CEC were also very important factors. CEC is a sign of soil fertility and nutrient retention capacity (Mukhopadhyay et al., 2019). Soil water affects nutrient transformation from unavailable to available forms and vice versa, and thereby the total nutrient uptake amount. It also influences the availability of applied nutrients and efficiency through its effect on various nutrient loss mechanisms such as volatilization, nitrification, and/or urease hydrolysis (Ullah et al., 2019).

Crop covariates were second in importance in the three methods. EVI in November & December and planting area index of melons and tubers decreased PNB_N, while planting area index of beans and cereal led to an increase. EVI in November & December also had a significantly negative correlation with PNB_N (Figure D.5b). This is as expected because long winter dormancy inhibits crop growth and has a negative effect on PNB_N (Gurung et al., 2009). Melons have a lower N requirement per unit of economic yield than beans (**Chapter 2**), which explains their negative influence on PNB_N. The planting area index of cereals was negatively correlated with PNB_N based on Pearson's correlation but showed positive influence in the SMLR model. This indicated that the influence of cereals on PNB_N was affected by other explanatory variables. Furthermore, cereal crops have various varieties. So NUE varied largely between different cereal crops and management practices (Xu et al., 2014b; Zhang et al., 2015).

Climate factors were important in RF models according to the %IncMSE and NodePurity methods. Ding et al. (2018) also found that climate and soil variation were dominant factors driving N leaching. This is because temperature influences the N availability to crops and changes crop N removal and N losses from the plant-soil system (Liang et al., 2018).

We had expected that increasing wealth would raise farmers' investment in farming and the government's attention to agriculture, and thus provide better agrotechnical services and increase NUE. Most countries, however, show an attenuated increase of maximum N removal from crops with increasing GDP, which adheres to the law of diminishing returns (Mogollón et al., 2018). Despite that, the temptation of higher interests might lead to excessive fertilization and reduce NUE. In Japan, a negative relationship between GDP and maximum N removal from crops was found (Mogollón et al., 2018). Proper guidance by the government and related scientific research workers is needed to avoid the negative effect of agricultural GDP.

The main difference in the relative importance of variables between the SMLR and RF models for PNB_N was that the RF model had higher relative importance of soil and climatic covariates, and lower relative importance of crop covariates. Similar to that observed for PFP_N, this may be because crop covariates have a simple linear relationship with PNB_N while soil and climate covariates are more complex and nonlinearly related with PNB_N. For example, the relative importance of EVI in November & December and planting area index of beans were higher than 15% in the SMLR model, while they were lower than 5% in the RF model. In contrast, the relative importance of saturated water content, bulk density and CEC was higher in the RF than in the SMLR model; the annual average surface temperature was important in the RF model, but not in the SMLR model. This could be because EVI in November & December has a linear relation with PNB_N, which SMLR captures well. However, soil covariates and temperature have a complex non-linear relation with PNB_N, by influencing N uptake and denitrification, which may explain why these were more important in the RF model (Fan et al., 2014).

We compared different methods that quantify the relative importance of explanatory variables. Although the methods used for this are not entirely new, their comparison is rarely made and certainly not in the context of NUE modelling. Considering that the ranger package is much faster, we found that the permutation and node impurity method used in the ranger package were more efficient than in the randomForest package. However, it was difficult to recognize the negative or positive influence of covariates and interpret results with the ranger package. More enhanced methods, such as deep learning, transfer learning, and wavelet phase harmonics (WPH) statistics are needed (Allys et al., 2020). One thing that attracted our approval was that the %IncMSE method in the randomForest package and permutation method in the ranger package were unscaled, because scaling of variable importance is misleading (Strobl et al., 2009).

4.4.4. Crop-specific analyses at county scale

Many studies have focused only on a single crop or grain crop, and rarely included cash crops (Xu et al., 2014b; Li et al., 2020a). As a result, these studies do not reflect the entire local NUE trends, which inhibits their use for formulating appropriate policies and technical guidance. This chapter made up for these shortcomings by aggregating all crops and including crop types as explanatory variables. Analysis at a higher level of aggregation is important for decision and policy makers who often need integrated information that shows general patterns and trends (**Chapter 2**). However, in some cases decision makers may benefit from crop-specific analyses. Such analysis was beyond the scope of this work and is also hampered by the fact that crop-specific N application data at county scale are not readily available in the governmental database (i.e., in the municipal and provincial Yearbook of northeast China). Future research could look into ways to collect such data and apply a similar analysis as we did to reveal trends and patterns for specific crops and crop categories, such as staples and cash crops. Our analysis was also conducted at a county level, thus enabling recommendations and policies, such as N fertilization regulation, at a much finer scale than the provincial level. This is important because our results indicated that there were considerable differences in NUE trends between counties in the same province.

4.5. Conclusions

We performed a first and novel study analyzing spatial and temporal variation of NUE at a county scale in three provinces in northeast China using statistical regression. The NUE indicators decreased in most counties during the study period and were higher in Heilongjiang province than in the other two provinces. The RF model was superior in performance than the SMLR model, indicating that many covariates had a non-linear relation with NUE. We note that both models smooth the reality and underpredict high extremes and overpredict low extremes. Considering that the ranger package is much faster than the randomForest package, we found that the permutation and node impurity methods in the ranger package were more efficient for analyzing covariate importance in the RF model.

The relative importance of crop covariates was much higher in the SMLR than in the RF model, while soil and climatic covariates were more important in the RF model, indicating the difference between linear and non-linear relationships between dependent and explanatory variables. These novel findings are particularly valuable when put into action in supporting land-use management and policymaking.

Supplementary material

The supplementary materials, Appendix D, can be downloaded as Appendices A-G from: https://wageningenur4-my.sharepoint.com/:w:/g/personal/yingxia_liu_wur_nl/EQLnRgk6ERRPjygndH-miiYBCx793tl1hqLuvQ_B7xKQWQ?email=liuyingxia91%40163.com&e=jySl9r.

Chapter 5

Uncertainty quantification of nitrogen use efficiency prediction in China using Monte Carlo simulation and quantile regression forest

Nitrogen use efficiency (NUE) plays an essential role in food security and environmental sustainability. This chapter analyzed uncertainty in NUE predictions obtained from a random forest machine learning model. We quantified the uncertainty of the input data and model and used Monte Carlo simulation in three scenarios and quantile regression forests (QRF) to analyze how these uncertainties propagate to the NUE predictions for 31 provinces in China from 1978 to 2015. We considered two specific NUE indicators, the partial factor productivity of nitrogen (PFP_N) and the partial nutrient balance of nitrogen (PNB_N). The prediction uncertainty for both NUE indicators decreased over time. In 2015, PFP_N had a higher 90% prediction interval ratio (PIR_{90}) of input data in south and west China and a higher 90% prediction interval width (PIW_{90}) in south and east-costal China, while PNB_N had a higher PIR_{90} in north China and a higher PIW_{90} in northeast China. The NUE prediction uncertainty propagated from QRF models has similar spatial patterns as input data. NUE in most provinces had smaller input uncertainty than model uncertainty, except PNB_N, which had larger model uncertainty than input uncertainty after 2010. Generally, PNB_N had higher input uncertainty contributions than PFP_N in 2015, especially in south and northeast China. Overall, the uncertainties in NUE predictions was substantial. A series of recommendations were made to improve the accuracy of NUE prediction. These may be applied by the government, in order to inform sustainable nitrogen management in food systems.

Based on:

Liu, Y., Heuvelink, G. B. M., Bai, Z., He, P. (2022). Uncertainty quantification of nitrogen use efficiency prediction in China using Monte Carlo simulation and quantile regression forest. (submitted to *Journal of Cleaner Production*)

5.1. Introduction

The Sustainable Development Goals defined by the United Nations call to action to ensure agricultural and environmental sustainability. Nitrogen (N) is one of the most important nutrients for crop growth, contributing more than 50% of the crop yield increase (Stewart. et al., 2005; Zhang et al., 2015; Dimkpa et al., 2020). Studies demonstrated that excessive N application caused diminishing returns for increased crop yields and N removals, especially in China (Zhang et al., 2015). In other words, excessive N application leads to decreased nitrogen use efficiency (NUE). NUE indicators are a major detection instrument related to food security, environmental pollution, economic development and resource use, and widely and increasingly used by agronomists, environmental scientists, biogeochemists, policymakers and other stakeholders at various temporal and spatial scales (Quan et al., 2021). The partial factor productivity of N (PFP_N, in kilograms of grain per kilogram of N applied) and partial nutrient balance of N (PNB_N, in kilograms of N removal by aboveground crop per kilogram of N applied) are two popular indicators to reflect NUE averages and trends (Dobermann, 2007). Specifically, NUE (38 and 0.4 kg kg^{-1} yr^{-1} for PFP_N and PNB_N, respectively) (**Chapter 2**) in China is much lower than in well managed systems in other parts of the world (>60 and 0.7-0.9 kg kg^{-1} yr^{-1} for PFP_N and PNB_N, respectively) (Roberts, 2007; Fixen et al., 2015). Therefore, improving NUE indicators to a reasonable range is vital to environmental sustainability in China.

Many studies revealed that NUE has large spatial and temporal variation as a consequence of different crops, climate, soil properties and management practices (Zhang et al., 2015). For example, PFP_N changed from 86, 56 and 74 kg kg^{-1} in 1960 to 26, 30 and 69 kg kg^{-1} in 2009 in China, India and USA, respectively (Lassaletta et al., 2014). Statistical modelling can support the understanding of the causes of the spatial and temporal variations (Li et al., 2020b; Zhang et al., 2021b). Random forest (RF), as one of the machine learning algorithms, stands out due to its flexibility and competitive explanatory performance. Li et al. (2020b) used RF and found that mean annual temperature was the most critical factor of PFP_N improvement. Ren et al. (2019) applied RF to the effects of manure on recovery efficiency of N, and found that soil properties contributed up to 55%, followed by climatic factors that contributed of 32%. N application rate and crop types only contributed 8.3% and 4.6%, respectively. Correndo et al. (2021) explored the driving factors of maize yield without N fertilizer using RF, and indicated that previous crop, irrigation and soil organic matter were the most relevant factors for maize yield.

Even though the explanatory performance of RF models is often high compared to other models, no model is perfect and model predictions have errors, which might influence their explanatory ability. Models are useful tools to support increasing policy interest in monitoring and reporting performance of sustainable agriculture and sustainable nutrient management practices, but only if their accuracy is sufficient. So far, current research rarely addressed prediction uncertainty of NUE models, even though this restricts their application. Additionally, the sources of uncertainty and their contribution to the model output uncertainty have been poorly understood. Generally speaking, the main uncertainty sources originate from model inputs, model parameters and model structure (Heuvelink et al., 1989; Kay et al., 2009). Consequently, all three will propagate to the model output.

The uncertainty of model inputs includes uncertainties in measurements, such as yield and fertilizer inputs, and uncertainties in calculation parameters, such as straw return and manure coefficients. When collecting data, measurement uncertainty can be caused by many factors, such as sampling error, human error, instrument error, and parameter error in data processing (e.g., N content in manure, straw and grain ratio). These uncertainties propagate to the model output (i.e., NUE predictions). Zhang et al. (2021b) discussed the uncertainty sources of the N budget, such as calibration parameters, data sources and calculation methods. Zheng et al. (2012) applied Monte Carlo simulation to identify key uncertainty sources of atmospheric ammonia emission by simulating input data (i.e., N fertilizer application, human being, livestock, biomass burning, waste treatment, fuel combustion and sewage treatment) based on the reliability and accuracy of data sources, estimation methods used, and uncertainty in emission factors and by propagating uncertainties in model inputs through source-based emission models. Miller et al. (2020) used the Monte Carlo method to propagate the uncertainty of N application measurements to the overall N application rate. Kros et al. (2012) also used the Monte Carlo method to quantify the propagation of input uncertainty from initial values, model parameters, and environmental constants and variables to model output (N fluxes). Although computationally demanding, the Monte Carlo has shown to be an effective method for quantifying the propagation of input and parameter uncertainty to model outputs.

In addition to uncertainty in model input, there is also uncertainty caused by the model itself, because even if the inputs are error-free, the output is still imperfect. This is because a model is only a simplified representation of reality (Heuvelink, 1998a). de Vries et al. (2011) estimated the model structural uncertainty for N budgets for European agriculture by using an ensemble of modelling approaches. In machine learning, the quantile regression forests (QRF) algorithm was developed and used to assess the uncertainty associated with model-derived predictions (Meinshausen, 2006; Córdoba et al., 2021). Lalitha et al. (2021) used QRF to quantify the model uncertainty of soil depth prediction, while Vaysse and Lagacherie (2017) showed that QRF provided more accurate and interpretable predicted patterns of uncertainty than regression kriging in operational digital soil mapping.

In this chapter we aim to quantify the uncertainty and uncertainty source contributions in random forest predictions of NUE using a Monte Carlo analysis and QRF for 31 provinces in China, for the years 1978 to 2015. Specifically, the objectives of this chapter are to: 1) use expert judgement and scenarios to quantify the uncertainty (probability distributions) in calculations of PFP_N and PNB_N (through propagating uncertainty in crop yield, N removal, and N input); 2) analyze how uncertainty in the inputs propagates through the RF model using a Monte Carlo uncertainty propagation analysis; 3) analyze how model uncertainty leads to uncertainty in model outputs using QRF; and 4) compare the contributions of input uncertainty and model uncertainty to the overall model output uncertainty. To facilitate the robustness of the results, we define three scenarios for input uncertainty quantification. We compare the input uncertainty with model uncertainty for one of these scenarios (i.e., the reference scenario). Our results can be used to benchmark new national food security projections and quantitative scenario studies and inform policy analysis and the public debate on improved data collection and model building.

5.2. Materials and methods

5.2.1. Study area and NUE indicators

The study area is defined by 31 provinces of China (excluding Hong Kong, Macao and Taiwan), for a time period from 1978 to 2015. The target variables studied were two indicators for long-term trends of NUE: PFP_N and PNB_N (Table 5.1) (Dobermann & Cassman, 2005; Dobermann, 2007). The dataset to calculate these indicators stems from the (National Bureau of Statistics of China, 2019). It mainly includes economic yield of different crops, human and livestock numbers, chemical fertilizer and related parameters, as listed in Table 5.1. More details are given in (He et al., 2018).

Table 5.1: Definitions of nitrogen use efficiency indicators used in this chapter, with parameter descriptions (Dobermann, 2007; He et al., 2018).

Definitions and equations	Parameter descriptions
$PFP_N = \frac{Yield}{N_{input}}$	PFP_N represents the partial factor productivity of nitrogen in each province (kg kg^{-1} yr^{-1}); *Yield* is the economic yield of total crops in each province (kg yr^{-1}); N_{input} represents the amount of N input (i.e., chemical fertilizer, manure, cake fertilizer and straw return) to the soil in each province (kg N ha^{-1} yr^{-1}).
$PNB_N = \frac{N_{crop\_removal}}{N_{input}}$	PNB_N represents the partial nutrient balance of nitrogen in each province (kg kg^{-1} yr^{-1}); $N_{crop\_removal}$ represents the amount of N removal from total crops in each province (kg N ha^{-1} yr^{-1}).
$Yield = \sum_{k=1}^{p} Yield_k$	*p* means the number of crop species in each province.
$N_{crop\_removal} = \sum_{k=1}^{p} N_{crop\_removal\_k}$	See above.
$N_{crop\_removal\_k} = Yield_k \times Crop_k$	$Crop_k$ means the amount of N requirement for unit of economic yield of different crops in N kg t^{-1}.
$N_{input} = N_{fert} + N_{man} + N_{cake} + N_{straw}$	N_{fert} is total N from chemical fertilizer (kg N ha^{-1} yr^{-1}); N_{man} is manure fertilizer from humans and livestock (kg N ha^{-1} yr^{-1}); N_{cake} is cake fertilizer returns to soil (kg N ha^{-1} yr^{-1}); N_{straw} is straw N returns to soil (kg N ha^{-1} yr^{-1}).
$N_{fert} = N_{fert\_N} + N_{fert\_Com} \times Ratio_{Com}$	N_{fert_N} and N_{fert_Com} are the total chemical fertilizer amount of single and compound (kg N ha^{-1} yr^{-1}), respectively; $Ratio_{Com}$ is the N ratio of compound fertilizer (%).
$N_{man} = \sum_{k=1}^{n}(Num_{animal\_k} \times NExcre_{rate\_k} \times MRate_{retur\_k})$	*n* is the number of human and livestock species; Num_{animal} is the number of humans and livestock; $NExcre_{rate}$ is the amount of N content in excretion and urine in different humans and livestock (kg N head^{-1} yr^{-1}); $MRate_{retur}$ is the returning field rate (%).
$N_{cake} = \sum_{k=1}^{m}(Yield_k \times Ratio_{cake\_k} \times NCake_k \times CRate_{retur\_k})$	*m* is the number of crop species for cake fertilizer; $Ratio_{cake}$ is the cake ratio; *NCake* is the amount of N content in different crops for cake fertilizer (%); $CRate_{retur}$ is the returning field rate (%).
$N_{straw} = \sum_{k=1}^{i}(Yield_k \times Ratio_{straw\_k} \times NStraw_k \times SRate_{retur\_k})$	*i* is the number of crop species for straw returning; $Ratio_{straw}$ is the ratio of straw to grain; NS_{traw} is the amount of N content in different straw (%); $SRate_{retur}$ is the returning field rate (%).

5.2.2. Uncertainty of input data

Measurement errors represented by probability distributions

Technically, our research used official data which should be trustworthy. However, no measurement is error-free due to the use of different equipment in field and lab, and sampling errors. In addition, the uncertainty of the input parameters used to calculate N removal and N input can be different between provinces and over different years. Moreover, measurement errors can also be correlated in space and time, because similar biases can be made in subsequent years or in neighboring provinces.

Mathematically, measurement error can be represented as follows:

$$Y_t = Y_m + \varepsilon \tag{5.1}$$

where Y_t and Y_m are the true and measured values of an input parameter, respectively; and ε is a random measurement error (i.e., the difference between the true and the measured value).

The actual measurement errors are unknown and therefore represented by probability distributions. A major task is then to quantify the statistical parameters of these probability distributions. We made the following assumptions to be able to do this for yield, N removal and N input:

- All measurement errors are assumed to be normally distributed. This substantially facilitates the subsequent statistical analysis and is defendable because of the central limit theorem (Burt et al., 2009).
- The mean (i.e., expected value) of the error distributions are assumed to be zero in all cases. In other words, we assume that there are no systematic measurement errors over the provinces and years.
- With the two assumptions above, the marginal distributions of the measurement errors are completely specified once the standard deviation of the distribution is known. We assume that these standard deviations are proportional to the measured value.
- We further assume that errors between different inputs (such as between yield and N removal) are uncorrelated.
- We allow measurement errors of the inputs to be spatially autocorrelated and we characterize the spatial autocorrelation with a negative exponential function of the geographic distance between the centroids of the provinces.
- Temporal autocorrelation of measurement errors is also incorporated; this is represented by a negative exponential function of the time difference in years.

Quantifying the standard deviation of input errors

As mentioned above, the standard deviation of the input errors is assumed to be proportional to the measured value:

$$\sigma_{ijk} = \frac{PE_{ijk}}{100\%} \cdot Y_{m,ijk} \tag{5.2}$$

Here, σ_{ijk}, PE_{ijk} and $Y_{m,ijk}$ are the standard deviation, the proportional error (expressed as a percentage) and the measured value for province i, year j and input k, respectively.

The input errors mainly include errors from crop yield, N removal and N input. Specifically, N removal was calculated from yield and N requirement for each crop. However, the N requirement may be different for different crop varieties. For the uncertainty of N requirement in the reference scenario we refer to He et al. (2018). They summarized mean values and standard deviations of N requirement from previous publications. The N input was obtained by aggregating N from chemical fertilizer, manure, cake fertilizer and straw return. Each of these was calculated by the equations given in Table 5.1, with corresponding parameters given in Table 5.2. For example, the chemical fertilizer as one of the N inputs, was composed of single N fertilizer and N from compound fertilizer (Table 5.1). Due to the N proportion in compound fertilizer being different between diverse fertilizer companies and brands among different years, we used different parameters for different regions and years. The proportional error (PE) of single N fertilizer and compound fertilizer can be derived from "N input", while the PE from N ratio in compound fertilizer is derived from the "N input coefficient" (Table 5.2). Manure, as the second N input, included even more parameters: the number of human and livestock species, the amount of N content in excretion and urine per year (i.e., daily excretion/urine, N content, population structure, and feeding period) and the returning rate to field (Table 5.1). To simplify our calculations, we used a single coefficient from N input coefficient to represent the aggregated parameters of manure (i.e., the amount of N content in excretion and urine and returning rate to field) (Table 5.2).

Table 5.2: Values used for proportional errors (Popescu et al.) in three scenarios to assess the robustness of the uncertainty analysis.

NUE measurements	Year	Types	Optimistic	Reference	Pessimistic
Yield	1978-1988	Cereals	2.5	5	10
		Crops with dry matters except cereals	5	10	20
		Sugar crops, vegetables, fruits and melons	10	20	40
	1989-1996	Crops with dry matter	4	8	16
		Sugar crops, vegetables, fruits and melons	8	16	32
	1997-2009	Crops with dry matter	3	6	12
		Sugar crops, vegetables, fruits and melons	6	12	24
	2010-2015	Crops with dry matter	2.5	5	10
		Sugar crops, vegetables, fruits and melons	5	10	20
N input	1978-1983	Single N fertilizer	5	10	20
		Compound fertilizer	5	10	20
		Number of livestock and human	5	10	20
	1984-1996	Single N fertilizer	4	8	16
		Compound fertilizer	4	8	16
		Number of livestock and human	4	8	16
	1997-2009	Single N fertilizer	3	6	12
		Compound fertilizer	3	6	12
		Number of livestock and human	3	6	12
	2010-2015	Single N fertilizer	2.5	5	10
		Compound fertilizer	2.5	5	10
		Number of livestock and human	2.5	5	10

N input coefficient	N ratio in compound fertilizer	2.5	5	10
	Manure coefficient	4	8	16
	Cake fertilizer coefficient	5	10	20
	Straw returning coefficient	5	10	20
N require	Rice			
	Early rice	7.7	15.3	30.7
	Middle rice	6.7	13.4	26.8
	Late rice	7.7	15.4	30.8
	Single-season rice	6.6	13.1	26.2
	Wheat			
	Winter Wheat	3.4	6.8	13.6
	Spring Wheat	3.0	6.0	12.0
	Maize			
	Spring maize	8.5	17.1	34.1
	Summer maize	5.7	11.3	22.6
	Millet	3.8	7.6	15.2
	Sorghum/Jowar	0.5	1.0	2.0
	Barley	4.1	8.1	16.3
	Other cereals	0.5	1.0	2.1
	Potatoes	6.1	12.1	24.3
	Other potatoes	0.6	1.1	2.2
	Soybeans	2.6	5.2	10.4
	Other bean crops	2.7	5.4	10.8
	Cotton	4.0	7.9	15.9
	Fiber crops	2.5	5	10
	Peanut	9.7	19.4	38.8
	Rapeseed	7.9	15.7	31.5
	Sunflower/Helianthus	12.9	25.8	51.7

Other oil crops	1.3	2.7	5.3
Tobacco	4.5	9.1	18.2
Tea	2.2	4.4	8.8
Sugarcane	0.3	0.6	1.1
Sugar beet/Beetroots	5.4	10.8	21.7
Vegetables	8.5	16.9	33.8
Melons	3.9	7.9	15.8
Garden fruits	5	10	20
Banana	0.3	0.5	1.1
Apple	14.9	29.7	59.5
Oranges	6.2	12.4	24.8
Pear	1.5	3.1	6.2
Grape	6.1	12.1	24.3
Pineapple	1.0	1.9	3.8
Dates	5.0	10.0	20.0
Peach	0.4	0.9	1.8
Persimmon	5.3	10.6	21.3
Kiwi	9.5	19.0	38.0
Litchi	3.8	7.7	15.4
Longan	4.8	9.5	19.0

Note: Cake fertilizer, straw return and manure coefficients represent the aggregated parameters (except yield) calculating cake fertilizer, straw return and manure. The PE for N requirement is derived from He et al. (2018).

Generally, activity data collected from the official statistics or from first-hand measurements have a smaller PE (likely between 2.5 and 10%) (e.g., crop yield, animal numbers, single N fertilizer application amount) (Huang et al., 2012; Gu et al., 2015; Luo et al., 2019; Xu et al., 2019) than other activity data and parameters collected or summarized from published studies, which are prone to have a larger PE range (0.3-80%) (e.g., N requirement, N input) (He et al., 2018). The PE can be different between crops, years and provinces.

The PE in the optimistic and pessimistic scenarios were chosen as half and double the PE values of the reference scenario, respectively. Due to the large difference among the water content of sugar crops, vegetables, fruits and melons, we assumed that the uncertainty of these crops was larger (i.e., double) than for dry matter crops, such as cereals and oil crops. Cereal yield information was obtained from a complete survey data for the period before 1989. We assumed that the PE of cereal yield was small (i.e., 5% for the reference scenario), while other crops with dry matters had a larger PE (i.e., twice as large as cereal crops: 10% for the reference scenario). Sugar crops, vegetables, fruits, melons had a twice as large PE than dry matter crops, as mentioned before. From 1989 onward sample survey data were used (National Bureau of Statistics of China, 2019). Therefore, the cereal yield uncertainty was larger than before, which was influenced by sample quality and planting area (because total production is the product of sample yield and planting area). In addition, the improvement of statistical methods and laws decreased the uncertainty of statistical data over the years. The statistics law of the People's Republic of China has been enacted in 1983, and amended in 1996, 2009. Based on this, we used different PEs between different crops and we decreased it for more recent years. That is to say, the PE of dry matter crops from 1989 to 1996 (8% for the reference scenarios) was smaller than in 1978-1988, but larger than cereal crops in the previous period that was based on a complete survey data.

The above showed that quantification of the PEs relied largely on expert judgements, which is not as reliable as quantitative assessment of measurement errors. We therefore decided to use three scenarios to quantify the uncertainty of measurements of yield, N removal and N input: Optimistic (O), Reference (R) and Pessimistic (P) (see Table 5.2). Comparison of results between these scenarios provides insight into how sensitive the results of the uncertainty propagation analysis are to the values of the PEs.

Spatial and temporal correlation of the errors

Uncertainty about spatially distributed input data tends to be positively spatially correlated, and this influences the degree to which uncertainties cancel out by spatial aggregation (Kros et al., 2012). We used centroids of provinces to determine the distances between provinces. The uncertainty about input data might also be positively correlated over time, since in nearby years the same errors can be repeated by using the same tools, methods or operators to measure or collect data. We represented the spatial and temporal autocorrelation of the measurement errors by a negative exponential function:

$$\rho(\Delta s, \Delta t) = e^{-\alpha * \Delta s - \beta * \Delta t} \tag{5.3}$$

where Δs and Δt refer to distance in space in kilometer and distance in time in years, respectively. As before we used three scenarios, because the parameters α and β could only be derived using expert judgement (Table 5.3). We assumed that α and β were the same for all uncertain inputs.

Table 5.3: Parameter values of spatial and temporal autocorrelation function for three robustness scenarios.

Parameter value	Optimistic	Reference	Pessimistic
α (kilometer^{-1})	0.002	0.001	0.0005
β (year^{-1})	0.5	0.2	0.1

With defining the PEs and spatial and temporal autocorrelation functions we had fully characterized the joint distribution of all uncertain inputs. The vector of all input errors had a multivariate normal distribution, whose mean vector was zero and whose variance-covariance matrix could be derived from the autocorrelation functions, the PE values given in Table 5.2, and the measured inputs. Since we assumed that there was no cross-correlation between uncertain inputs, we could work with joint normal distributions for each input separately. Note that the joint distribution was still complex, because the vector of normal variates for each input consisted of 31 x 38 = 1178 elements. We derived the 1178 x 1178 covariance matrices for all inputs as described above and checked that all were positive-definite.

5.2.3. Uncertainty propagation

Propagation of input uncertainty

Monte Carlo method

The Monte Carlo method is by far the most often used tool for uncertainty propagation analysis since it is transparent, easily implemented and generally applicable (Heuvelink, 1998b). It can approximate the probability distribution of the uncertain output at an arbitrary accuracy level as long as the number of simulations is large enough (Heuvelink, 1998a). The Monte Carlo simulation approach starts by sampling a large number of 'possible realities' from the probability distributions of the uncertain inputs, using a pseudo-random number generator. In this chapter, we used 500 Monte Carlo runs. As the joint distribution of the uncertain inputs was multivariate normal, we could use the mvrnorm function of the MASS package (version 7.3-54) (Venables & Ripley, 2002).

Propagation of input uncertainty to NUE calculations

The yields of all crops were summed for each Monte Carlo run to get the total yield of each province and year for that run. The same was done for N removal. N input was already recorded for total crops, hence summation for all crops was not needed. If the summed simulations of total yield or N removal or simulated N input were smaller than zero, then these values were replaced by zero.

The PFP_N and PNB_N values for each Monte Carlo run were next computed from the yield, N removal and N input simulations for each province and year. The frequency distributions of the 500 values of PFP_N and PNB_N represent the propagation of input uncertainty to uncertainty in NUE calculations.

The random forest model

The RF model, developed by (Breiman, 2001), tends to have higher accuracy compared with other machine learning approaches (Hengl et al., 2015). It belongs to the family of ensemble machine learning algorithms that predicts a dependent variable (i.e., NUE indicators) from a set of explanatory variables (training data) selected randomly using a bootstrapping technique (sampling with replacement). Once a model is calibrated, it can be used to predict the dependent variable from only the explanatory variables. The explanatory variables used to build the RF model included crop type, topography, soil, climate, economy and AMP. They were obtained from the Climatic Research Unit (1978-2015), Harris et al. (2014), the National Bureau of Statistics of China (1978-2015) and published articles

(Shangguan et al., 2014; Hengl et al., 2017; Poggio et al., 2020). More details about these covariates are provided in **Chapter 1**. All explanatory variables in raster format were aggregated to provincial scale by taking the provincial average. Crop and soil types were transformed to continuous-numerical variables by computing area proportions. Annual temperatures at night and daytime were also added. In summary, there were 1159 NUE observations (30 provinces from 1978 to 2015; Chongqing province was established in 1997) and 108 explanatory variables.

Propagation of input uncertainty through the NUE random forest model

After obtaining the 500 PFP_N and PNB_N Monte Carlo simulations for each province and year, we fitted the RF (Breiman, 2001) model for each case. In other words, we obtained 500 different RF models because each model was calibrated with a different PFP_N and PNB_N dataset. The differences between these models and their predictions showed how measurement uncertainty in inputs propagated to the model output. If the differences are large, then this indicates that uncertainty in input data (i.e., yield data, human and livestock numbers, chemical fertilizer data and corresponding coefficients) has a large impact on the model output. If it is small, then the effect of input uncertainty on model predictions of NUE indicators is small. We characterized the uncertainty in RF predictions by the 90% Prediction Interval Width (PIW_{90}) and the ratio of the inter-quantile range over the median (Prediction Interval Ratio, PIR_{90}). Note that these uncertainty metrics were computed for every province and year. Their equations are as follows (Poggio et al., 2020):

$$PIW_{90} = q_{0.95} - q_{0.05} \tag{5.4}$$

$$PIR_{90} = \frac{q_{0.95} - q_{0.05}}{q_{0.50}} \tag{5.5}$$

Here, $q_{0.05}$, $q_{0.50}$ and $q_{0.95}$ are the 0.05 quantile, median and the 0.95 quantile of the 500 NUE predictions, respectively.

Model uncertainty

The RF model also causes uncertainty, because the explanatory variables cannot explain all of the variation of the NUE indicators. For instance, in **Chapter 4** we found that the RF model could explain only 84 and 89% of the variation of PFP_N and PNB_N, respectively.

We used QRF to quantify the model uncertainty (Meinshausen, 2006). QRF is similar as RF, but it gives a non-parametric and accurate way of estimating conditional quantiles of the dependent variables (i.e., NUE indicators). It keeps the value of all observations in each node of each tree, not just their mean, and assesses the conditional distribution based on this information. It is invoked in the ranger function of the ranger package (version 0.13.1) with quantreg option set to TRUE. In this case, the prediction is not a single value, i.e., the average of the predictions from the group of decision trees in the random forest, but a cumulative probability distribution of the PFP_N and PNB_N for each province and year. We obtained the 0.05, 0.50 and 0.95 quantiles from this distribution to compute the PIW_{90} and PIR_{90} associated with model uncertainty.

5.2.4. Uncertainty sources contributions

Both input uncertainty and model uncertainty lead to uncertainty in RF predictions of NUE indicators. In previous sections we have explained how these uncertainties can be derived, by using Monte Carlo simulation for input uncertainty and QRF for model uncertainty. The magnitude of the uncertainties was quantified with PIW_{90} and PIR_{90}. Once these are obtained it is an easy step to compare them. We computed the uncertainty contributions of each source by dividing each by the sum. Such analysis can provide highly relevant information. For instance, if input uncertainty is the principal uncertainty source (with larger PIW_{90} or PIR_{90}), then it is important to improve the accuracy of the yield data, human and livestock numbers, chemical fertilizer data and corresponding coefficients. If model uncertainty is the main uncertainty source, then there is little gain in putting a large effort in collecting more accurate data and coefficients. Instead, improvements can best be achieved by obtaining more and better explanatory covariates and models. Note that the uncertainty contributions were computed for all provinces and years and can be different between provinces and years.

5.3. Results

In this section we will only present results of the uncertainty analysis. Spatial patterns and time series of PFP_N and PNB_N calculations and RF model predictions are given in (**Chapter 2**) and (**Chapter 4**). Due to space limitations, we will concentrate on three provinces (Heilongjiang, Henan and Sichuan) and the year 2015 when presenting results in this section. Results for other provinces and the entire period are mostly presented in the Appendix A. We selected these three provinces because these are main food productive provinces in China, while their environmental conditions and geographical locations are different and fairly representative for the whole of China. The results for other provinces and scenarios can be found in Figures E.1 and E.2.

5.3.1. Uncertainty of NUE calculations for different scenarios

We used probability distributions width between the quantile 0.05 and 0.95 to describe the uncertainty of PFP_N and PNB_N calculations for each province in each year. The probability distribution is estimated by the frequency distribution of the PFP_N and PNB_N calculations that were derived from simulating 500 Monte Carlo realizations of yield, N removal, and N input for each province and year and computing PFP_N and PNB_N using the equations given in Table 5.1. Results showed that as expected that the PFP_N and PNB_N uncertainty of the reference scenario is larger than that of the optimistic and smaller than that of the pessimistic scenario (Figure 5.1a and b). The differences between the three scenarios are large, since the probability densities are much narrower for the optimistic scenario and much wider for the pessimistic scenario than for the reference scenario. The three selected provinces had similar distribution width between the quantile 0.05 and 0.95 for PFP_N (3, 6 and 12 kg kg^{-1} for the optimistic, reference and pessimistic scenarios, respectively) (Figure 5.1). For PNB_N, the distribution width of Heilongjiang (0.15, 0.3 and 0.6 kg kg^{-1} for the optimistic, reference and pessimistic scenarios, respectively) is three times that of Henan and Sichuan (Figure 5.2). All distributions had small skewness and did not deviate much from normal distributions.

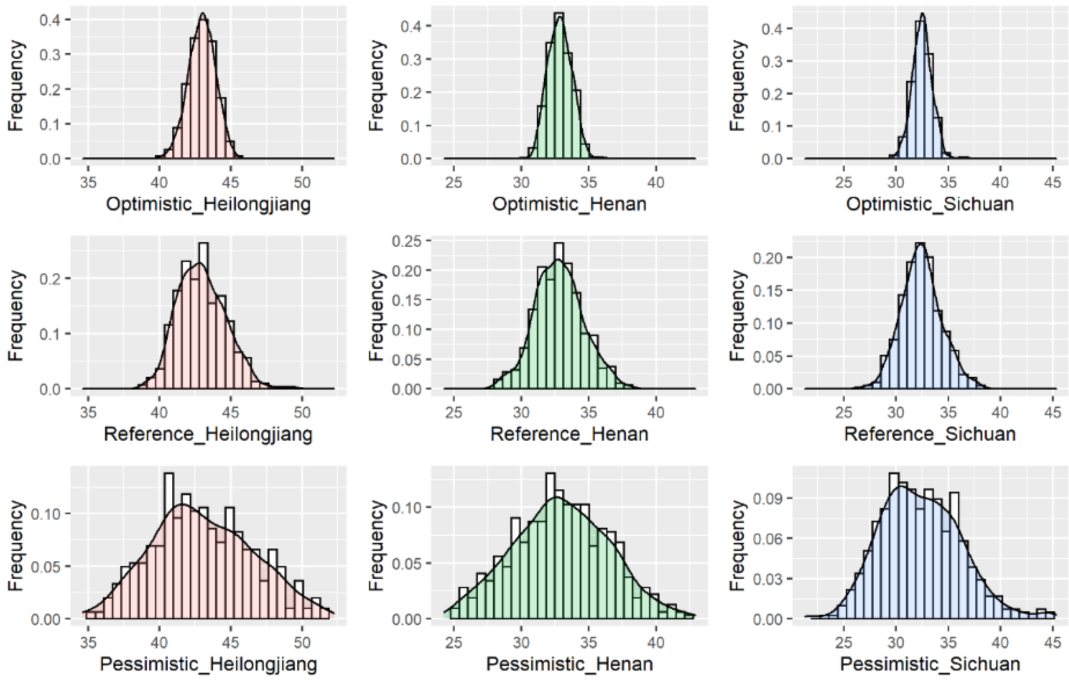


Figure 5.1: Propagation of uncertainty in crop yield, N removal and N input measurements to calculated partial factor productivity (PFP_N) in Heilongjiang, Henan and Sichuan provinces for the optimistic (top), reference (middle) and pessimistic (bottom) scenarios in 2015. Histograms show frequency distributions of 500 Monte Carlo simulations. The kernel density estimate (continuous line) was obtained with geom_density function of the ggplot2 package (version 3.3.5) in R. The x-axis unit is kg kg^{-1}.

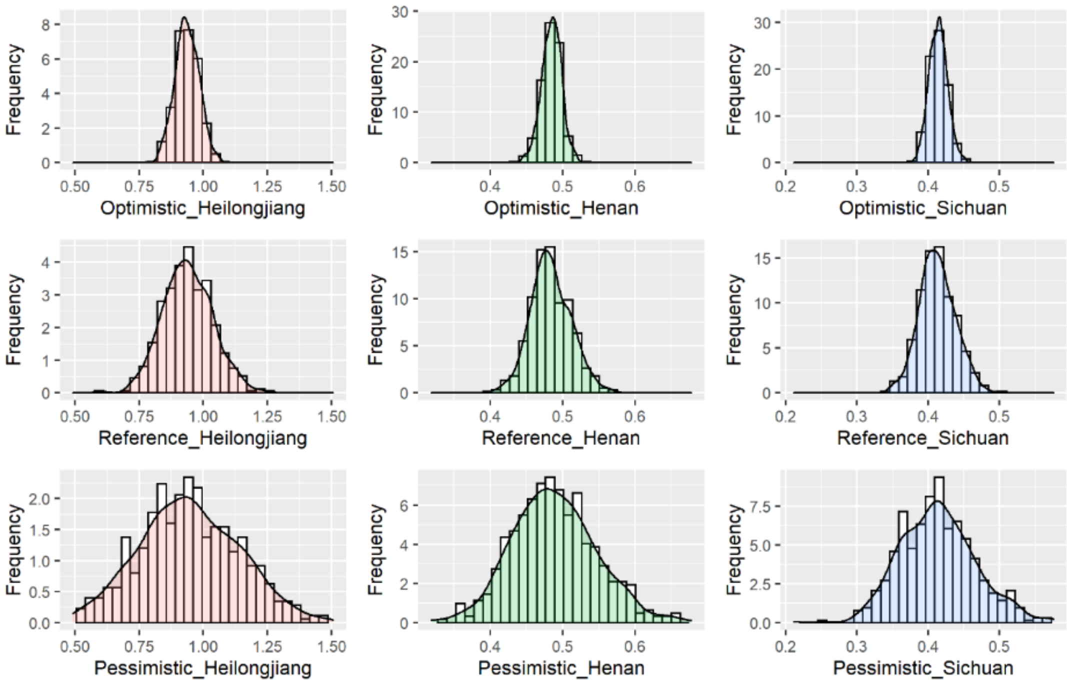


Figure 5.2: Propagation of uncertainty in crop yield, N removal and N input measurements to calculated partial nutrient balance (PNB_N) in Heilongjiang, Henan and Sichuan provinces for the optimistic (top), reference (middle) and pessimistic (bottom) scenarios in 2015. Histograms show frequency distributions of 500 Monte Carlo simulations. The kernel density estimate (continuous line) was obtained with geom_density function of the ggplot2 package (version 3.3.5) in R. The x-axis unit is kg kg^{-1}.

For PFP_N, Guangxi and Shanghai had the largest distribution width between the quantile 0.05 and 0.95 (19 and 14 kg kg^{-1} in reference scenario, respectively), while Jilin and Inner Mongolia (both were 4 kg kg^{-1} in reference scenario) had the smallest distribution width (Figures E.1 and E.2). For PNB_N, Heilongjiang and Jilin (0.3 and 0.2 kg kg^{-1} in the reference scenario, respectively) had the largest distribution width, while Beijing and Hainan (both were 0.06 kg kg^{-1} in reference scenario) had the smallest distribution width.

The temporal correlation of each NUE indicator for each province is shown for Heilongjiang, Henan and Sichuan provinces among several years in the appendix (Figures E.3 and E.4). The results show a higher temporal correlation in the pessimistic scenario and a lower correlation in the optimistic scenario. This is as expected because the inputs to the calculations have a higher and lower temporal correlation in the pessimistic and optimistic scenarios, respectively (Table 5.3). The correlation decreases with increasing time lag. The NUE indicators are autocorrelated up to about four years in the pessimistic scenario, two years in the reference scenario, and only about one year in the optimistic scenario. The temporal autocorrelation in Sichuan was higher than in the other two provinces (Figures E.3c and E.4c).

5.3.2. Propagation of input calculations uncertainty to RF model output

The 500 PFP_N and PNB_N calculations obtained with the Monte Carlo method and presented in Section 5.3.1 were used to calibrate 500 RF models. Thus, the differences in the outputs of these 500 RF models reflect how uncertainty in the measured yield, N input and N removal propagate through a RF model that predicts PFP_N and PNB_N from environmental covariates. This was again done for the three scenarios and for all provinces and years. In this section we present and interpret time series and spatial maps of the uncertainties of these RF model outputs.

PFP_N uncertainty

The magnitude and temporal variation of PFP_N prediction uncertainty in Heilongjiang, Henan and Sichuan provinces is shown in Figure 5.3 (PIW_{90}) and 5.4 (PIR_{90}). Overall, the PIW_{90} in the three provinces decreased with time until 2000, after which it was stable in Heilongjiang and increased in Henan and Sichuan provinces (Figure 5.3). The PIW_{90} was higher in Heilongjiang than in the other two provinces before 2005. The PIR_{90} in Heilongjiang and Henan had a decreasing trend while it decreased before and increased after 2000 in Sichuan (Figure 5.4). The PIR_{90} was largest in Heilongjiang and smallest in Sichuan in 1978, while it was the converse in 2015. The temporal variation of PFP_N prediction uncertainty in 31 provinces in China is shown in Figures E.5 (PIW_{90}) and E.6 (PIR_{90}). The temporal variation of PIW_{90} was different in different provinces and most provinces had a decreasing trend before 2000, after which it increased or stayed stable, except for Tianjin province, which had a decreasing trend (Figure E.5). The PIR_{90} in most provinces had a declining trend over time (Figure E.6).

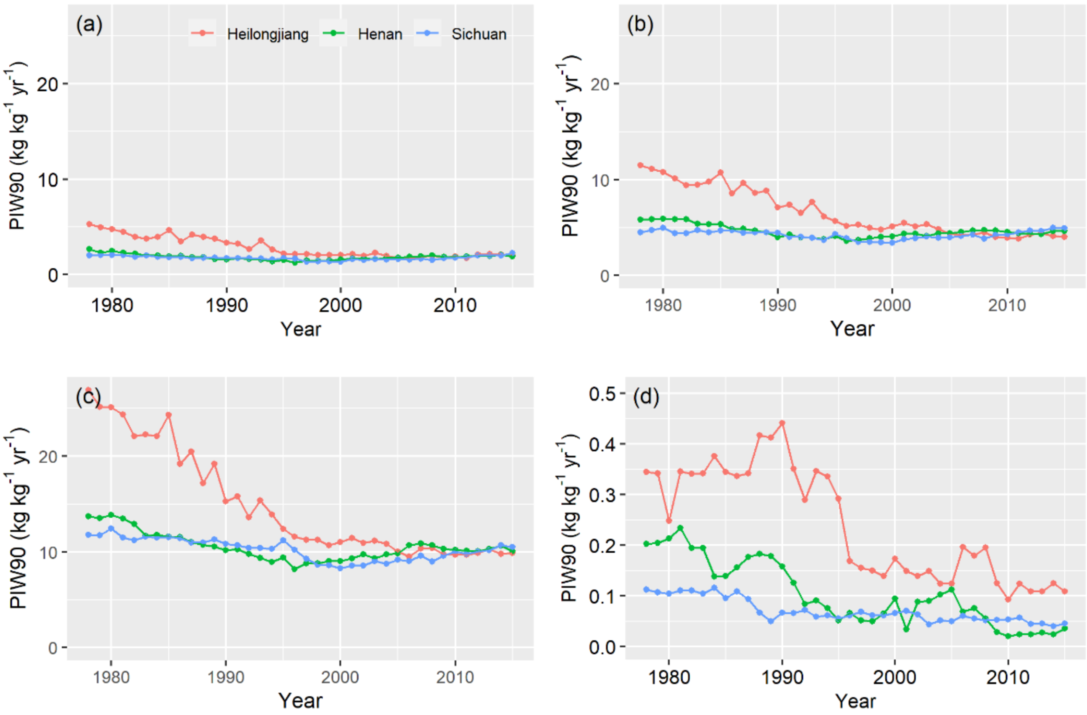


Figure 5.3: Time series from 1978 to 2015 of the partial factor productivity (PFP_N) 90% prediction interval width (PIW_{90}) of the random forest model outputs in Heilongjiang, Henan and Sichuan provinces, as resulting from measurement uncertainty in input data for the optimistic (a), reference (b), and pessimistic (c) scenarios, and as resulting from model uncertainty (d).

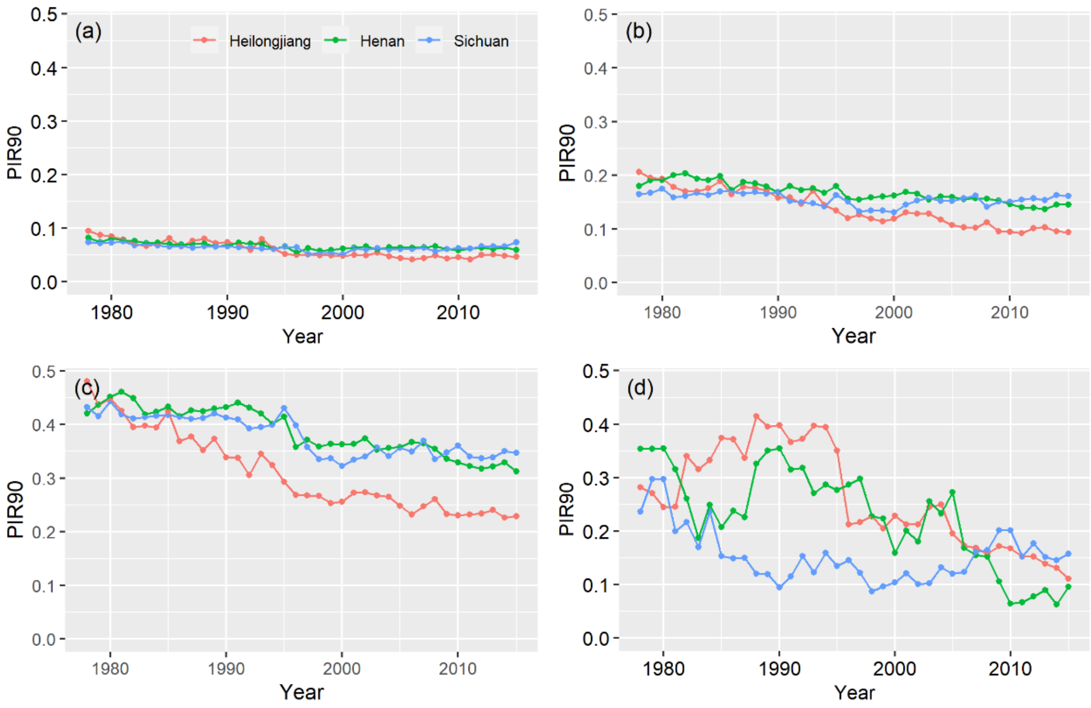


Figure 5.4: Time series from 1978 to 2015 of the partial factor productivity (PFP_N) 90% prediction interval ratio (PIR_{90}) of the random forest model outputs in Heilongjiang, Henan and Sichuan provinces, as resulting from measurement uncertainty in input data for the optimistic (a), reference (b), and pessimistic scenarios (c), and as resulting from model uncertainty (d).

The spatial variation of PFP_N prediction uncertainty in Heilongjiang, Henan and Sichuan provinces is shown in Figures 5 (PIR_{90}) and E.7 (PIW_{90}). PIR_{90} was higher in south and west China than in other regions in 2015 (Figure 5.5). The average PIR_{90} from 1978 to 2015 in Tianjin, Shanghai, Hainan and Beijing was higher than in other provinces. The average PIR_{90} in Inner Mongolia and Heilongjiang was lower than in other provinces. In 2015, PIR_{90} was lower than 0.10 in the optimistic scenario (Figure 5.5a); between 0.10 and 0.20 in the reference scenario for most provinces, except for Heiloangjiang, where it was 0.09 (Figure 5.5b); and between 0.20 and 0.40 in the pessimistic scenario (Figure 5.5c). For PIW_{90}, south and east-costal China had higher values than in other regions in 2015 (Figure E.7). The average PIW_{90} from 1978 to 2015 in Tianjin, Guangxi and Hainan was higher than in other provinces. The average PIW_{90} in Inner Mongolia and Gansu was lower than in other provinces. In 2015, the PIW_{90} of most provinces in the optimistic scenario were smaller than 5 kg kg^{-1} yr^{-1}, except for Guangxi (5.6 kg kg^{-1} yr^{-1}) (Figure E.7a). In addiation, it was lower than 10 kg kg^{-1} yr^{-1} in the reference scenario, except Guangxi (12.4 kg kg^{-1} yr^{-1}) and Shanghai (10.5 kg kg^{-1} yr^{-1}) (Figure E.7b). As expected, the largest uncertainty of PIW_{90} occured in the pessimistic scenario (7-28 kg kg^{-1} yr^{-1}) (Figure E.7c).

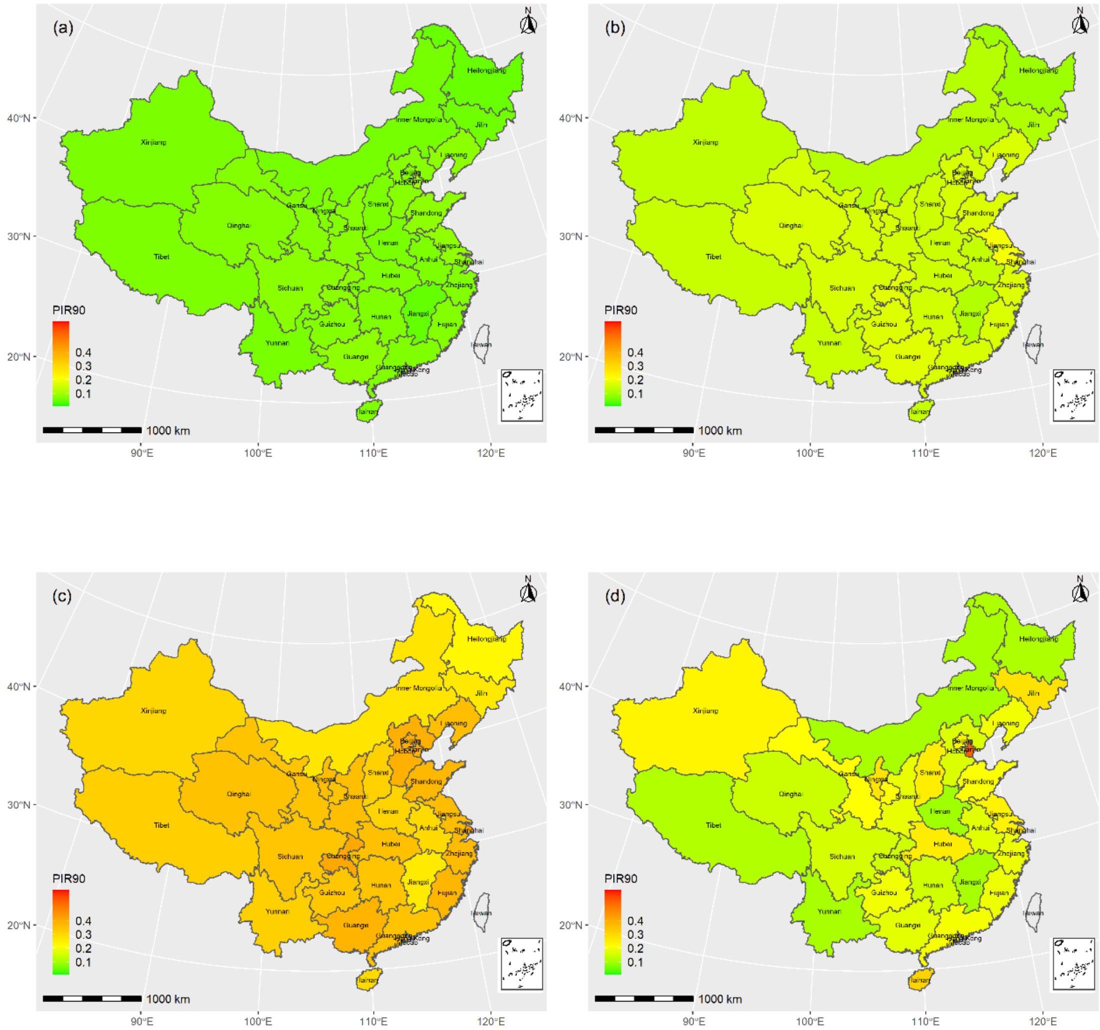


Figure 5.5: Spatial distribution of the 90% prediction interval ratio (PIR_{90}) of partial factor productivity (PFP_N) in 31 provinces of China in 2015, in the optimistic (a), reference (b) and pessimistic scenarios (c), obtained by propagating measurement uncertainty in crop yield, N removal and N input measurements through the RF model. Map of PIR_{90} caused by model uncertainty as quantified by quantile regression forests (d).

PNB_N uncertainty

The magnitude and temporal variation of PNB_N prediction uncertainty for the three input uncertainty scenarios in Heilongjiang, Henan and Sichuan provinces is shown in Figures 5.6 (PIW_{90}) and 5.7 (PIR_{90}). The PIW_{90} in Heilongjiang, Henan and Sichuan generally decreased over time. After 2000 it increased dramatically in

Heilongjiang and increased slightly in Henan and Sichuan. Overall, the PIW_{90} was larger in Heilongjiang and smaller in Sichuan than in Henan (Figure 5.6). The PIR_{90} showed a downward trend in Henan and Sichuan and fluctuations in Heilongjiang, but it went up after 2000. The PIR_{90} in Heilongjiang was lower than in the other two provinces before 2000, and larger after 2000. It should be noted that Sichuan had higher PIR_{90} than Henan all the time (Figure 5.7). The temporal variation of PIW_{90} and PIR_{90} in 31 provinces for PNB_N was similar as that for PFP_N (Figures E.8 and E.9).

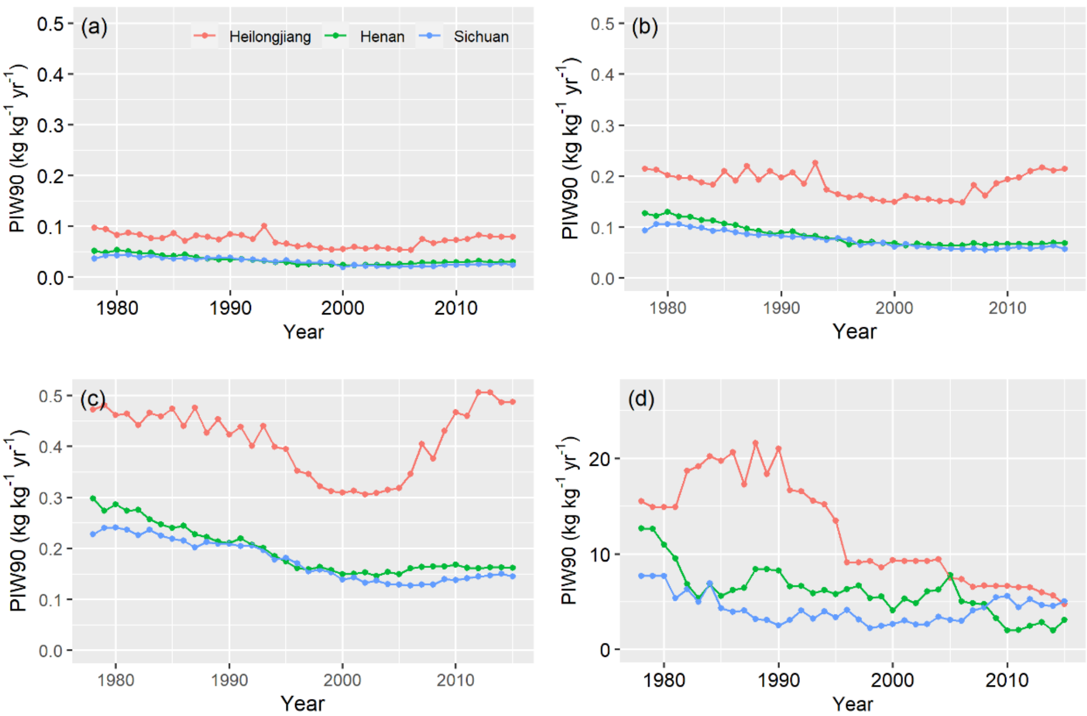


Figure 5.6: Time series from 1978 to 2015 of the 90% prediction interval width (PIW90) of partial nutrient balance (PNB_N) for Heilongjiang, Henan and Sichuan provinces in the optimistic (a), reference (b), and pessimistic scenario (c), obtained by propagating measurement uncertainty in crop yield, N removal and N input measurements through the RF model. Times series of PIW90 as caused by model uncertainty and quantified by quantile regression forests (d).

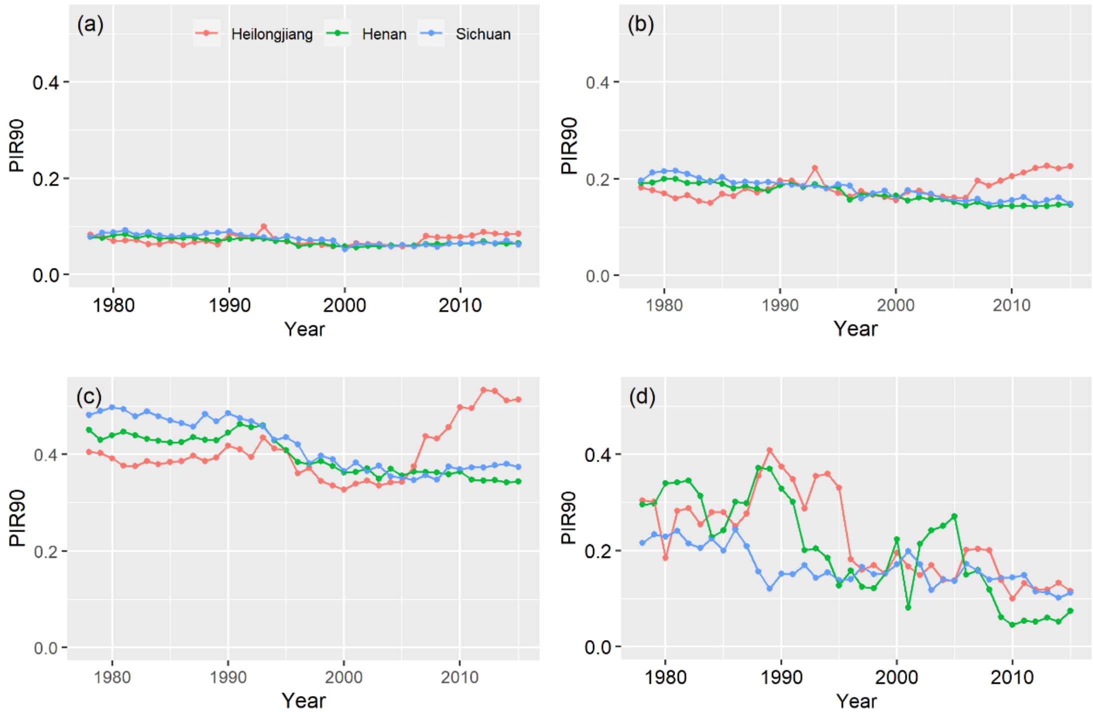


Figure 5.7: Time series from 1978 to 2015 of the 90% prediction interval ratio (PIR_{90}) of partial nutrient balance (PNB_N) for Heilongjiang, Henan and Sichuan provinces in the optimistic (a), reference (b), and pessimistic scenario (c), obtained by propagating measurement uncertainty in crop yield, N removal and N input measurements through the RF model. Times series of PIR_{90} as caused by model uncertainty and quantified by quantile regression forests (d).

For PNB_N, the PIR_{90} in 2015 was higher in northeast China than in other regions (Figure 5.8). In Jilin and Liaoning province, they were 0.11 and 0.09 in the optimistic, 0.26 and 0.22 in the reference and 0.61 and 0.53 in the pessimistic scenarios, respectively. The PIR_{90} in Gansu and Tibet was lower than in other provinces. They were 0.06 and 0.06 in the optimistic, 0.14 and 0.13 in the reference, and 0.32 and 0.31 in the pessimistic scenario, respectively. For the optimistic scenario, the PIR_{90} was lower than 0.1, except in Tianjin, Inner Mongolia and Jilin. The PIR_{90} of most provinces in the reference scenario was between 0.1 and 0.2. In the pessimistic scenario, most provinces had PIR_{90} within 0.2 and 0.5, apart from Jilin (0.61), Inner Monglia (0.54), Liaoning (0.52) and Heilongjiang (0.51) provinces. The PIW_{90} had similar spatial distribution patterns as PIW_{90} in 2015 (Figure E.10). Specifically, the annual average PIW_{90} from 1978 to 2015 in Heilongjiang and Jilin was higher than in other provinces and were 0.07 and 0.06 kg kg^{-1} yr^{-1} in the optimistic, 0.19 and 0.16 kg kg^{-1} yr^{-1} in the reference and 0.41 and 0.36 kg kg^{-1} yr^{-1} in the pessimistic scenarios, respectively. The average PIW_{90} in Xinjiang and Gansu was lower than in other provinces. These were 0.02

and 0.02 kg kg^{-1} yr^{-1} in the optimistic, 0.06 and 0.06 kg kg^{-1} yr^{-1} in the reference, and 0.13 and 0.14 kg kg^{-1} yr^{-1} in the pessimistic scenarios, respectively. In 2015, the PIW_{90} of most provinces in the optimistic scenario were smaller than 0.1 kg kg^{-1}. It was also lower than 0.1 kg kg^{-1} in the reference scenario, except for Heilongjiang (0.21 kg kg^{-1}), Jilin (0.13 kg kg^{-1}) and Inner Mongolia (0.11 kg kg^{-1}). It was between 0.1 and 0.3 kg kg^{-1} in the pessimistic scenario, except for Hainan (0.096 kg kg^{-1}), Heilongjiang (0.49 kg kg^{-1}) and Jilin (0.34 kg kg^{-1}).

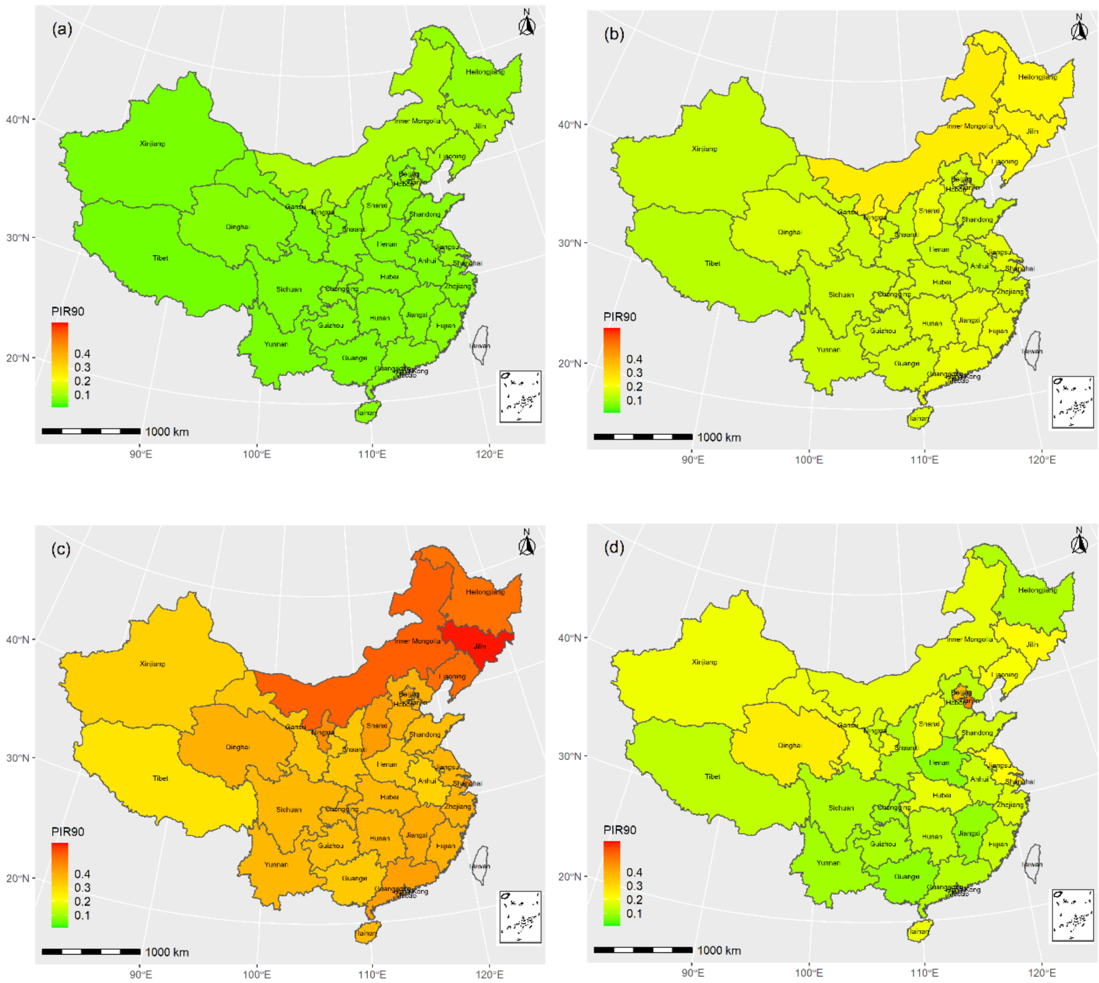


Figure 5.8: Spatial distribution of the 90% prediction interval ratio (PIR_{90}) of partial nutrient balance (PNB_N) in 31 provinces of China in 2015, in the optimistic (a), reference (b) and pessimistic scenarios (c), obtained by propagating measurement uncertainty in crop yield, N removal and N input measurements through the RF model. Map of PIR_{90} caused by model uncertainty as quantified by quantile regression forests (d).

5.3.3. Model uncertainty

PFP_N model uncertainty

The PIW_{90} of the QRF model in Heilongjiang was higher than that in the other two provinces (Figure 5.3d). It was higher in Henan than in Sichuan from 1978 to 2008, but reversed from 2008 to 2015. For PIR_{90}, Heilongjiang had considerably higher values than Henan and Sichuan during 1982-1995 (Figure 5.4d). Henan had a lower model uncertainty PIR_{90} than Sichuan from 2007 onwards. The temporal variation of model uncertainty (PIW_{90} and PIR_{90}) had more fluctuations than that of input uncertainty (Figures E.5 and E.6). Note that the uncertainty resulting from model uncertainty was larger than that obtained in the optimistic and reference input error scenarios, but smaller than that of the pessimistic input error scenario (see also Section 5.3.4).

The PIR_{90} for the QRF model in northwest and east China was higher than in other regions in 2015 (Figure 5.5d). The average PIR_{90} in Shanghai and Tianjin was higher than in other provinces from 1978 to 2015 and were 0.42 and 0.40, respectively. The PIR_{90} in Yunnan, Sichuan and Jiangxi was lower than in other provinces and was 0.13, 0.16 and 0.16, respectively. In 2015, Tianjin (0.54) and Hainan (0.31) had a high PIR_{90}, while Henan had the lowest value (0.10). The PIW_{90} was higher in northwest and southeast China in 2015 (Figure E.7d). At provincial scale, the annual average PIW_{90} from 1978 to 2015 in Tianjin (17.3 kg kg^{-1} yr^{-1}), Guangxi (16.1 kg kg^{-1} yr^{-1}) and Hainan (14.7 kg kg^{-1} yr^{-1}) was higher than in other provinces. The PIW_{90} in Yunnan and Sichuan was lower than in other provinces and was 3.8 and 4.3 kg kg^{-1} yr^{-1}, respectively. In 2015, the PIW_{90} in Tianjin, Shanghai, Guangxi, Hainan and Shanxi was larger than 10 kg kg^{-1}, while Inner Mongolia, Henan, Tibet, Yunnan, Jiangxi and Heilongjiang had smaller values (< 5 kg kg^{-1}).

PNB_N model uncertainty

The PIW_{90} for the QRF model in Heilongjiang was higher than in other provinces (Figure 5.6d). It was higher in Henan than in Sichuan from 1978 to 2008, but reversed from 2008 to 2015. Heilongjiang and Henan had higher PIR_{90} than Sichuan before 1995, and thereafter Heilongjiang had similar PIR_{90} as Sichuan, while Henan had the lowest PIR_{90} (Figure 5.7d). The model uncertainty (PIW_{90} and PIR_{90}) had a decreasing trend over time and showed more fluctuations than uncertainty caused by input measurements (Figures E.8 and E.9).

The PIR_{90} in 2015 was higher in north China than in south China in 2015 (Figure 5.8d). The average PIR_{90} from 1978 to 2015 in Tianjin, Shanghai, Beijing and Jilin was higher than in other provinces (> 0.3). They were 0.42, 0.39, 0.32 and 0.30, respectively. The average PIR_{90} in Guangxi, Sichuan, Jiangxi and Chongqing was lower than in other provinces. They were 0.14, 0.17, 0.17 and 0.17, respectively. In 2015, Tianjin (0.49) and Beijing (0.36) had a high PIR_{90}, while Henan, Guangxi, Jiangxi, Chongqing and Yunnan had lower values (< 0.1). For PNB_N, the PIW_{90} of the QRF model was higher in north and west China in 2015 (Figure E.10d). The annual average PIW_{90} from 1978 to 2015 in Heilongjiang (0.24 kg kg^{-1} yr^{-1}) and Tianjin (0.21 kg kg^{-1} yr^{-1}) was higher than in other provinces. The average PIW_{90} in Chongqing, Guangxi and Guangdong was lower than in other provinces (< 0.06 kg kg^{-1} yr^{-1}). In 2015, the PIW_{90} in Tianjin, Qinghai, and Jilin was larger than 0.12 kg kg^{-1}, while Guangxi, Guangdong and Yunnan had smaller values (< 0.03 kg kg^{-1}).

5.3.4. Uncertainty source contributions

The NUE uncertainty was caused by input measurement uncertainty and model uncertainty. In this section we compare the contributions of these two uncertainty sources for the reference scenario case, and using PIR_{90} as uncertainty metric. Figures 5.9, 5.10 and E.11 show the input uncertainty contributions, presented as a percentage of the total uncertainty (i.e., the sum of the PIR_{90} of both sources). Input uncertainty is the dominant source of uncertainty if it is above the horizontal dashed lines in Figures 5.9 and E.11, otherwise model uncertainty has a bigger contribution.

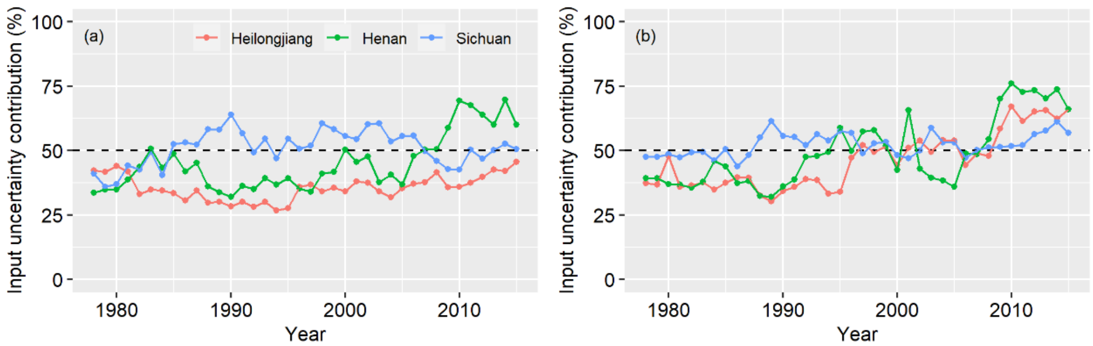


Figure 5.9: Time series from 1978 to 2015 of input uncertainty contributions (90% prediction interval ratio in reference scenario) of partial factor productivity (PFP_N, a) and partial nutrient balance (PNB_N, b) in Heilongjiang, Henan and Sichuan provinces. Dashed line represents a case in which input uncertainty and model uncertainty have equal contribution.

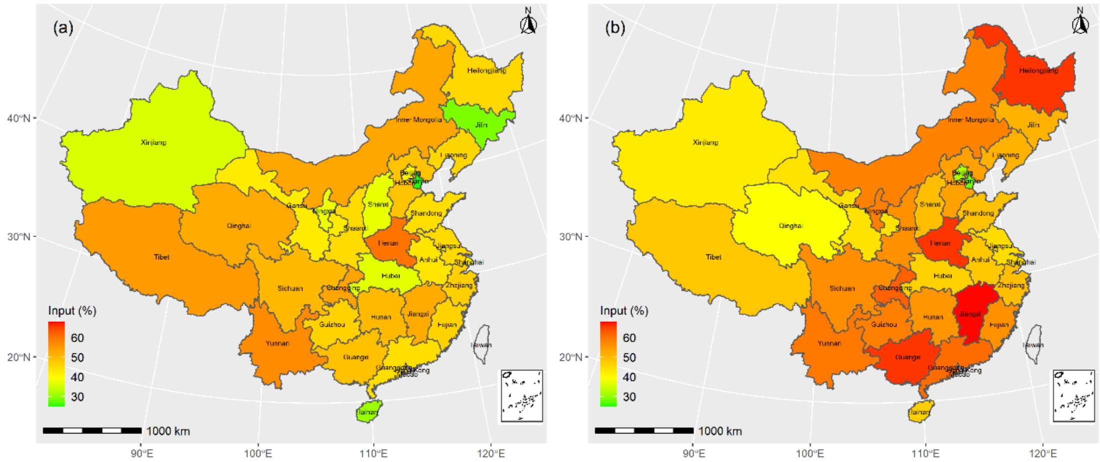


Figure 5.10: Spatial distribution of input uncertainty contributions (90% prediction interval ratio in reference scenario) for partial factor productivity (PFP_N, a) and partial nutrient balance (PNB_N, b) in 31 provinces in 2015.

Uncertainty sources contribution for PFP_N

The input uncertainty contribution in Heilongjiang was higher than in Henan and Sichuan until 1981, after which it was lower than in the other provinces (Figure 5.9a). The input uncertainty contribution in Sichuan was higher than in Henan at first but after 2008 it was lower than that in Henan. The input uncertainty contribution in Heilongjiang was smaller than the model uncertainty contribution during the entire period from 1980 to 2015, while Sichuan and Henan had an input uncertainty contribution higher than 50% in the 1985-2006 and 2008-2015 periods, respectively. The contributions of input uncertainty were within 25 and 60% from 1978 to 2015 for most provinces; most provinces had an input uncertainty contribution smaller than 50% (Figure E.11a). There was also a large variation in spatial patterns of input uncertainty contribution. The average input uncertainty contribution from 1978 to 2015 was highest in Yunnan (56%), Beijing (54%) and Fujian (54%), and lowest in Jilin (33%) and Ningxia (33%). In 2015, the input uncertainty contribution in Henan, Yunnan and Tibet was larger than 55%, while Tianjin and Jilin had the smallest values (< 30%) (Figure 5.10a).

Uncertainty sources contribution for PNB_N

The input uncertainty contribution in Sichuan was higher than in other provinces before 1995, and it was lower than in other provinces after 2009 (Figure 5.9b). The temporal variation trend of input uncertainty contribution in Henan and Heilongjiang was very similar: fluctuating and on average upward. The input uncertainty contribution in Sichuan was close to or higher than 50% over time, while Henan and Heilongjiang mostly had input uncertainty contribution higher than 50% after 1995. The contributions of input uncertainty were within 30 and 60% in 1978 for most provinces and increased to the range of 40-75% in 2015; most provinces had input uncertainty contributions lower than 50% before 2010 (Figure E.11b). Large variation existed in spatial distribution of input uncertainty contribution (Figure 5.10b). The annual average input uncertainty contribution from 1978 to 2015 was higher in Jiangxi (60%), Guangxi (59%) and Fujian (58%), and lower in Tianjin (33%) and Tibet (35%). In 2015, the input uncertainty contribution in Jiangxi, Henan, Guangxi and Heilongjiang was higher than 66%, while it was lower than 40% in Tianjin, Beijing and Qinghai (Figure 5.10b).

5.4. Discussion

5.4.1. Uncertainty of NUE predictions

Admittedly, most of the uncertainty assessment of input data was based on expert knowledge. This makes them unreliable, and we addressed this problem by using three scenarios. The results indicated large difference among scenarios, which meant that the results of the uncertainty analysis are quite sensitive to the quantification of input uncertainty. Therefore, it is important that the government pays more attention to assessing uncertainty in the inputs to the model and NUE calculations. Moreover, yield aggregation at spatial and crop scale leads to additional input uncertainty. Porwollik et al. (2017) demonstrated that aggregation uncertainty of gridded yields at national scale can be up to 37% (maize, South Africa), 43% (wheat, Pakistan), 51% (rice, Japan), and 427% (Soybean, Bolivia). On the other hand, crop aggregation leads to uncertainty of provincial yield, since there are large differences between dry matter crop yield and fresh weight crop yield. Zhang et al. (2021a) reported that uncertainties of N inputs and outputs were 8% and 12%, respectively. Chemical fertilizer, as the most important N input, is uncertain because of uncertainty in fertilizer rate collected and N ratio in compound fertilizer. For manure, process wastewater (water that contacts milk, feed, manure, or cattle during dairy operations) was the dominant source of uncertainty in manure application N, contributing 64-94% to the overall uncertainty (Miller et al., 2020). Other uncertainty might arise from N content in manure and livestock numbers. These parameters used in our calculations lack validation and may have large uncertainty.

Interestingly, model uncertainty had similar spatial and temporal trends as input uncertainty. A plausible explanation is that the development of science and technology improve the accuracy of input data and model simultaneously. Temporal variation of model uncertainty had more fluctuations than input uncertainty during the study period. This might be because the annual discrepancies in some explanatory variables were neglected due to lack of information. This leads to a weaker relation between the target variable (i.e., NUE) and explanatory variables, which means a larger model uncertainty in time. But it is difficult to determine which explanatory variables cause this effect since the QRF model is a highly complex model in which it is difficult to see the effect of covariates on the predictions. Wang et al. (2017) and Lobell and Field (2007) indicated that the uncertainty of temperature measurement or prediction will influence the accuracy of crop yield prediction. Additionally, the explanatory variables could not explain all variation in NUE indicators because crucial explanatory variables were lacking, such as fertilizer management information. Model uncertainty also increased because we did not have information about water supply condition, which is known to have a significant influence on NUE indicators (Lemaire & Ciampitti, 2020). (Helfenstein et al., 2022) and (Hounkpatin et al., 2022) indicated that the QRF model may slightly overestimate the prediction uncertainty. Therefore, other advanced models should also be applied in order to find a more suitable model for NUE uncertainty. Nigon et al. (2020) showed that the performance among Lasso, SVR, and PLSR was comparable when predicting N uptake in maize, while the performance of random forest was substantially inferior with higher error values.

We found that the temporal and spatial variation of NUE uncertainty (PIW_{90}) (Figures E.5, E.7, E.8 and E.10) from uncertainty in input data was related to the variation of PFP_N values as shown in Figure 2.1a in **Chapter 2**. This phenomenon was most obvious in the pessimistic scenario. A possible reason for the relation is that our assumption for parameters/statistical data for computing input data (crop yield, N removal and N input) was based on proportional errors. This effect shows up in the PIW_{90} but not in the PIR_{90}, because PIR_{90} is a relative error metric (with median in the denominator).

Overall, PIR_{90} decreased over time. This may be explained from the improvement of technology and statistical policy over the years. A very interesting result was that northeast China had a lower PIR_{90} for PFP_N, but a higher PIR_{90} for PNB_N in 2015. Since these were calculated by the same N input, the difference can only be explained from differences in crop yield and N removal uncertainties. An explanation is that the uncertainty of crop yield was lower in northeast China, while

N removal uncertainty was higher in northeast China (Figure E.12). That might be caused by the crop types that are different between these regions. Northeast China has much less fruits, vegetables and melons, which were high yield uncertainty crops.

5.4.2. Contribution of uncertainty sources to NUE prediction uncertainty

In this chapter, the input uncertainty contribution in the reference scenario had a similar distribution range for PFP_N and PNB_N. Overall, input uncertainty had a somewhat lower contribution than model uncertainty in most provinces, except for PNB_N after 2010 (Figure E.11). The results also showed that input uncertainty had a larger contribution over time, possibly because the QRF model improved over time (i.e., the explanatory variables were more accurate over time). This also meant that the QRF model was more competitive for PNB_N in the 2010s. We should focus on improving the accuracy of explanatory variables or the advancement of models to decrease the model uncertainty propagation to PFP_N, while simultaneously improving the input data accuracy when computing PNB_N. Both investments pay off because both sources of uncertainty have a substantial contribution.

It is worth noting that Tianjin and Jilin had smaller input uncertainty for PFP_N. In other words, in these cases the model uncertainty was limiting. More suitable models need to be applied in these provinces if we cannot collect more explanatory variables to improve the model performance. For PNB_N, provinces in south and northeast China had higher input uncertainty than model uncertainty. This might be because N removal had larger uncertainty in northeast China with higher uncertainty crops (e.g., rice, maize). The more developed provinces in south China have better economy and technology, which might improve the accuracy of the explanatory variables. Moreover, there are more crop types and seasons in south China, which leads to high uncertainty in N removal and N input. Del Grosso et al. (2010) showed that model uncertainty and input data accounted for 83% and 17% of the total uncertainty of N_2O emission, respectively. In such case, it is best to put more effort in model improving. In our study, both uncertainty sources had large contributions in the recent past. Therefore, it is advised to put more effort into obtaining reliable information about statistical data and crop parameters as well as improving the model performance.

5.4.3. Recommendations for reducing NUE prediction uncertainty

Importance of reducing NUE uncertainty

NUE is an important indicator of scientific and efficient fertilization. Quantifying NUE prediction accurately is beneficial for improved fertilization and mitigating environmental pollution. If the NUE can be accurately predicted, then we can better determine the optimal N fertilizer rate and fertilization tactics. This is because with an accurate NUE prediction, the N input can be calculated accurately as well based on the expected yield and N removal, which are relatively stable for a given crop and can be obtained from historical data. Thus, accurate information about NUE can enhance productivity of crops and decrease the fertilization costs of farmers (Lobell, 2007b). Meanwhile, the fertilization model could be developed and optimized. Wang and Zhou (2014) indicated that suitable NUE calculation methods could decrease the NUE uncertainty and were favorable to guiding fertilization and management. Such analyses could prevent excessive N application and waste of N resources in China. Reducing NUE uncertainty is also helpful to preventing environmental pollution and taking alleviation measures. For example, even though the PNB_N is higher in northeast China than in other parts of the country (**Chapter 2**), the NUE uncertainty is also higher in this area. Therefore, we cannot be certain that northeast China has less environmental pollution. We should reduce the NUE prediction uncertainty in northeast China, using strategies outlined before.

Recommendations for decreasing NUE uncertainty

Most studies on uncertainty related to crop research assess uncertainty through treatment replications (e.g., coefficient of variation, standard deviation). This assesses the uncertainty about the response of crops to fertilizers, but does not represent uncertainty in measurements (Yang et al., 2017). As far as we know, this chapter is the first attempt to quantify the NUE uncertainty at provincial scale, while including spatial and temporal correlation between measurement errors, which is crucial in regional scale assessments and when aggregating results to national level. But we still need more efforts to improve the NUE uncertainty sources, both from input data and model. Firstly, it is essential to formulate standard procedures and guidelines about data collection and defining their associated uncertainties, especially for key parameters (e.g., crop N content, livestock excretion, crop-specific fertilizer application, N ratio in compound fertilizer). Secondly, important management practices need to be recorded by the government, such as irrigation water volume (only including irrigation area so far), manure nutrient content for agriculture, fertilizer times, fertilizer date. These may serve as covariates in an improved machine learning model. Thirdly, the technological advancement of experimental observations and improved bookkeeping of human activity data must be achieved by the statistical bureau, and be shared publicly in a clear and accessible format, to reduce errors in secondary inputs. Finally, we should build a scientists-farmers network. Scientists should cooperate with farmers and train them to finish the survey accurately and online. They can also provide better management practices for farmers in order to increase NUE. In this case, we can get more accurate field parameters and enhance NUE accuracy. Although these recommendations cannot be completely achieved within a short term, to realize them in the long run they should be put on the agenda.

Outlook

The results from uncertainty analyses are often scale dependent (Heuvelink et al., 1989). This implies that results obtained for one scale cannot directly be extrapolated to other scales (Nol et al., 2010). Therefore, we call for more research about NUE uncertainty at finer resolution and enhance the basic agricultural data statistics work by building a multi-scale data quality management system. Meanwhile, future research should focus on field work validation, and further explore other NUE indicators (e.g., agronomic efficiency, recovery efficiency). In addition, crop specific NUE uncertainty should be more valuable when guiding fertilizer application, especially when provinces differ largely in crop types and area.

5.5. Conclusion

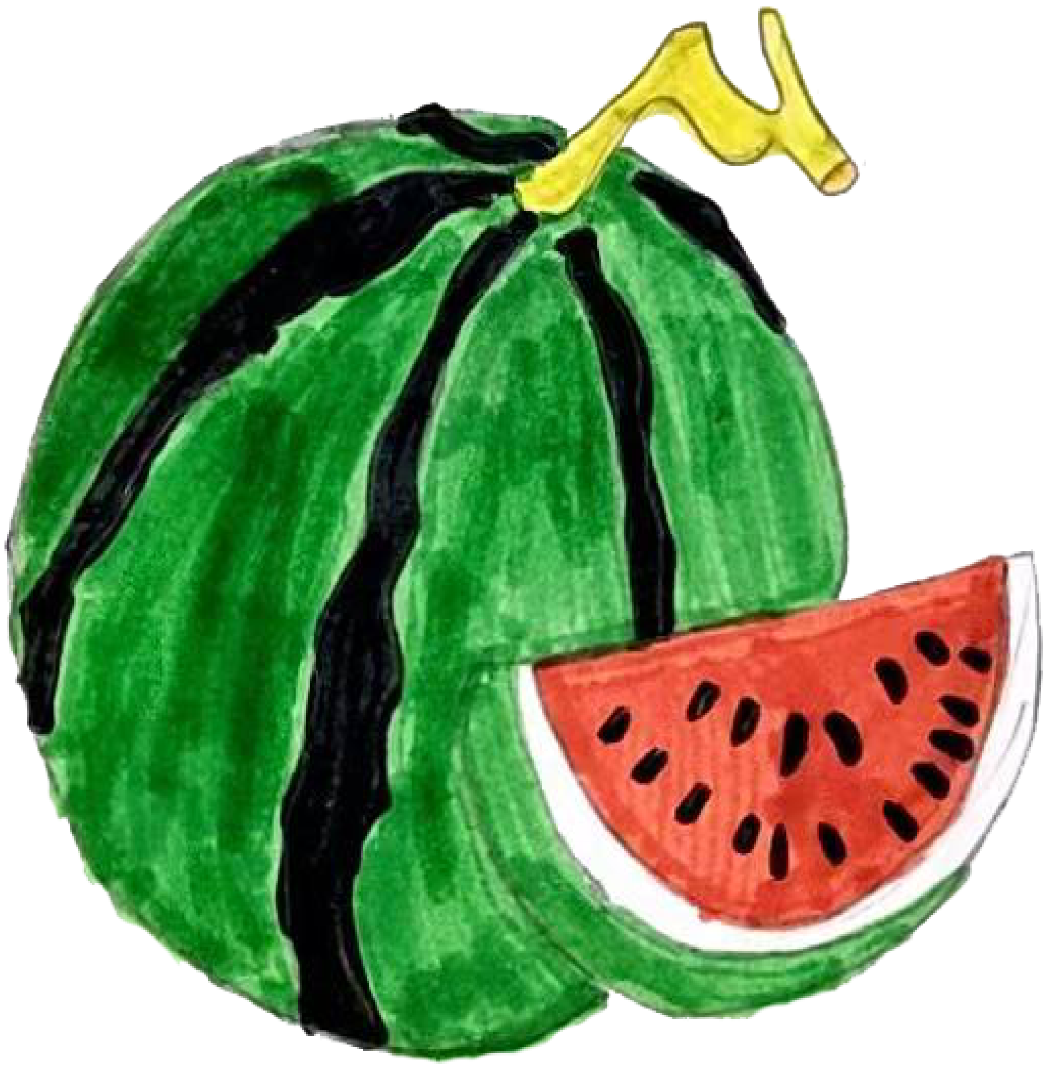
This chapter conducted a comprehensive uncertainty analysis for NUE prediction with a consideration of the spatial and temporal correlation of measurement errors in inputs to NUE calculations. It also analyzed the contribution of input and model errors to NUE predictions of machine learning models. The results showed, as expected, that the NUE calculations uncertainty of the reference scenario was larger than that of the optimistic and smaller than that of the pessimistic scenario. The differences were large, which indicated that proper quantification of input errors is important. For PFP_N calculations, Guangxi and Shanghai had the largest probability distribution width between the quantile 0.05 and 0.95, while Jilin and Inner Mongolia had the smallest. For PNB_N calculations, Heilongjiang and Jilin had the largest distribution width between the quantile 0.05 and 0.95, while Beijing and Hainan had the smallest. Results also revealed that the temporal variation of NUE prediction uncertainty (PIW_{90} and PIR_{90}) had a downward trend due to the improvement of technology and policy. In 2015, the PFP_N had lower uncertainty in northeast China, while PNB_N had higher uncertainty in northeast China. This was likely caused by the difference in major crop types between these regions. NUE had smaller input uncertainty than model uncertainty in most provinces, except for PNB_N, which showed converse results after 2010. This means the QRF model had a better performance for PNB_N in the 2010s. Future work should focus on bookkeeping of detailed field data and accurate collection of crop parameters and explanatory variables.

Supplementary material

The supplementary materials, Appendix E, can be downloaded as Appendix A from: https://wageningenur4-my.sharepoint.com/:w:/g/personal/yingxia_liu_wur_nl/ERBtICvWsJlKl3hgxK7hUGEBS4Cw0pN1BF8Odx90zDy1Dw.

Chapter 6

Synthesis



This thesis addressed four key objectives aiming at enhancing our understanding of the spatial and temporal variation in crop yield and NUE at provincial and county scales in China. To achieve these objectives and understand the role and importance of agricultural, environmental and economic factors as explanatory variables of the variation I made use of stepwise multiple linear regression (SMLR) and random forest (RF) models, as well as uncertainty propagation in nitrogen use efficiency (NUE) prediction using Monte Carlo simulation and quantile regression forests (QRF). Figure 6.1 depicts the roadmap that I used to achieve the defined objectives. Section 6.1 summarizes the main findings of this thesis. In Section 6.2 I address their implications and make recommendations on how the results of this thesis research can be used in policy. The innovations and limitations of this PhD research are presented in Section 6.3.

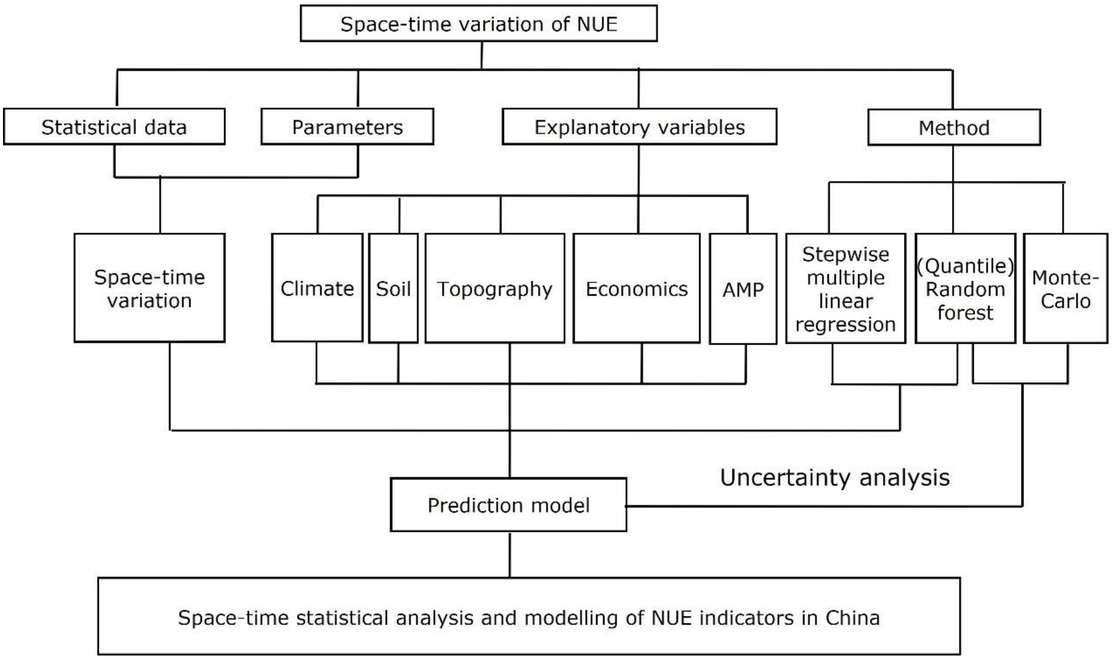


Figure 6.1: From data to information: Roadmap of this PhD research.

6.1. Main findings

Chapter 2 used SMLR to analyze the influence of explanatory variables on NUE and quantify their importance. NUE varies considerably in space and time in China. The partial factor productivity of nitrogen (PFP_N) was larger in north- and south-east China than in other regions. The partial nutrient balance of nitrogen (PNB_N) was lower in south China than in other parts of China. The national PFP_N declined slightly from 32 kg kg^{-1} in 1978 to 27 kg kg^{-1} in 1995 and went up gradually to reach 38 kg kg^{-1} in 2015. The national PNB_N decreased from 0.53 to 0.36 kg kg^{-1}

from 1978 to 2003; thereafter it stabilized at around 0.40 kg kg^{-1} yr^{-1} between 2004 and 2015. Multiple linear regression explained 74% of the NUE variation. Crop types and various soil properties were identified as major influential factors in the PFP_N model; while crop types, climate and soil properties accounted for most of the PNB_N variation. These findings should be considered by policy makers when agricultural land development decisions are made to balance NUE and productivity (i.e., agronomy and environment).

Chapter 3 employed the same method as used in **Chapter 2** but explored the variation in aggregated yield instead of NUE indicators at a provincial scale for the whole of China. It also evaluated the driving factors of aggregate yield variation. The aggregated yield for all crops was higher in provinces in the eastern coast and south China than in the inland west and north China. The SMLR explained 95% of the aggregated yield spatio-temporal variation. Crop types, soil covariates, economic variables and agricultural management practices were the major explanatory variables in modelling the aggregated crop yield. I concluded that enhancing economic growth could be an adequate solution to meet the growing food demand for an increasing population with limited agricultural land resources, in combination with better management practices, crop composition, breeding and planting technologies.

Chapter 4 further explored the importance of the explanatory variables of NUE at a finer spatial resolution i.e., at county scale in northeast China. Analysis at the finer county scale revealed different patterns and variable importances than the coarse provincial scale analysis. NUE decreased in most of the counties during the study period and was highest in Heilongjiang province. The soil, crop and climatic covariates had higher relative importance in the SMLR model of NUE variation. The RF model (model efficiency coefficient: 0.84 and 0.89 for PFP_N and PNB_N, respectively) has a superior performance than the SMLR model (model efficiency coefficient: 0.44 and 0.67 for PFP_N and PNB_N, respectively), which indicated a non-linear relation between explanatory variables and NUE. All modelling was done in the R language for statistical computing. The ranger package was more efficient than the randomForest package for building the RF model.

Chapter 4 showed that RF performed better than SMLR for modelling NUE variation, but the analysis also revealed that neither of the two models were perfect. In **Chapter 5** I therefore quantified the uncertainty of the NUE predictions and determined contributions of uncertainty sources for the 31 provinces from 1978 to 2015. The uncertainty of the NUE prediction caused by measurement uncertainty in yield, N input and N removal was quantified using Monte Carlo simulation in three scenarios, while model uncertainty was assessed using QRF. The prediction uncertainty for both NUE indicators decreased over time. In 2015, PFP_N had a higher 90% prediction interval ratio (PIR_{90}) of input data in south and west China and a higher 90% Prediction Interval Width (PIW_{90}) in south and east-costal China, while PNB_N had a higher PIR_{90} in north China and a higher PIW_{90} in northeast China. The NUE prediction uncertainty propagated from QRF models has similar spatial patterns as input data. NUE in most provinces had smaller input uncertainty than model uncertainty, except PNB_N, which had larger model uncertainty than input uncertainty after 2010. Generally, PNB_N had higher input uncertainty contributions than PFP_N in 2015, especially in south and northeast China. Overall, the uncertainties in NUE predictions were substantial. A series of recommendations were made to improve the accuracy of NUE prediction. These may be applied by the government, in order to inform sustainable nitrogen management in food systems.

The provincial model for NUE (**Chapter 2**) had higher explanatory ability than the county model (**Chapter 4**) since errors partly cancelled out with spatial aggregation. Predictions outside the area of applicability should be handled with care or be left out from further consideration because the environmental properties may differ too strongly from those observed in the training data. Although RF had a superior performance than SMLR, the uncertainty of NUE predictions from the RF model was still larger than that from input uncertainty, especially for PFP_N. I also noticed that the scenarios had largely different results, indicating that proper assessment of input uncertainty deserves attention. Input uncertainties were largely derived from expert judgement, which is less reliable than deriving these from direct measurements.

6.2. Implications and recommendations

Based on the findings of this thesis, implications and recommendations are given as follows:

- First, the findings of **Chapter 5** showed that if we want to obtain more accurate results and conclusions, we should improve the accuracy of input data/parameters. Furthermore, we should use suitable models that meet the research objectives and take the merits and demerits of models into consideration. Model accuracy is not the only important issue because we may also want to interpret it with explicit equations (i.e., this is where SMLR outperforms RF). This thesis also showed that the explanatory variables of crop yield and NUE are scale- and crop-dependent. It is therefore important to analyze the scale effect in more detail and improve our understanding of the causes of the scale-dependency. Since crop type is an important explanatory variable for variations in crop yield and NUE, the government could encourage farmers to grow high NUE crops. Crop breeders have committed to breed high N nutrition index crops and obtained significant achievements e.g., in hybrid maize (Fernandez & Ciampitti, 2019; Lemaire & Ciampitti, 2020).

- Second, crop types, climate and soil properties should be taken into consideration in policymaking for agricultural land development in order to balance NUE and productivity (i.e., agronomy and environment). This thesis showed that farmer’s income has a significant positive correlation with crop yield. However, the gross domestic product in agriculture has a negative impact on NUE in northeast China, indicating that northeast China is still in the first stage of Environmental Kuznets Curve, that is a stage where N pollution first increases and then decreases with economic growth (Dinda, 2004). In this case, I suggest that policy makers could macro-control the N fertilizer production and import and give fertilizer recommendations for proper application to farmers. In this respect it is worth mentioning that the education level of farmers is increasing in China under the implementation of the nine-year compulsory education law.
- Third, the government could promote the current agricultural management mode from small-size farm holders to large-size farming. The latter is conducive not only to the development of mechanization and reduction of costs, but also to improvement of NUE for easy adoption of new technologies which promote NUE. For soils with poor fertility, profit and NUE could be improved by choosing suitable crops, varieties and agronomic practices.

6.3. Innovations and limitations

This thesis achieved various innovative aspects, the most important of which are:

1) It analyzed the relative importance of explanatory variables of NUE at different scales. The results at a provincial level could support the provincial governments to compare their yields with other provinces and formulate more effective policies to increase crop yield, such as high-yield fertilization, balanced fertilization and soil-test based fertilization; while comprehensive county scale analyses may be feasible for large regions and provide insights that are obscured by provincial scale analyses (Lu et al., 2019b);

2) It explored the relative importance of explanatory variables of yield for different crops i.e., single crop yields and three crop aggregated yields (**Chapter 3**). To make results economically tangible, dry weight of staples crops (including beans) and fresh weight of cash crops (except for beans) were considered. Results showed that soil properties and AMP were more important for single crop yields than for aggregated crops, and climatic covariates were important for the staples and cash models, but not for the aggregate yield model for all crops. This integrated information provides general patterns and trends about different crop aggregations for decision and policy makers.

3) I applied (geo)statistical modelling and analysis of NUE in plant nutrition and compared model uncertainty with input uncertainty. This was novel because most of the methods and models reported in the literature only evaluate NUE from an agronomic perspective by optimizing nutrient management strategies and mainly include agronomic explanatory variables in modelling, such as by using crop yield models (Xu et al., 2013), precision agriculture (Diacono et al., 2012), site-specific nutrient management (Dobermann et al., 2003), 4R nutrient stewardship (i.e., right source, right rate, right time and right place) (Johnston & Bruulsema, 2014), Nutrient Expert Systems (Xu et al., 2014b) and soil testing (He et al., 2009);

4) In this thesis I included multiple agricultural, environmental, geographical and economic variables in the data-driven modelling in order to tease out which of those explanatory variables play major roles in NUE. Explanatory variables of NUE, such as socio-economic variables (e.g., income), AMP (e.g., irrigated area, agricultural machinery) and environmental variables (e.g., soil, climate) turned out to be important for explaining the variation of NUE in space and time, developing strategies to balance crop yield, profitability and environmental sustainability, and achieving suitability-based high-efficient agricultural management. Although there are studies on modelling of spatial and temporal variation of NUE, little work has been done to explore its explanatory variables and model NUE through deriving empirical relationships with explanatory variables (Ma et al., 2012; He et al., 2018).

Yet this thesis also had some limitations:

1) It lacked field verification due to time limitation. The models built were evaluated using cross-validation. The model performance for specific fields

and crops might be worse as I concluded that the model is crop- and scale-dependent;

2) I only analyzed two NUE indicators (i.e., the partial factor productivity of nitrogen and the partial nutrient balance of nitrogen) as other indicators need field experimental data for treatment without N application (blank control treatment), which were not available in the statistical data used in this thesis;
3) I assessed the importance of NUE explanatory variables for aggregated crops but not for specific crops. In many cases decision makers may need crop-specific analyses, which was beyond the scope of this thesis. Such analyses are hampered by the fact that crop-specific N application data at county scale are not readily available in the governmental database (i.e., in the county and provincial Yearbook of northeast China);
4) Model uncertainty had a large influence on NUE prediction uncertainty and more advanced and suitable models remain to be explored. While the SMLR is a useful method to quantify the effect of explanatory variables on the dependent variables and allows interpretation because of the simplicity of the resulting models, it is worth noting that the SMLR only reflects linear and additive relationships. The machine learning approach (e.g., random forests) can solve this problem (James et al., 2013) but it is a black box and has low interpretability. Thus, this thesis was limited in the sense that it did not use a model that was both flexible and non-linear, and that at the same time was easy to interpret and transparent.

Improving NUE in a sustainable way without impeding crop productivity is challenging and complex. This thesis was devoted to shed light on this research question. I found that there was considerable space-time variation in crop yield and NUE in China. The relative importance of explanatory variables can be diverse at different scales and for different crops between yield and NUE. Northeast China is a good example that had low NUE with high crop production. I found that the main reason for this was that northeast China has fertile soil and suitable crops (varieties). Economic variables and agricultural management practices were also important for crop production, while the effect of climate is different between crops. I also concluded that uncertainties in measured input data and models have a significant impact on results. Considering the uncertainty contributions of input data and models for NUE prediction, I encourage the government to standardize the data collection process and I stimulate scientists to explore available data better using statistical tools and develop more suitable models.

In spite of the many complexities and challenges, in this PhD research I hope to have presented valuable information, insights and guidelines to support policy makers to take better decisions on developing agricultural land management and agronomic policies in a resource use efficient way, based on food security and environmental sustainability principles.

References

Aghdaei, N., Kokogiannakis, G., Daly, D., & McCarthy, T. (2017). Linear regression models for prediction of annual heating and cooling demand in representative Australian residential dwellings. *Energy Procedia*, *121*, 79-86.

Alencar, F. A. R., Filho, C. M., Silva, D. G. D., & Wolf, D. F. (2014). Pedestrian Classification Using K-means and Random Decision Forests. Robotics: Sbr-lars Robotics Symposium & Robocontrol,

Alexander, R. G., Schmidt, J., & Zelinsky, G. J. (2014). Are summary statistics enough? Evidence for the importance of shape in guiding visual search. *Visual Cognition*, *22*(3-4), 595-609.

Allys, E., Marchand, T., Cardoso, J.-F., Villaescusa-Navarro, F., Ho, S., & Mallat, S. (2020). New interpretable statistics for large-scale structure analysis and generation. *Physical Review D*, *102*(10), 103506.

Barraclough, P. B., Howarth, J. R., Jones, J., Lopez-Bellido, R., Parmar, S., Shepherd, C. E., & Hawkesford, M. J. (2010). Nitrogen efficiency of wheat: genotypic and environmental variation and prospects for improvement. *European Journal of Agronomy*, *33*(1), 1-11.

Bavajigari, S. K. K., & Singh, C. (2019). Investigation of Computational Advantage of using Importance sampling in Monte Carlo Simulation. 2019 North American Power Symposium (NAPS),

Bengtsson, G., Bengtson, P., & Månsson, K. F. (2003). Gross nitrogen mineralization-, immobilization-, and nitrification rates as a function of soil C/N ratio and microbial activity. *Soil Biology and Biochemistry*, *35*(1), 143-154.

Bettinger, P. (2021). Random Forest Regression Model for Estimation of the Growing Stock Volumes in Georgia, USA, Using Dense Landsat Time Series and FIA Dataset. *Remote Sensing*, *13*.

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*, 5-32.

Brentrup, F., & Pallière, C. (2010). Nitrogen use efficiency as an agro-environmental indicator. Proceedings of the OECD Workshop on Agrienvironmental Indicators, March,

Bullen, P. S. (2013). *Handbook of means and their inequalities* (Vol. 560). Springer Science & Business Media.

Burt, J. E., Barber, G. M., & Rigby, D. L. (2009). *Elementary statistics for geographers*. Guilford Press.

Cassman, K. G., Dobermann, A., & Walters, D. T. (2002). Agroecosystems, nitrogen-use efficiency, and nitrogen management. *AMBIO: A Journal of the*

Human Environment, *31*(2), 132-140.

Cen, H., Wan, L., Zhu, J., Li, Y., Li, X., Zhu, Y., Weng, H., Wu, W., Yin, W., Xu, C., Bao, Y., Feng, L., Shou, J., & He, Y. (2019). Dynamic monitoring of biomass of rice under different nitrogen treatments using a lightweight UAV with dual image-frame snapshot cameras. *Plant Methods*, *15*, 32. https://doi.org/10.1186/s13007-019-0418-8

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

Chang, V., & Xu, Q. (2021). Analysis of influencing factors of grain yield based on multiple linear regression. *International Journal of Business and Systems Research*, *15*.

Chen, C., Baethgen, W. E., Wang, E., & Yu, Q. (2011a). Characterizing spatial and temporal variability of crop yield caused by climate and irrigation in the North China Plain. *Theoretical and Applied Climatology*, *106*(3-4), 365-381. https://doi.org/10.1007/s00704-011-0440-x

Chen, C., Lei, C., Deng, A., Qian, C., Hoogmoed, W., & Zhang, W. (2011b). Will higher minimum temperatures increase corn production in Northeast China? An analysis of historical data over 1965–2008. *Agricultural and Forest Meteorology*, *151*(12), 1580-1588.

Chen, X., Cui, Z., Fan, M., Vitousek, P., Zhao, M., Ma, W., Wang, Z., Zhang, W., Yan, X., Yang, J., Deng, X., Gao, Q., Zhang, Q., Guo, S., Ren, J., Li, S., Ye, Y., Wang, Z., Huang, J., Tang, Q., Sun, Y., Peng, X., Zhang, J., He, M., Zhu, Y., Xue, J., Wang, G., Wu, L., An, N., Wu, L., Ma, L., Zhang, W., & Zhang, F. (2014). Producing more grain with lower environmental costs. *Nature*, *514*(7523), 486-489. https://doi.org/10.1038/nature13609

Chen, Y., Zhang, Z., Tao, F., Wang, P., & Wei, X. (2017). Spatio-temporal patterns of winter wheat yield potential and yield gap during the past three decades in North China. *Field Crops Research*, *206*, 11-20.

Chien, S., Prochnow, L., & Cantarella, a. H. (2009). Recent developments of fertilizer production and use to improve nutrient efficiency and minimize environmental impacts. *Advances in agronomy*, *102*, 267-322.

China, N. B. o. S. o. (2021). *Chinese Statistical Yearbook*. China statistics press.

Climatic Research Unit. (1978-2015). University of East Auglia, Norwich, UK. http://www.cru.uea.ac.uk/data

Climatic Research Unit. (1990-2015). University of East Auglia, Norwich, UK. http://www.cru.uea.ac.uk/data

Climatic Research Unit. (2015). University of East Auglia, Norwich, UK.

http://www.cru.uea.ac.uk/data

Cook, T. D., Campbell, D. T., & Shadish, W. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin Boston, MA.

Córdoba, M., Carranza, J. P., Piumetto, M., Monzani, F., & Balzarini, M. (2021). A spatially based quantile regression forest model for mapping rural land values. *Journal of Environmental Management*, *289*, 112509.

Correndo, A. A., Rotundo, J. L., Tremblay, N., Archontoulis, S., & Ciampitti, I. A. (2021). Assessing the uncertainty of maize yield without nitrogen fertilization. *Field Crops Research*, *260*(1), 107985.

de Beaufort, H., Nauta, F., Conti, M., Cellitti, E., Trentin, C., Faggiano, E., van Bogerijen, G., Figueroa, C., Moll, F., & van Herwaarden, J. (2017). Extensibility and distensibility of the thoracic aorta in patients with aneurysm. *European Journal of Vascular and Endovascular Surgery*, *53*(2), 199-205.

de Carvalho Junior, W., Lagacherie, P., da Silva Chagas, C., Calderano Filho, B., & Bhering, S. B. (2014). A regional-scale assessment of digital mapping of soil attributes in a tropical hillslope environment. *Geoderma*, *232*, 479-486.

De Reu, J., Bourgeois, J., Bats, M., Zwertvaegher, A., Gelorini, V., De Smedt, P., Chu, W., Antrop, M., De Maeyer, P., & Finke, P. (2013). Application of the topographic position index to heterogeneous landscapes. *Geomorphology*, *186*, 39-49.

de Vries, W., Leip, A., Reinds, G. J., Kros, J., Lesschen, J. P., & Bouwman, A. (2011). Comparison of land nitrogen budgets for European agriculture by various modeling approaches. *Environmental Pollution*, *159*(11), 3254-3268.

Degiuli, Nastia, Barbalić, Marijan, N., & Goran. (2007). Causes of Sampling Measurement Uncertainties when Determining the Particle Concentration in a Gaseous Environment. *Strojniski Vestnik*.

Del Grosso, S. J., Ogle, S. M., Parton, W. J., & Breidt, F. J. (2010). Estimating uncertainty in N2O emissions from U.S. cropland soils. *Global Biogeochemical Cycles*, *24*(1), n/a-n/a. https://doi.org/10.1029/2009gb003544

Deng, N., Grassini, P., Yang, H., Huang, J., Cassman, K. G., & Peng, S. (2019). Closing yield gaps for rice self-sufficiency in China. *Nat Commun*, *10*(1), 1725. https://doi.org/10.1038/s41467-019-09447-9

Diacono, M., Rubino, P., & Montemurro, F. (2012). Precision nitrogen management of wheat. A review. *Agronomy for Sustainable Development*, *33*(1), 219-241. https://doi.org/10.1007/s13593-012-0111-z

Diebold, C. (1954). Permeability and intake rates of medium textured soils in relation to silt content and degree of compaction. *Soil Science Society of America Journal*, *18*(3), 339-343.

Dimkpa, C. O., Fugice, J., Singh, U., & Lewis, T. D. (2020). Development of fertilizers for enhanced nitrogen use efficiency - Trends and perspectives. *Sci Total Environ*, *731*, 139113. https://doi.org/10.1016/j.scitotenv.2020.139113

Dinda, S. (2004). Environmental Kuznets curve hypothesis: a survey. *Ecological economics*, *49*(4), 431-455.

Ding, W., Xu, X., He, P., Ullah, S., Zhang, J., Cui, Z., & Zhou, W. (2018). Improving yield and nitrogen use efficiency through alternative fertilization options for rice in China: A meta-analysis. *Field Crops Research*, *227*, 11-18. https://doi.org/10.1016/j.fcr.2018.08.001

Ding, W., Xu, X., He, P., Zhang, J., Cui, Z., & Zhou, W. (2020). Estimating regional N application rates for rice in China based on target yield, indigenous N supply, and N loss. *Environ Pollut*, *263*(Pt B), 114408. https://doi.org/10.1016/j.envpol.2020.114408

Dobermann, A. (2007). *Nutrient use efficiency–measurement and management*. IFA.

Dobermann, A., & Cassman, K. G. (2005). Cereal area and nitrogen use efficiency are drivers of future nitrogen fertilizer consumption. *Sci China C Life Sci*, *48 Suppl 2*, 745-758. https://doi.org/10.1007/BF03187115

Dobermann, A., Witt, C., Abdulrachman, S., Gines, H., Nagarajan, R., Son, T., Tan, P., Wang, G., Chien, N., & Thoa, V. (2003). Estimating indigenous nutrient supplies for site‐specific nutrient management in irrigated rice. *Agronomy Journal*, *95*(4), 924-935.

Du, Y., Xu, Y., Zhang, L., & Song, S. (2020). Can China's food production capability meet her peak food demand in the future? Based on food demand and production capability prediction till the year 2050. *International Food and Agribusiness Management Review*, *23*(1), 1-17.

Dutta, D., Das, P. K., Bhunia, U. K., Singh, U., Singh, S., Sharma, J. R., & Dadhwal, V. K. (2015). Retrieval of tea polyphenol at leaf level using spectral transformation and multi-variate statistical approach. *International Journal of Applied Earth Observation and Geoinformation*, *36*, 22-29.

Efroymson, M. A. (1960). Multiple regression analysis. *Mathematical methods for digital computers*, 191-203.

ESA. (2015). *ESA/CCI Viewer* http://maps.elie.ucl.ac.be/CCI/viewer/download.php

EU Nitrogen Expert Panel. (2015). Nitrogen Use Efficiency (NUE) an indicator for the utilization of nitrogen in food systems. *Wageningen University, Alterra, Wageningen, Netherlands*.

Everingham, Y., Sexton, J., Skocaj, D., & Inman-Bamber, G. (2016). Accurate prediction of sugarcane yield using a random forest algorithm. *Agronomy for Sustainable Development*, *36*(2). https://doi.org/10.1007/s13593-016-0364-z

Fageria, N. K., & Baligar, V. (2005). Enhancing nitrogen use efficiency in crop plants.

Advances in agronomy, *88*, 97-185.

Fan, X., Xu, D., Wang, Y., Zhang, X., Cao, S., Mou, S., & Ye, N. (2014). The effect of nutrient concentrations, nutrient ratios and temperature on photosynthesis and nutrient uptake by Ulva prolifera: implications for the explosion in green tides. *Journal of Applied Phycology*, *26*(1), 537-544.

Fang, Q., Zhang, X., Chen, S., Shao, L., & Sun, H. (2017). Selecting traits to increase winter wheat yield under climate change in the North China Plain. *Field Crops Research*, *207*, 30-41.

FAO. (2018). *Food and Agriculture Organization of the United Nations* http://www.fao.org/home/en/

FAO, I., UNICEF, WFP and WHO. (2021). *The State of Food Security and Nutrition in the World 2021. Transforming food systems for food security, improved nutrition and affordable healthy diets for all. Rome, FAO* https://doi.org/10.4060/cb4474en

Feng, X., Liu, G., Chen, J., Chen, M., Liu, J., Ju, W., Sun, R., & Zhou, W. (2007). Net primary productivity of China's terrestrial ecosystems from a process model driven by remote sensing. *Journal of Environmental Management*, *85*(3), 563-573.

Fernandez, J., & Ciampitti, I. (2019). Effect of late nitrogen fertilization on grain yield and grain filling in corn. *Kansas Agricultural Experiment Station Research Report s*, 26.

Fixen, P., Brentrup, F., Bruulsema, T., Garcia, F., Norton, R., & Zingore, S. (2015). Nutrient/fertilizer use efficiency: measurement, current situation and trends. *Managing water and fertilizer for sustainable agricultural intensification*, *270*.

Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2003). *Geographically weighted regression: the analysis of spatially varying relationships*. John Wiley & Sons.

Fraiwan, L., Lweesy, K., Khasawneh, N., Wenz, H., & Dickhaus, H. (2012). Automated sleep stage identification system based on time–frequency analysis of a single EEG channel and random forest classifier. *Computer methods and programs in biomedicine*, *108*(1), 10-19.

Fu, M., Liu, B., Wang, J., Sun, Y., Wang, X., & Chen, F. (2019). Temporal and spatial changes of sunflower production in China from 1985 to 2015. *Journal of Henan Agricultural University*.

Galloway, J. N., Townsend, A. R., Erisman, J. W., Bekunda, M., Cai, Z., Freney, J. R., Martinelli, L. A., Seitzinger, S. P., & Sutton, M. A. (2008). Transformation of the nitrogen cycle: recent trends, questions, and potential solutions. *Science*, *320*(5878), 889-892.

Gao, B., Huang, W., Wang, L., Huang, Y., Ding, S., & Cui, S. (2019). Driving forces of nitrogen flows and nitrogen use efficiency of food systems in seven Chinese cities, 1990 to 2015. *Sci Total Environ*, *676*, 144-154. https://doi.org/10.1016/j.scitotenv.2019.04.136

Garamszegi, L. Z. (2016). A simple statistical guide for the analysis of behaviour when data are constrained due to practical or ethical reasons. *Animal Behaviour*, *120*, 223-234.

Gislason, P. O., Benediktsson, J. A., & Sveinsson, J. R. (2006). Random forests for land cover classification. *Pattern recognition letters*, *27*(4), 294-300.

Grömping, U. (2007). Relative importance for linear regression in R: the package relaimpo. *Journal of Statistical Software*, *17*, 1-27.

Grömping, U. (2009). Variable Importance Assessment in Regression: Linear Regression versus Random Forest. *The American Statistician*, *63*(4), 308-319. https://doi.org/10.1198/tast.2009.08199

Gu, B., Ju, X., Chang, J., Ge, Y., & Vitousek, P. M. (2015). Integrated reactive nitrogen budgets and future trends in China. *Proc Natl Acad Sci U S A*, *112*(28), 8792-8797. https://doi.org/10.1073/pnas.1510211112

Guo, E., Liu, X., Zhang, J., Wang, Y., Wang, C., Wang, R., & Li, D. (2017). Assessing spatiotemporal variation of drought and its impact on maize yield in Northeast China. *Journal of Hydrology*, *553*, 231-247. https://doi.org/10.1016/j.jhydrol.2017.07.060

Gurung, R. B., Breidt, F. J., Dutin, A., & Ogle, S. M. (2009). Predicting Enhanced Vegetation Index (EVI) curves for ecosystem modeling applications. *Remote Sensing of Environment*, *113*(10), 2186-2193.

Hamoud, Y. A., Shaghaleh, H., Sheteiwy, M., Guo, X., Elshaikh, N. A., Khan, N. U., Oumarou, A., & Rahim, S. F. (2019). Impact of alternative wetting and soil drying and soil clay content on the morphological and physiological traits of rice roots and their relationships to yield and nutrient use-efficiency. *Agricultural Water Management*, *223*, 105706.

Harris, I., Jones, P. D., Osborn, T. J., & Lister, D. H. (2014). Updated high - resolution grids of monthly climatic observations–the CRU TS3. 10 Dataset. *International journal of climatology*, *34*(3), 623-642.

He, P., Li, S., Jin, J., Wang, H., Li, C., Wang, Y., & Cui, R. (2009). Performance of an Optimized Nutrient Management System for Double - Cropped Wheat - Maize Rotations in North - Central China. *Agronomy Journal*, *101*(6), 1489-1496.

He, P., Yang, L., Xu, X., Zhao, S., Chen, F., Li, S., Tu, S., Jin, J., & Johnston, A. M. (2015). Temporal and spatial variation of soil available potassium in China (1990–2012). *Field Crops Research*, *173*, 49-56.

He, W., Jiang, R., He, P., Yang, J., Zhou, W., Ma, J., & Liu, Y. (2018). Estimating soil nitrogen balance at regional scale in China's croplands from 1984 to 2014. *Agricultural Systems*, *167*, 125-135. https://doi.org/10.1016/j.agsy.2018.09.002

Helfenstein, A., Mulder, V. L., Heuvelink, G. B., & Okx, J. P. (2022). Tier 4 maps of soil pH at 25 m resolution for the Netherlands. *Geoderma*, *410*, 115659.

Hengl, T., Heuvelink, G. B., Kempen, B., Leenaars, J. G., Walsh, M. G., Shepherd, K. D., Sila, A., MacMillan, R. A., Mendes de Jesus, J., Tamene, L., & Tondoh, J. E. (2015). Mapping Soil Properties of Africa at 250 m Resolution: Random Forests Significantly Improve Current Predictions. *PLoS One*, *10*(6), e0125814. https://doi.org/10.1371/journal.pone.0125814

Hengl, T., Mendes de Jesus, J., Heuvelink, G. B., Ruiperez Gonzalez, M., Kilibarda, M., Blagotic, A., Shangguan, W., Wright, M. N., Geng, X., Bauer-Marschallinger, B., Guevara, M. A., Vargas, R., MacMillan, R. A., Batjes, N. H., Leenaars, J. G., Ribeiro, E., Wheeler, I., Mantel, S., & Kempen, B. (2017). SoilGrids250m: Global gridded soil information based on machine learning. *PLoS One*, *12*(2), e0169748. https://doi.org/10.1371/journal.pone.0169748

Heuvelink, G. B. (1998a). Uncertainty analysis in environmental modelling under a change of spatial scale. *Nutrient Cycling in Agroecosystems*, *50*(1-3), 255-264.

Heuvelink, G. B., Burrough, P. A., & Stein, A. (1989). Propagation of errors in spatial modelling with GIS. *International Journal of Geographical Information System*, *3*(4), 303-322.

Heuvelink, G. B. M. (1998b). *Error propagation in environmental modelling with GIS*. CRC press.

Ho, T. K. (1995). Random decision forests. Document Analysis and Recognition, 1995., Proceedings of the Third International Conference on,

Ho, T. K. (1998). The random subspace method for constructing decision forests. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, *20*(8), 832-844.

Hounkpatin, K. O., Bossa, A. Y., Yira, Y., Igue, M. A., & Sinsin, B. A. (2022). Assessment of the soil fertility status in Benin (West Africa)–Digital soil mapping using machine learning. *Geoderma Regional*, *28*, e00444.

Huang, B., Wu, B., & Barry, M. (2010). Geographically and temporally weighted regression for modeling spatio-temporal variation in house prices. *International Journal of Geographical Information Science*, *24*(3), 383-401. https://doi.org/10.1080/13658810802672469

Huang, J. X., Zeng, C. F., Xu, L., Chen, X. W., Jiang, L. B., & Hong, C. M. (2018). Standardized cultivation technology of cantaloupe in Hainan greenhouse *China Vegetables*(6), 91-94. (In Chinese).

Huang, X., Song, Y., Li, M., Li, J., Huo, Q., Cai, X., Zhu, T., Hu, M., & Zhang, H.

(2012). A high-resolution ammonia emission inventory in China. *Global Biogeochemical Cycles*, *26*(1), n/a-n/a. https://doi.org/10.1029/2011gb004161

Ichami, S. M., Shepherd, K. D., Sila, A. M., Stoorvogel, J. J., & Hoffland, E. (2019). Fertilizer response and nitrogen use efficiency in African smallholder maize farms. *Nutrient Cycling in Agroecosystems*, *113*(1), 1-19.

Iizumi, T., & Ramankutty, N. (2016). Changes in yield variability of major crops for 1981–2010 explained by climate change. *Environmental Research Letters*, *11*(3). https://doi.org/10.1088/1748-9326/11/3/034003

Iqbal, A., He, L., Khan, A., Wei, S., Akhtar, K., Ali, I., Ullah, S., Munsif, F., Zhao, Q., & Jiang, L. (2019). Organic manure coupled with inorganic fertilizer: An approach for the sustainable production of rice by improving soil properties and nitrogen use efficiency. *Agronomy*, *9*(10), 651.

Ishaq, M., Ibrahim, M., Hassan, A., Saeed, M., & Lal, R. (2001). Subsoil compaction effects on crops in Punjab, Pakistan. *Soil and Tillage Research*, *60*(3-4), 153-161.

Ishaq, M., Ibrahim, M., & Lal, R. (2002). Tillage effects on soil properties at different levels of fertilizer application in Punjab, Pakistan. *Soil and Tillage Research*, *68*(2), 93-99.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning* (Vol. 112). Springer.

Jansen, M., Rossing, W., & Daamen, R. A. (1994). *Monte Carlo estimation of uncertainty contributions from several independent multivariate sources*. Monte Carlo Estimation of Uncertainty Contributions from Several Independent Multivariate Sources.

Jiang, P., & Thelen, K. (2004). Effect of soil and topographic properties on crop yield in a North‐Central corn–soybean cropping system. *Agronomy Journal*, *96*(1), 252-258.

Jin, D., Gao, J., Jiang, P., Lv, X., Wang, Y., & Zhang, W. (2017). Nitrogen use efficiency and rice yield of different locations in northeast China. *National Academy Science Letters*, *40*(4), 227-232.

Johnston, A. M., & Bruulsema, T. W. (2014). 4R Nutrient Stewardship for Improved Nutrient Use Efficiency. *Procedia Engineering*, *83*, 365-370. https://doi.org/10.1016/j.proeng.2014.09.029

Kabacoff, R. (2011). R in Action: Data Analysis and Graphics With R. Manning, Shelter Island.

Kahrl, F., Li, Y., Su, Y., Tennigkeit, T., Wilkes, A., & Xu, J. (2010). Greenhouse gas emissions from nitrogen fertilizer use in China. *Environmental science & policy*, *13*(8), 688-694.

Kant, S., Bi, Y.-M., & Rothstein, S. J. (2011). Understanding plant response to nitrogen limitation for the improvement of crop nitrogen use efficiency. *Journal of experimental Botany*, *62*(4), 1499-1509.

Kaur, H., Jalota, S., Kanwar, R., & Bhushan Vashisht, B. (2012). Climate change impacts on yield, evapotranspiration and nitrogen uptake in irrigated maize (Zea mays)-wheat (Triticum aestivum) cropping system: A simulation analysis. *Indian Journal of Agricultural Sciences*, *82*(3), 213.

Kay, A., Davies, H., Bell, V., & Jones, R. (2009). Comparison of uncertainty sources for climate change impacts: flood frequency in England. *Climatic Change*, *92*(1), 41-63.

Khan, U. T., Valeo, C., & He, J. (2013). Non-linear fuzzy-set based uncertainty propagation for improved DO prediction using multiple-linear regression. *Stochastic Environmental Research & Risk Assessment*, *27*(3), 599-616.

Khanal, S., Fulton, J., Klopfenstein, A., Douridas, N., & Shearer, S. (2018). Integration of high resolution remotely sensed data and machine learning techniques for spatial prediction of soil properties and corn yield. *Computers and Electronics in Agriculture*, *153*, 213-225.

Kitchen, N., Drummond, S., Lund, E., Sudduth, K., & Buchleiter, G. (2003). Soil electrical conductivity and topography related to yield for three contrasting soil–crop systems. *Agronomy Journal*, *95*(3), 483-495.

Kliopova, I., Baranauskaitė-Fedorova, I., Malinauskienė, M., & Staniškis, J. K. (2016). Possibilities of increasing resource efficiency in nitrogen fertilizer production. *Clean technologies and environmental policy*, *18*(3), 901-914.

Kolehmainen, E., & Knuutinen, J. (1981). Multiple linear regression. *14*(11), 626-628.

Kros, J., Heuvelink, G. B. M., Reinds, G. J., Lesschen, J. P., Ioannidi, V., & De Vries, W. (2012). Uncertainties in model predictions of nitrogen fluxes from agro-ecosystems in Europe. *Biogeosciences*, *9*(11), 4573-4588. https://doi.org/10.5194/bg-9-4573-2012

Kukal, M. S., & Irmak, S. (2018). Climate-Driven Crop Yield and Yield Variability and Climate Change Impacts on the U.S. Great Plains Agricultural Production. *Sci Rep*, *8*(1), 3450. https://doi.org/10.1038/s41598-018-21848-2

Kutner, M. H., Nachtsheim, C. J., Neter, J., & Wasserman, W. (2004). *Applied linear regression models* (Vol. 4). McGraw-Hill/Irwin New York.

Lalitha, M., Dharumarajan, S., Suputhra, A., Kalaiselvi, B., Hegde, R., Reddy, R., Prasad, C., Harindranath, C., & Dwivedi, B. (2021). Spatial prediction of soil depth using environmental covariates by quantile regression forest model. *Environmental Monitoring and Assessment*, *193*(10), 1-10.

Lassaletta, L., Billen, G., Grizzetti, B., Anglade, J., & Garnier, J. (2014). 50 year trends in nitrogen use efficiency of world cropping systems the relationship between yield and nitrogen input to cropland. *Environmental Research Letters*, *9*(10), 105011.

Latiri-Souki, K., Nortcliff, S., & Lawlor, D. W. (1998). Nitrogen fertilizer can increase dry matter, grain production and radiation and water use efficiencies for durum wheat under semi-arid conditions. *European Journal of Agronomy*, *9*(1), 21-34.

LeBreton, J. M., & Tonidandel, S. (2008). Multivariate relative importance: extending relative weight analysis to multivariate criterion spaces. *Journal of Applied Psychology*, *93*(2), 329.

Lemaire, G., & Ciampitti, I. (2020). Crop mass and N status as prerequisite covariables for unraveling nitrogen use efficiency across genotype-by-environment-by-management scenarios: a review. *Plants*, *9*(10), 1309.

Lendasse, A., Wertz, V., & Verleysen, M. (2003). Model selection with cross-validations and bootstraps - Application to time series prediction with RBFN models. *CiteSeer*.

Li, B., Bi, Z., & Xiong, Z. (2017). Dynamic responses of nitrous oxide emission and nitrogen use efficiency to nitrogen and biochar amendment in an intensified vegetable field in southeastern China. *Gcb Bioenergy*, *9*(2), 400-413.

Li, C., Hoffland, E., Kuyper, T. W., Yu, Y., Zhang, C., Li, H., Zhang, F., & van der Werf, W. (2020a). Syndromes of production in intercropping impact yield gains. *Nat Plants*, *6*(6), 653-660. https://doi.org/10.1038/s41477-020-0680-9

Li, S., He, P., & Jin, J. (2013). Nitrogen use efficiency in grain production and the estimated nitrogen input/output balance in China agriculture. *J Sci Food Agric*, *93*(5), 1191-1197. https://doi.org/10.1002/jsfa.5874

Li, X. Y., Zhou, X. G., Zeng, Y. G., Jiang, S. T., & Bai, J. F. (2016a). Spatial-temporal changes analysis of grain production in central plains economic zone in recent 30 years. *J. Nanyang Normal Univ.*, *15*(12), 53-57. (In Chinese). https://doi.org/10.3969/j.issn.1671-6132.2016.12.012

Li, Y., Cui, S., Zhang, Z., Zhuang, K., Wang, Z., & Zhang, Q. (2020b). Determining effects of water and nitrogen input on maize (Zea mays) yield, water- and nitrogen-use efficiency: A global synthesis. *Sci Rep*, *10*(1), 9699. https://doi.org/10.1038/s41598-020-66613-6

Li, Y., Liu, H., & Huang, G. (2016b). The effect of nitrogen rates on yields and nitrogen use efficiencies during four years of wheat–maize rotation cropping seasons. *Agronomy Journal*, *108*(5), 2076-2088.

Li, Y., Ma, J., Xiao, C., & Li, Y. (2020c). Effects of climate factors and soil properties on soil nutrients and elemental stoichiometry across the Huang–Huai–Hai River Basin, China. *Journal of Soils and Sediments*, *20*(4), 1970-1982.

Liang, H., Qi, Z., Hu, K., Prasher, S. O., & Zhang, Y. (2016). Can nitrate contaminated groundwater be remediated by optimizing flood irrigation rate with high nitrate water in a desert oasis using the WHCNS model? *J Environ Manage*, *181*, 16-25. https://doi.org/10.1016/j.jenvman.2016.05.082

Liang, L., Zhao, X., Yi, X., Chen, Z., Dong, X., Chen, R., & Shen, R. (2013). Excessive application of nitrogen and phosphorus fertilizers induces soil acidification and phosphorus enrichment during vegetable production in Y angtze R iver D elta, C hina. *Soil Use and Management*, *29*(2), 161-168.

Liang, S., Li, Y., Zhang, X., Sun, Z., Sun, N., Duan, Y., Xu, M., & Wu, L. (2018). Response of crop yield and nitrogen use efficiency for wheat-maize cropping system to future climate change in northern China. *Agricultural and Forest Meteorology*, *262*, 310-321. https://doi.org/10.1016/j.agrformet.2018.07.019

Lie, M., Glaser, B., & Huwe, B. (2012). Uncertainty in the spatial prediction of soil texture: Comparison of regression tree and Random Forest models. *Geoderma*, *170*(3), 70–79.

Lin, L. (1989). A concordance correlation coefficient to evaluate reproducibility. *Biometrics*, *45*, 255-268.

Lindeman, R. H., Merenda, P., & Gold, R. (1980). Introduction to Bivariate and Multivariate Analysis; Scott. *Foresman Co.: Glenview, UK*.

Liu, M., Liu, X., Liu, D., Ding, C., & Jiang, J. (2015a). Multivariable integration method for estimating sea surface salinity in coastal waters from in situ data and remotely sensed data using random forest algorithm. *Computers & Geosciences*, *75*, 44-56.

Liu, S., Yang, J. Y., Zhang, X. Y., Drury, C. F., Reynolds, W. D., & Hoogenboom, G. (2013a). Modelling crop yield, soil water content and soil temperature for a soybean–maize rotation under conventional and conservation tillage systems in Northeast China. *Agricultural Water Management*, *123*, 32-44. https://doi.org/10.1016/j.agwat.2013.03.001

Liu, X., Chen, L., Hua, Z., Mei, S., Wang, P., & Wang, S. (2020). Comparing ammonia volatilization between conventional and slow-release nitrogen fertilizers in paddy fields in the Taihu Lake region. *Environmental Science and Pollution Research*, *27*(8), 8386-8394.

Liu, X., He, P., Jin, J., Zhou, W., Sulewski, G., & Phillips, S. (2011). Yield gaps, indigenous nutrient supply, and nutrient use efficiency of wheat in China. *Agronomy Journal*, *103*(5), 1452-1463.

Liu, Y., Gao, M., Wu, W., Tanveer, S. K., Wen, X., & Liao, Y. (2013b). The effects of conservation tillage practices on the soil water-holding capacity of a non-irrigated apple orchard in the Loess Plateau, China. *Soil and Tillage Research*, *130*, 7-12.

Liu, Y., Zhou, Z., Zhang, X., Xu, X., Chen, H., & Xiong, Z. (2015b). Net global warming potential and greenhouse gas intensity from the double rice system with integrated soil–crop system management: A three-year field study. *Atmospheric Environment*, *116*, 92-101.

Liu, Z., Gao, J., Gao, F., Dong, S., Liu, P., Zhao, B., & Zhang, J. (2018). Integrated agronomic practices management improve yield and nitrogen balance in double cropping of winter wheat-summer maize. *Field Crops Research*, *221*, 196-206.

Lobell, D. B. (2007a). Changes in diurnal temperature range and national cereal yields. *Agricultural and Forest Meteorology*, *145*(3-4), 229-238.

Lobell, D. B. (2007b). The cost of uncertainty for nitrogen fertilizer management: A sensitivity analysis. *Field Crops Research*, *100*(2-3), 210-217.

Lobell, D. B., & Field, C. B. (2007). Global scale climate–crop yield relationships and the impacts of recent warming. *Environmental Research Letters*, *2*(1). https://doi.org/10.1088/1748-9326/2/1/014002

Lu, C., Zhang, J., Cao, P., & Hatfield, J. L. (2019a). Are We Getting Better in Using Nitrogen?: Variations in Nitrogen Use Efficiency of Two Cereal Crops Across the United States. *Earth's Future*, *7*(8), 939-952. https://doi.org/10.1029/2019ef001155

Lu, Y.-y., Liu, F., Zhao, Y.-g., Song, X.-d., & Zhang, G.-l. (2019b). An integrated method of selecting environmental covariates for predictive soil depth mapping. *Journal of Integrative Agriculture*, *18*(2), 301-315. https://doi.org/10.1016/s2095-3119(18)61936-7

Lucas, R. E. (1982). Oraganic soils (Histosols) Formation, distribution, Physical and chemical properties and management for crop production. *Mich. Agr. Ext. Bull.*, *435*, 72.

Luo, Z., Lam, S. K., Fu, H., Hu, S., & Chen, D. (2019). Temporal and spatial evolution of nitrous oxide emissions in China: Assessment, strategy and recommendation. *Journal of Cleaner Production*, *223*, 360-367. https://doi.org/10.1016/j.jclepro.2019.03.134

Lv, Z., Liu, X., Cao, W., & Zhu, Y. (2017). A model-based estimate of regional wheat yield gaps and water use efficiency in main winter wheat production regions of China. *Scientific reports*, *7*(1), 1-15.

Ma, L., Velthof, G., Wang, F., Qin, W., Zhang, W., Liu, Z., Zhang, Y., Wei, J., Lesschen, J., & Ma, W. (2012). Nitrogen and phosphorus use efficiencies and losses in the food chain in China at regional scales in 1980 and 2005. *Science of The Total Environment*, *434*, 51-61.

Masuda, K. (2016). Measuring eco-efficiency of wheat production in Japan: a combined application of life cycle assessment and data envelopment analysis. *Journal of Cleaner Production*, *126*, 373-381.

https://doi.org/10.1016/j.jclepro.2016.03.090

Meena, S. K., Rakshit, A., & Meena, V. S. (2016). Effect of seed bio-priming and N doses under varied soil type on nitrogen use efficiency (NUE) of wheat (Triticum aestivum L.) under greenhouse conditions. *Biocatalysis and agricultural biotechnology*, *6*, 68-75.

Meinshausen, N. (2006). Quantile Regression Forests. *Journal of Machine Learning Research*, *7*, 983–999.

Meng, Q., Sun, Q., Chen, X., Cui, Z., Yue, S., Zhang, F., & Römheld, V. (2012). Alternative cropping systems for sustainable water and nitrogen use in the North China Plain. *Agriculture, Ecosystems & Environment*, *146*(1), 93-102.

Meyer, H., & Pebesma, E. (2021). Predicting into unknown space? Estimating the area of applicability of spatial prediction models. *Methods in Ecology and Evolution*, *12*(9), 1620-1633.

Miao, Y., Mulla, D. J., & Robert, P. C. (2006). Identifying important factors influencing corn yield and grain quality variability using artificial neural networks. *Precision Agriculture*, *7*(2), 117-135. https://doi.org/10.1007/s11119-006-9004-y

Miao, Y., Stewart, B. A., & Zhang, F. (2010). Long-term experiments for sustainable nutrient management in China. A review. *Agronomy for Sustainable Development*, *31*(2), 397-414. https://doi.org/10.1051/agro/2010034

Miller, C. M. F., Waterhouse, H., Harter, T., Fadel, J. G., & Meyer, D. (2020). Quantifying the uncertainty in nitrogen application and groundwater nitrate leaching in manure based cropping systems. *Agricultural Systems*, *184*. https://doi.org/10.1016/j.agsy.2020.102877

Mogollón, J., Lassaletta, L., Beusen, A., Van Grinsven, H., Westhoek, H., & Bouwman, A. (2018). Assessing future reactive nitrogen inputs into global croplands based on the shared socioeconomic pathways. *Environmental Research Letters*, *13*(4), 044008.

Montgomery, D. C., Peck, E. A., & Vining, G. G. (2021). *Introduction to linear regression analysis*. John Wiley & Sons.

Mueller, L., Schindler, U., Shepherd, T. G., Ball, B. C., Smolentseva, E., Hu, C., Hennings, V., Schad, P., Rogasik, J., Zeitz, J., Schlindwein, S. L., Behrendt, A., Helming, K., & Eulenstein, F. (2012). A framework for assessing agricultural soil quality on a global scale. *Archives of Agronomy and Soil Science*, *58*(sup1), S76-S82. https://doi.org/10.1080/03650340.2012.692877

Mukhopadhyay, S., Masto, R. E., Tripathi, R. C., & Srivastava, N. K. (2019). Application of Soil Quality Indicators for the Phytorestoration of Mine Spoil Dumps. In *Phytomanagement of Polluted Sites* (pp. 361-388). https://doi.org/10.1016/b978-0-12-813912-7.00014-4

National Bureau of Statistics of China. (1978-2015). *China Statistics Press, Beijing* http://www.stats.gov.cn/

National Bureau of Statistics of China. (1991-2016). *China Statistics Press, Beijing* http://www.stats.gov.cn/

National Bureau of Statistics of China. (2019). *China Statistics Press, Beijing* (Agriculture) http://www.stats.gov.cn/tjsj/zbjs/201912/t20191202_1713049.html

National Bureau of Statistics of China. (2022). *China Statistics Press, Beijing* http://www.stats.gov.cn/

Nicodemus, K. K., Malley, J. D., Strobl, C., & Ziegler, A. (2010). The behaviour of random forest permutation-based variable importance measures under predictor correlation. *BMC Bioinformatics*, *11*(1), 1-13.

Nigon, T. J., Yang, C., Paiao, G. D., Mulla, D. J., & Fernández, F. (2020). Prediction of Early Season Nitrogen Uptake in Maize Using High-Resolution Aerial Hyperspectral Imagery. *Remote Sensing*, *12*(8), 1234.

Nol, L., Heuvelink, G. B. M., Veldkamp, A., de Vries, W., & Kros, J. (2010). Uncertainty propagation analysis of an N2O emission model at the plot and landscape scale. *Geoderma*, *159*(1-2), 9-23. https://doi.org/10.1016/j.geoderma.2010.06.009

Omara, P., Aula, L., Oyebiyi, F., & Raun, W. R. (2019). World Cereal Nitrogen Use Efficiency Trends: Review and Current Knowledge. *Agrosystems, Geosciences & Environment*, *2*(1), 1-8. https://doi.org/10.2134/age2018.10.0045

Orlenko, A., & Moore, J. (2020). Improving the Interpretability of Random Forest Models of Genetic Association in the Presence of Non-Additive Interactions.

Pan, J., Liu, Y., Zhong, X., Lampayan, R. M., Singleton, G. R., Huang, N., Liang, K., Peng, B., & Tian, K. (2017). Grain yield, water productivity and nitrogen use efficiency of rice under different water management and fertilizer-N inputs in South China. *Agricultural Water Management*, *184*, 191-200.

Pan, X., Baquy, M., Guan, P., Yan, J., Wang, R., Xu, R., & Xie, L. (2020). Effect of soil acidification on the growth and nitrogen use efficiency of maize in Ultisols. *Journal of Soils and Sediments*, *20*(3), 1435-1445.

Peerlinck, A., Sheppard, J., & Maxwell, B. (2018). Using deep learning in yield and protein prediction of winter wheat based on fertilization prescriptions in precision agriculture. International Conference on Precision Agriculture (ICPA),

Peng, S., Buresh, R. J., Huang, J., Yang, J., Zou, Y., Zhong, X., Wang, G., & Zhang, F. (2006). Strategies for overcoming low agronomic nitrogen use efficiency in irrigated rice systems in China. *Field Crops Research*, *96*(1), 37-47. https://doi.org/10.1016/j.fcr.2005.05.004

Peng, S., Huang, J., Sheehy, J. E., Laza, R. C., Visperas, R. M., Zhong, X., Centeno, G. S., Khush, G. S., & Cassman, K. G. (2004). Rice yields decline with higher night temperature from global warming. *Proceedings of the National Academy of Sciences*, *101*(27), 9971-9975.

Peng, X., Maharjan, B., Yu, C., Su, A., Jin, V., & Ferguson, R. B. (2015). A laboratory evaluation of ammonia volatilization and nitrate leaching following nitrogen fertilizer application on a coarse‐textured soil. *Agronomy Journal*, *107*(3), 871-879.

Pereira, S., Meier, R., Mckinley, R., Wiest, R., & Reyes, M. (2017). Enhancing interpretability of automatically extracted machine learning features: application to a RBM-Random Forest system on brain lesion segmentation. *Medical Image Analysis*, *44*, 228.

Peters, J., Verhoest, N., Samson, R., Boeckx, P., & De Baets, B. (2008). Wetland vegetation distribution modelling for the identification of constraining environmental variables. *Landscape ecology*, *23*(9), 1049-1065.

Piazza, E. A., Sweeny, T. D., Wessel, D., Silver, M. A., & Whitney, D. (2013). Humans Use Summary Statistics to Perceive Auditory Sequences. *Psychological Science*, *24*(8), 1389-1397.

Pilbeam, D. J. (2015). Breeding crops for improved mineral nutrition under climate change conditions. *Journal of experimental Botany*, *66*(12), 3511-3521.

Poggio, L. M., de Sousa, L., Batjes, N. H., Heuvelink, G. B. M., Kempen, B., Riberio, E., & Rossiter, D. (2020). SoilGrids 2.0 producing quality-assessed soil information for the globe. *Soil Discuss*, *1*. https://doi.org/10.5194/soil-2020-65

Popescu, P. S., Mihaescu, M. C., Popescu, E., & Mocanu, M. (2016). Using Ranking and Multiple Linear Regression to Explore the Impact of Social Media Engagement on Student Performance. IEEE International Conference on Advanced Learning Technologies,

Porwollik, V., Müller, C., Elliott, J., Chryssanthacopoulos, J., Iizumi, T., Ray, D. K., Ruane, A. C., Arneth, A., Balkovič, J., Ciais, P., Deryng, D., Folberth, C., Izaurralde, R. C., Jones, C. D., Khabarov, N., Lawrence, P. J., Liu, W., Pugh, T. A. M., Reddy, A., Sakurai, G., Schmid, E., Wang, X., de Wit, A., & Wu, X. (2017). Spatial and temporal uncertainty of crop yield aggregations. *European Journal of Agronomy*, *88*, 10-21. https://doi.org/10.1016/j.eja.2016.08.006

Prasad, A. M., Iverson, L. R., & Liaw, A. (2006). Newer classification and regression tree techniques: bagging and random forests for ecological prediction. *Ecosystems*, *9*(2), 181-199.

Qian, C., Yu, Y., Gong, X., Jiang, Y., Zhao, Y., Yang, Z., Hao, Y., Li, L., Song, Z., & Zhang, W. (2016). Response of grain yield to plant density and nitrogen rate in spring maize hybrids released from 1970 to 2010 in Northeast China. *The Crop Journal*, *4*(6), 459-467.

Qiu, S., He, P., Zhao, S., Li, W., Xie, J., Hou, Y., Grant, C., Zhou, W., & Jin, J. (2015).

Impact of nitrogen rate on maize yield and nitrogen use efficiencies in northeast China. *Agronomy Journal*, *107*(1), 305-313.

Quan, Z., Zhang, X., Fang, Y., & Davidson, E. A. (2021). Different quantification approaches for nitrogen use efficiency lead to divergent estimates with varying advantages. *Nature Food*, *2*(4), 241-245. https://doi.org/10.1038/s43016-021-00263-3

R Core Team. (2018). *R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria* http://www.r-project.org/

R Core Team. (2021). *R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria* http://www.r-project.org/

Ramírez, E., & Reheul, D. (2009). Statistical modelling of nitrogen use efficiency of dairy farms in Flanders. *Agronomy for Sustainable Development*, *29*(2), 339-352. https://doi.org/10.1051/agro/2008065

Ray, D. K., Gerber, J. S., MacDonald, G. K., & West, P. C. (2015). Climate variation explains a third of global crop yield variability. *Nat Commun*, *6*, 5989. https://doi.org/10.1038/ncomms6989

Reidsma, P., Feng, S., van Loon, M., Luo, X., Kang, C., Lubbers, M., Kanellopoulos, A., Wolf, J., van Ittersum, M. K., & Qu, F. (2012). Integrated assessment of agricultural land use policies on nutrient pollution and sustainable development in Taihu Basin, China. *Environmental science & policy*, *18*, 66-76.

Ren, C., Jin, S., Wu, Y., Zhang, B., Kanter, D., Wu, B., Xi, X., Zhang, X., Chen, D., & Xu, J. (2021). Fertilizer overuse in Chinese smallholders due to lack of fixed inputs. *Journal of Environmental Management*, *293*, 112913.

Ren, C., Wang, Z., Song, K., Zhang, B., Liu, D., Yang, G., & Liu, Z. (2011). Spatial variation of soil organic carbon and its relationship with environmental factors in the farming-pastoral ecotone of Northeast China. *Fresenius Environmental Bulletin*, *20*(1A), 253-261.

Ren, K. Y., Duan, Y. H., Minggang, X. U., & Zhang, X. B. (2019). Effect of Manure Application on Nitrogen Use Efficiency of Crops in China: A Meta-Analysis. *Scientia Agricultura Sinica*.

Rencher, A., & Christensen, W. (2012). Multivariate regression. *Methods of multivariate analysis*, 339-384.

Richardson, H. J., Hill, D. J., Denesiuk, D. R., & Fraser, L. H. (2017). A comparison of geographic datasets and field measurements to model soil carbon using random forests and stepwise regressions (British Columbia, Canada). *GIScience & Remote Sensing*, *54*(4), 573-591.

Roberts, T. (2007). Right product, right rate, right time and right place… the foundation of best management practices for fertilizer. *Fertilizer best management practices*, *29*, 1-8.

Rodriguez-Galiano, V. F., Ghimire, B., Rogan, J., Chica-Olmo, M., & Rigol-Sanchez, J. P. (2012). An assessment of the effectiveness of a random forest classifier for land-cover classification. *Isprs Journal of Photogrammetry & Remote Sensing*, *67*(Jan.), 93-104.

Ryan, T. P. (2008). *Modern regression methods* (Vol. 655). John Wiley & Sons.

Sakamoto, T. (2020). Incorporating environmental variables into a MODIS-based crop yield estimation method for United States corn and soybeans through the use of a random forest regression algorithm. *ISPRS Journal of Photogrammetry and Remote Sensing*, *160*, 208-228. https://doi.org/10.1016/j.isprsjprs.2019.12.012

Sänger, A., Geisseler, D., & Ludwig, B. (2011). Effects of moisture and temperature on greenhouse gas emissions and C and N leaching losses in soil treated with biogas slurry. *Biology and Fertility of Soils*, *47*(3), 249-259.

Shangguan, W., Dai, Y., Duan, Q., Liu, B., & Yuan, H. (2014). A global soil data set for earth system modeling. *Journal of Advances in Modeling Earth Systems*, *6*(1), 249-263.

Sharma, A., Weindorf, D. C., Wang, D., & Chakraborty, S. (2015). Characterizing soils via portable X-ray fluorescence spectrometer: 4. Cation exchange capacity (CEC). *Geoderma*, *239*, 130-134.

Sharma, L. K., & Bali, S. K. (2018). A review of methods to improve nitrogen use efficiency in agriculture. *Sustainability*, *10*(1), 51.

Simmonds, M. B., Plant, R. E., Peña-Barragán, J. M., van Kessel, C., Hill, J., & Linquist, B. A. (2013). Underlying causes of yield spatial variability and potential for precision management in rice systems. *Precision Agriculture*, *14*(5), 512-540. https://doi.org/10.1007/s11119-013-9313-x

Singh, G., Williard, K., Schoonover, J., Nelson, K. A., & Kaur, G. (2019). Cover crops and landscape position effects on nitrogen dynamics in plant-soil-water pools. *Water*, *11*(3), 513.

Snyder, C., & Bruulsema, T. (2007). Nutrient use efficiency and effectiveness in North America. *Publ. Int. Plant Nutr. Inst. IPNI*.

Soergel, B., Kriegler, E., Weindl, I., Rauner, S., Dirnaichner, A., Ruhe, C., Hofmann, M., Bauer, N., Bertram, C., Bodirsky, B. L., Leimbach, M., Leininger, J., Levesque, A., Luderer, G., Pehl, M., Wingens, C., Baumstark, L., Beier, F., Dietrich, J. P., Humpenöder, F., von Jeetze, P., Klein, D., Koch, J., Pietzcker, R., Strefler, J., Lotze-Campen, H., & Popp, A. (2021). A sustainable development pathway for climate action within the UN 2030 Agenda. *Nature Climate Change*, *11*(8), 656-664. https://doi.org/10.1038/s41558-021-01098-3

SoilGrids. (2018). *International Soil Reference and Information Centre-World Soil Information.* https://soilgrids.org/

Sousa, S., Martins, F. G., Alvim-Ferraz, M., & Pereira, M. C. (2007). Multiple linear regression and artificial neural networks based on principal components to predict ozone concentrations. *Environmental Modelling & Software*, *22*(1), 97-103.

Stewart., W. M., Dibb, D. W., Johnston, A. E., & Smyth, T. J. (2005). The Contribution of Commercial Fertilizer Nutrients to Food Production. *Agronomy Journal*, *97*(1), 1-6.

Stone, M. (1977). An asymptotic equivalence of choice of model by cross-validation and Akaike's criterion. *Journal of the Royal Statistical Society*, *39*(1).

Strobl, C., Malley, J., & Tutz, G. (2009). An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychol Methods*, *14*(4), 323-348. https://doi.org/10.1037/a0016973

Sun, Y., Wang, X., Wang, N., Chen, Y., & Zhang, S. (2014). Changes in the yield and associated photosynthetic traits of dry-land winter wheat (Triticum aestivum L.) from the 1940s to the 2010s in Shaanxi Province of China. *Field Crops Research*, *167*, 1-10.

Svetnik, V. (2003). Random forest: a classification and regression tool for compound classification and QSAR modeling. *Journal of Chemical Information & Computer Sciences*, *43*.

Tang, Y. H. (2016). Discussion on evaluation factors of cultivated land production potential in heilongjiang province. *Modern. Agric.*, *8*, 63-64. (In Chinese).

Tang, Z., Ma, J., Peng, H., Wang, S., & Wei, J. (2017). Spatiotemporal changes of vegetation and their responses to temperature and precipitation in upper Shiyang river basin. *Advances in Space Research*, *60*(5), 969-979.

Tao, F., Yokozawa, M., Liu, J., & Zhang, Z. (2008). Climate–crop yield relationships at provincial scales in China and the impacts of recent climate trends. *Climate Research*, *38*, 83-94. https://doi.org/10.3354/cr00771

Thongkam, J., Xu, G., & Zhang, Y. (2008). AdaBoost algorithm with random forests for predicting breast cancer survivability. IEEE International Joint Conference on Neural Networks,

Tian, D., & Niu, S. (2015). A global analysis of soil acidification caused by nitrogen addition. *Environmental Research Letters*, *10*(2), 024019.

Tian, X., Engel, B. A., Qian, H., Hua, E., & Wang, Y. (2021). Will reaching the maximum achievable yield potential meet future global food demand? *Journal of Cleaner Production*, *294*, 126285.

Ullah, H., Santiago-Arenas, R., Ferdous, Z., Attia, A., & Datta, A. (2019). Improving water use efficiency, nitrogen use efficiency, and radiation use efficiency in field crops under drought stress: A review. In (pp. 109-157).

https://doi.org/10.1016/bs.agron.2019.02.002

Vaysse, K., & Lagacherie, P. (2017). Using quantile regression forest to estimate uncertainty of digital soil mapping products. *Geoderma*, *291*, 55-64. https://doi.org/10.1016/j.geoderma.2016.12.017

Venables, W. N., & Ripley, B. D. (2002). *Modern Applied Statistics with S* (Fourth ed.). Springer.

Wang, D., Chakraborty, S., Weindorf, D. C., Li, B., Sharma, A., Paul, S., & Ali, M. N. (2015). Synthesized use of VisNIR DRS and PXRF for soil characterization: Total carbon and total nitrogen. *Geoderma*, *243*, 157-167.

Wang, E., Martre, P., Zhao, Z., Ewert, F., Maiorano, A., Rotter, R. P., Kimball, B. A., Ottman, M. J., Wall, G. W., White, J. W., Reynolds, M. P., Alderman, P. D., Aggarwal, P. K., Anothai, J., Basso, B., Biernath, C., Cammarano, D., Challinor, A. J., De Sanctis, G., Doltra, J., Dumont, B., Fereres, E., Garcia-Vila, M., Gayler, S., Hoogenboom, G., Hunt, L. A., Izaurralde, R. C., Jabloun, M., Jones, C. D., Kersebaum, K. C., Koehler, A. K., Liu, L., Muller, C., Naresh Kumar, S., Nendel, C., O'Leary, G., Olesen, J. E., Palosuo, T., Priesack, E., Eyshi Rezaei, E., Ripoche, D., Ruane, A. C., Semenov, M. A., Shcherbak, I., Stockle, C., Stratonovitch, P., Streck, T., Supit, I., Tao, F., Thorburn, P., Waha, K., Wallach, D., Wang, Z., Wolf, J., Zhu, Y., & Asseng, S. (2017). The uncertainty of crop yield projections is reduced by improved temperature response functions. *Nat Plants*, *3*, 17102. https://doi.org/10.1038/nplants.2017.102

Wang, F., Liu, Y., Kong, X., Chen, Y., & Pan, J. (2018a). Spatial and temporal variation of grain production and its influencing factors at the county level in China. *Economic Geography*, *38*(5), 142-151.

Wang, H., & Zhou, J. (2014). Calculation of real fertilizer use efficiency and discussion on fertilization strategies. *Acta Pedologica Sinica*, *51*(2), 216-225 (*In Chinese*).

Wang, M., Ma, L., Strokal, M., Chu, Y., & Kroeze, C. (2018b). Exploring nutrient management options to increase nitrogen and phosphorus use efficiencies in food production of China. *Agricultural Systems*, *163*, 58-72. https://doi.org/10.1016/j.agsy.2017.01.001

Wang, R., Wan, S., Kang, Y., Liu, S., & Zhang, Y. (2016a). Effects of water-fertilizer-salt regulation under dripfertigation on rapeseed yield and the water and fertilizer efficiency in Qinghai Province. *Water Saving Irri*, *4*, 18-23.

Wang, S., Jin, X., Adhikari, K., Li, W., Yu, M., Bian, Z., & Wang, Q. (2018c). Mapping total soil nitrogen from a site in northeastern China. *Catena*, *166*, 134-146.

Wang, S., Luo, S., Gao, Q., & Li, S. (2019). INFLUENCE OF N SPLIT APPLICATION ON NH3 VOLATILIZATION LOSSES AND N RECOVERY EFFICIENCY FROM

PLASTIC MULCHING MAIZE IN LOESS PLATEAU, CHINA. *Applied Ecology and Environmental Research*, *17*(4), 9215-9227.

Wang, X.-H., Jiang, Y.-L., Liu, Y., Lu, J., Yin, X.-G., Shi, L.-G., Huang, J., Chu, Q.-Q., & Chen, F. (2018d). Spatio-temporal Changes of Rice Production in China Based on County Unit. *Acta Agronomica Sinica*, *44*(11). https://doi.org/10.3724/sp.J.1006.2018.01704

Wang, Z., Ye, T., Wang, J., Cheng, Z., & Shi, P. (2016b). Contribution of climatic and technological factors to crop yield: empirical evidence from late paddy rice in Hunan Province, China. *Stochastic Environmental Research and Risk Assessment*, *30*(7), 2019-2030.

Webster, R., & Oliver, M. A. (2007). *Geostatistics for environmental scientists*. John Wiley & Sons.

Wright, M. N., & Ziegler, A. (2017). ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. *Journal of Statistical Software*, *77*(1). https://doi.org/10.18637/jss.v077.i01

Wu, L., Chen, X., Cui, Z., Wang, G., & Zhang, W. (2015). Improving nitrogen management via a regional management plan for Chinese rice production. *Environmental Research Letters*, *10*(9), 095011.

Xiao, H. J., Jiang, T. M., Xia, J. H., & Deng, Y. (2004). Analysis and calculation of main crop potential productivity in Guizhou. *Southwest China Journal of Agricultural Sciences*, *17*(5), 580-583. (In Chinese, with English abstract).

Xu, H., Rigazio, L., & Kryze, D. (2006). Vector Taylor series based joint uncertainty decoding. International Conference on Interspeech -icslp,

Xu, R., Tian, H., Pan, S., Prior, S. A., Feng, Y., Batchelor, W. D., Chen, J., & Yang, J. (2019). Global ammonia emissions from synthetic nitrogen fertilizer applications in agricultural systems: Empirical and process-based estimates and uncertainty. *Glob Chang Biol*, *25*(1), 314-326. https://doi.org/10.1111/gcb.14499

Xu, X., He, P., Pampolino, M. F., Chuan, L., Johnston, A. M., Qiu, S., Zhao, S., & Zhou, W. (2013). Nutrient requirements for maize in China based on QUEFTS analysis. *Field Crops Research*, *150*, 115-125. https://doi.org/10.1016/j.fcr.2013.06.006

Xu, X., He, P., Pampolino, M. F., Johnston, A. M., Qiu, S., Zhao, S., Chuan, L., & Zhou, W. (2014a). Fertilizer recommendation for maize in China based on yield response and agronomic efficiency. *Field Crops Research*, *157*, 27-34. https://doi.org/10.1016/j.fcr.2013.12.013

Xu, X., He, P., Pampolino, M. F., Li, Y., Liu, S., Xie, J., Hou, Y., & Zhou, W. (2016a). Narrowing yield gaps and increasing nutrient use efficiencies using the Nutrient

Expert system for maize in Northeast China. *Field Crops Research*, *194*, 75-82. https://doi.org/10.1016/j.fcr.2016.05.005

Xu, X., He, P., Qiu, S., Pampolino, M. F., Zhao, S., Johnston, A. M., & Zhou, W. (2014b). Estimating a new approach of fertilizer recommendation across small-holder farms in China. *Field Crops Research*, *163*, 10-17. https://doi.org/10.1016/j.fcr.2014.04.014

Xu, X., He, P., Zhang, J., Pampolino, M. F., Johnston, A. M., & Zhou, W. (2017). Spatial variation of attainable yield and fertilizer requirements for maize at the regional scale in China. *Field Crops Research*, *203*, 8-15. https://doi.org/10.1016/j.fcr.2016.11.013

Xu, X., He, P., Zhao, S., Qiu, S., Johnston, A. M., & Zhou, W. (2016b). Quantification of yield gap and nutrient use efficiency of irrigated rice in China. *Field Crops Research*, *186*, 58-65. https://doi.org/10.1016/j.fcr.2015.11.011

Yan, X., Xia, L., & Ti, C. (2022). Temporal and spatial variations in nitrogen use efficiency of crop production in China. *Environmental Pollution*, *293*, 118496.

Yang, F., Xu, X., Ma, J., He, P., Pampolino, M. F., & Zhou, W. (2017). Experimental validation of a new approach for rice fertiliser recommendations across smallholder farms in China. *Soil Research*, *55*(6). https://doi.org/10.1071/sr16328

Yang, R.-M., Zhang, G.-L., Liu, F., Lu, Y.-Y., Yang, F., Yang, F., Yang, M., Zhao, Y.-G., & Li, D.-C. (2016). Comparison of boosted regression tree and random forest models for mapping topsoil organic carbon concentration in an alpine ecosystem. *Ecological Indicators*, *60*, 870-878.

Yang, S. (2017). The influence of planting methods on rice yield in ningxia and discussion on increasing yield measures. *North Rice*, *47*(4), 56-60. (In Chinese). https://doi.org/10.3969/j.issn.1673-6737.2017.04.022

Yang, Z. M., Feng, X. L., Huang, X., Lai, X. L., & Wang, Z. W. (2019). Grain Crops in Eastern Agricultural Region of Qinghai During 1987-2017: Spatio-temporal Variation Characteristics of Potential Productivity and Yield Gaps. *Chinese Agricultural Science Bullentin*, *35*(3), 26-33. (In Chinese with English abstract).

Yong, L., & Liao, S. (2014). Preventing Over-Fitting of Cross-Validation with Kernel Stability. Th European Conference on Machine Learning & Knowledge Discovery in Databases,

Yousaf, M., Li, X., Zhang, Z., Ren, T., Cong, R., Ata-Ul-Karim, S. T., Fahad, S., Shah, A. N., & Lu, J. (2016). Nitrogen Fertilizer Management for Enhancing Crop Productivity and Nitrogen Use Efficiency in a Rice-Oilseed Rape Rotation System in China. *Front Plant Sci*, *7*, 1496. https://doi.org/10.3389/fpls.2016.01496

Zhang, M., Gao, Y., Zhang, Z., Liu, Y., Han, M., Hu, N., Wang, Z., Sun, Z., & Zhang, Y. (2020). Limited irrigation influence on rotation yield, water use, and wheat traits. *Agronomy Journal*, *112*(1), 241-256.

Zhang, W.-y. (2012). Effects of fertilization on spring rape in eastern part of Qinghai province. *Journal of Anhui Agricultural Sciences*.

Zhang, W., Cao, G., Li, X., Zhang, H., Wang, C., Liu, Q., Chen, X., Cui, Z., Shen, J., Jiang, R., Mi, G., Miao, Y., Zhang, F., & Dou, Z. (2016). Closing yield gaps in China by empowering smallholder farmers. *Nature*, *537*(7622), 671-674. https://doi.org/10.1038/nature19368

Zhang, X., Bol, R., Rahn, C., Xiao, G., Meng, F., & Wu, W. (2017). Agricultural sustainable intensification improved nitrogen use efficiency and maintained high crop yield during 1980–2014 in Northern China. *Science of The Total Environment*, *596*, 61-68.

Zhang, X., Davidson, E. A., Mauzerall, D. L., Searchinger, T. D., Dumas, P., & Shen, Y. (2015). Managing nitrogen for sustainable development. *Nature*, *528*(7580), 51-59. https://doi.org/10.1038/nature15743

Zhang, X., Ren, C., Gu, B., & Chen, D. (2021a). Uncertainty of nitrogen budget in China. *Environ Pollut*, *286*, 117216. https://doi.org/10.1016/j.envpol.2021.117216

Zhang, X., Zou, T., Lassaletta, L., Mueller, N. D., Tubiello, F. N., Lisk, M. D., Lu, C., Conant, R. T., Dorich, C. D., Gerber, J., Tian, H., Bruulsema, T., Maaz, T. M., Nishina, K., Bodirsky, B. L., Popp, A., Bouwman, L., Beusen, A., Chang, J., Havlík, P., Leclère, D., Canadell, J. G., Jackson, R. B., Heffer, P., Wanner, N., Zhang, W., & Davidson, E. A. (2021b). Quantification of global and national nitrogen budgets for crop production. *Nature Food*, *2*(7), 529-540. https://doi.org/10.1038/s43016-021-00318-5

Zhang, Z. (2016). Variable selection with stepwise and best subset approaches. *Annals of translational medicine*, *4*(7).

Zhao, Y. W. (2010). A discussion on yield-increase potential of fine-quality varieties in major grain crops of Guangxi. *J. Guangxi Agric.*, *25*(5), 45-48. (In Chinese). https://doi.org/10.3969/j.issn.1003-4374.2010.05.017

Zheng, J., Yin, S., Kang, D., Che, W., & Zhong, L. (2012). Development and uncertainty analysis of a high-resolution NH_3 emissions inventory and its implications with precipitation over the Pearl River Delta region, China. *Atmospheric Chemistry and Physics*, *12*(15), 7041-7058.

Zhu, J., Li, R., & Yang, X. (2012). Spatial and temporal distribution of crop straw resources in 30 years in China. *J. Northwest A & F University-Nat. Sci. Ed.*, *40*(4), 139-145. (In Chinese, with English abstract).

Zhu, Q., de Vries, W., Liu, X., Hao, T., Zeng, M., Shen, J., & Zhang, F. (2018). Enhanced acidification in Chinese croplands as derived from element budgets in the period 1980-2010. *Sci Total Environ*, *618*, 1497-1505. https://doi.org/10.1016/j.scitotenv.2017.09.289

Zscheischler, J., Orth, R., & Seneviratne, S. I. (2017). Bivariate return periods of temperature and precipitation explain a large fraction of European crop yields. *Biogeosciences*, *14*(13), 3309-3320.

Zuur, A., Ieno, E., Walker, N., Saveliev, A., & Smith, G. (2009). Mixed effects models and extensions in ecology with R; Gail M, Krickeberg K, Samet JM, Tsiatis A, Wong W, eds. *New York: Spring Science and Business Media*.

Summary

Crop yields in China increased substantially over the past decades, mainly driven by the increasing use of chemical fertilizers, improved crop varieties and agronomic management. Nitrogen (N), as a major constituent of chemical fertilizer, is applied to agricultural fields to improve the growth and yield of crop. However, excess N application not only decreases the economic efficiency of fertilizer application, but can also result in serious environmental problems, such as waterbody eutrophication, greenhouse gas emission and soil acidification. In other words, there is an urgent need to improve nitrogen use efficiency (NUE), since this would allow increasing yield and profits with minimal environmental impact.

Recent research showed that there is a large temporal and spatial variation of crop yield and NUE in China. But existing research did not perform a thorough analysis and interpretation of this phenomenon. Explanatory variables of NUE, such as socio-economic variables (e.g., income), agricultural management practice (e.g., irrigated area, agricultural machinery) and environmental variables (e.g., soil, climate) are crucial for explaining the variation of NUE in space and time, developing strategies to balance crop yield, profitability and environmental sustainability, and achieving suitability-based efficient agricultural management. Most existing research only concentrated on the influence of N application rate, crop variety and soil type on NUE by performing experiments for specific sites, which does not yield representative relationships between NUE and explanatory variables for the entire country. More advanced statistical methods are required to explore the influence factors of NUE. Stepwise multiple linear regression (SMLR) models the linear relation between a dependent variable (i.e., NUE indicator) and explanatory variables, by an iterative process that continues to add or remove variables from the regression equation until there is no improvement. Random forest (RF) is an ensemble learning method for classification, regression and other tasks that operates by constructing a multitude of decision trees. Both SMLR and RF are practical, meaningful, and informative methods for exploring the effect of explanatory variables on NUE and quantifying their relative importance in explaining NUE variability from a large data set. Policy makers are typically focused on overall patterns, and hence they are more interested in general findings for aggregated crops. However, there is no established scientific and unified method to aggregate yield and NUE among different crops regionally.

The main objective of this thesis was to apply (geo)statistical methods to analyze and explain space-time patterns of crop yield and NUE in China at two spatial scales, to support the development of effective strategies and policies. These policies should improve NUE in a sustainable way without impacting crop productivity. The objective was approached based on: 1) an overview and analysis of space-time variation of NUE and corresponding model predictions with agricultural, environmental and economic explanatory variables(**Chapter 2**); 2) an overview and analysis of space-time variation of different crop yields and their relation with explanatory variables (**Chapter 3**); 3) a finer resolution exploration and statistical modelling of space-time NUE in northeast China (**Chapter 4**); and 4) uncertainty quantification of NUE predictions using Monte Carlo simulation and quantile regression forests in China (**Chapter 5**). Chapters 2, 3 and 5 focused on analysis at provincial scale, while **Chapter 4** was carried out at county scale.

In **Chapter 2**, I collected yield, livestock, and fertilizer data and corresponding parameters at provincial scale. Spatial and temporal variation of NUE was analyzed and revealed in maps and graphs. In addition, I developed and calibrated multiple linear regression models that predict NUE indicators at provincial scale in China from explanatory variables (crop type, climate, topography, soil type and properties, economic variables and agricultural management practices (AMP)). The results showed substantial temporal and spatial variation of the partial factor productivity of N (PFP_N) and partial nutrient balance of N (PNB_N). PFP_N was larger in east and south China than in central and west China. It was also smaller than 30 kg kg^{-1} yr^{-1} in most provinces. PNB_N was low in south China (< 0.40 kg kg^{-1} yr^{-1}), moderate in most provinces (0.41-0.50 kg kg^{-1} yr^{-1}), and high in northeast and southwest China. The PFP_N in China decreased from 32 kg kg^{-1} in 1978 to 27 kg kg^{-1} in 1995, after which it increased to 38 kg kg^{-1} in 2015. PNB_N varied from 0.53 in 1978 to 0.38 kg kg^{-1} in 2000, after which it remained constant until 2015. SMLR proved to be an effective and powerful modelling approach to model and predict NUE and derive the major influencing factors of the dependent variables. The models derived in **Chapter 2** explained more than 70% of the variation of NUE. Crop types and various soil properties were influential factors of the PFP_N model, while crop types, climate and soil properties accounted for most of the variation of PNB_N. Although the models could explain a large part of the spatial and temporal variation, they may be improved by expanding the covariate set with additional relevant variables and by exploring the use of non-linear statistical models (as was done in **Chapter 4**). Suitable crop types, temperature and soil properties should be considered by policy makers when taking decisions on developing agricultural land management in an agricultural resource use efficient way, by balancing NUE, productivity, and the environment.

In **Chapter 3** I analyzed temporal and spatial variation of yield for multiple crop aggregations. Stepwise multiple linear regression was used to explore the relationships between crop yield and agricultural, environmental and economic explanatory variables. The temporal and spatial patterns of yields were different for different levels of crop aggregations. Most of the models explained more than 60% of the crop yield variance, except for rice, potato and cotton. AMP, soil and economic covariates were the most important factors in all models. Topography had an influence on the aggregate yield (provincial yield including all crops, calculated as provincial production divided by provincial cultivated land) but was not included in the staples and cash model. Instead, climatic covariates were important for the staples and cash models, but not for the aggregate yield model. Model performance for the aggregate yield was different for each province in individual years and residuals of the regression model had distinct spatial and temporal patterns. Hence, a more detailed analysis of model performance and residual analysis is needed to explore the causes of these patterns. The models could not predict the impact of natural hazards, plant diseases and insect pests due to lack of data. This may be improved in future research using a combination of natural disaster prediction and pest diagnosis analysis. With the increasing food requirement and limited agricultural land resources, enhancing economic growth might be a possible solution for China to safeguard food security, if this is combined with better management practices, breeding and planting technologies, and taking account of crop suitability (i.e., adaptability of crops to the local environment).

Since a provincial scale analysis may be too coarse for some policy decisions, I also analyzed spatial and temporal variation of NUE at county scale in the high NUE region in northeast China, expecting remaining potentials of NUE improvement (**Chapter 4**). Results demonstrated that the NUE indicators decreased in most counties during the study period and were higher in Heilongjiang than in the other two provinces of northeast China. SMLR and RF models were both applied in this chapter, to explore the explanatory variables of NUE from a more comprehensive and complex perspective. The RF model had a superior performance than the SMLR model, indicating that many covariates had a non-linear relation with NUE. Both models smoothed the reality and underpredicted high extremes and overpredicted low extremes. The relative importance of crop covariates was much higher in SMLR than in RF, while soil and climatic covariates were more important in RF, confirming a difference between linear and non-linear models of the relation between dependent and explanatory variables. These novel findings are particularly valuable when put into action in supporting land-use management and policymaking.

As we know, no model is perfect. In order to quantify the uncertainty of NUE prediction uncertainty in RF modelling, in **Chapter 5**, I conducted a comprehensive uncertainty analysis using Monte Carlo simulation and quantile regression forests (QRF), with a consideration of the spatial and temporal correlation of measurement errors. I used three scenarios (pessimistic, reference, and optimistic) to evaluate the sensitivity of the results for the magnitude of the measurement errors in yield, N input and N removal. The results showed, as expected, that NUE calculations uncertainty of the reference scenario was larger than that of the optimistic and smaller than that of the pessimistic scenario. The differences between scenarios were large, which indicates that proper quantification of input errors is important. For PFP_N calculations, Guangxi and Shanghai had the largest probability distribution width between the 0.05 and 0.95 quantiles, while Jilin and Inner Mongolia had the smallest. For PNB_N calculations, Heilongjiang and Jilin had the largest distribution width, while Beijing and Hainan had the smallest. Results also revealed that the temporal variation of NUE prediction uncertainty (90% Prediction Interval Width, PIW_{90} and Prediction Interval Ratio, PIR_{90}) had a downward trend due to the improvement of technology and policy. In 2015, the PFP_N had lower uncertainty in northeast China, while PNB_N had higher uncertainty in northeast China. This was likely caused by the difference in major crop types between these regions. NUE had smaller input uncertainty than model uncertainty in most provinces, except for PNB_N, which showed converse results after 2010. This means that the QRF model had a better performance for PNB_N in the 2010s. Future work should focus on bookkeeping of detailed field data and accurate collection of crop parameters and explanatory variables.

The thesis synthesis is given in **Chapter 6**. It discusses the main findings of this thesis, my personal implications and recommendations for government and policy makers, and points out the innovations and limitations of this thesis research. In conclusion, the relative importance of explanatory variables can be diverse at different scales and for different crops, and can be different between yield and NUE. Policy makers should make considerate decisions on agronomic policies based on food security and environmental sustainability, and for this they require adequate information and insights, which this thesis aimed to provide. Soil, crop and climatic covariates had high relative importance for NUE, while economic variables and agricultural management practices were also important for crop production. Considering the uncertainty contributions of input data and models for NUE prediction, we encourage the government to standardize the data collection process

and inspire scientists to explore available data better using statistical tools and to develop more suitable models.

Acknowledgements

I am sorry that COVID-19 has forced me to work at home, and prevented me from traveling to Europe. I owe my perseverance in finishing the PhD to the encouragement of my supervisors, family and friends. At the very beginning, to be honest, my progress was very slow because of my difficulties with spoken English. I also failed the IELTS exam many times. My co-promotor-Zhanguo, gave me great courage by analyzing my dilemma and actively seeking solutions from his rich storehouse of experience. Following his advice, I joined IELTS exam once more in the Netherlands. As expected, I passed. This gave me temporary relief. My Chinese supervisor Ping also encouraged me many times and helped me grow in confidence. I am very grateful for her trust and glad that I have such a considerate supervisor. My Dutch supervisor, Gerard invited me to the Netherlands before I pass the IELTS exam. After a few months, I realized that Gerard took a risk and had to fill out plenty of files to guarantee that my English was good enough. I am deeply touched and feel greatly blessed him. Not only that, but he also went to the train station to pick up me with his car. He also lent me his daughter’s bike to ensure that I could commute to work. By the way, I finally realized that Dutch people are indeed very tall, because that bike was too high for me. Speaking of which, I should thank Claudius for adjusting the bike seat and helping to improve my presentation skills during the courses. He gave me a lot of assistance to be a successful PhD candidate.

My trouble was not only from the language. Inter-professional knowledge asked me to learn more about the geographic information system (GIS). I remember that my first course was advanced GIS. The teacher removed me from this course because I did not join in on the group discussions. That was the first week I came to the Netherlands, after having had a fever during the flight. Gerard felt disappointed in me because of this course. He lent me a book and other materials and suggested that I learn by myself. I felt pain in my heart and even a little bitterness. Thankfully, my Chinese roommate Yan always invited me to her room, made dinner for me, and encouraged me that this was understandable and predictable in the Dutch education system, which is very different from the Chinese one. I quickly picked myself up from my disappointment and finished most of the courses in Training and Supervision Plan. Gerard felt satisfied and confident about my progress and give me a “GO” after I passed the IELTS exam in November 2018. Looking back, 2018 was a tough year for me, but I passed the English exam, submitted the TSP and proposal, finished most of the courses, learned R programming, and got a “GO”. What a fruitful year!

Acknowledgements

My three years in Wageningen were a pleasant experience with nice, kind colleagues in Soil Geography and Landscape group—Luc, Maricke, Simona, Cindy, Cynthia, Abbey, Anatol, Tijn, Selcuk, Alexandre, Marijn, Jasper, Arturo, and Eric. They brought immeasurable joy to my stressful PhD life. In addition, I want to especially thank Dainius, who was my mentoring teacher for R programming and my co-author for my papers. I also have lots of Chinese friends in the Netherlands, such as Yan, Panpan, Xiaoxiao, Qian, Lina, Yusi, Lei, Dengke, Xiulu, Ran, Xue, Jingjing, Qingbo, Jingyan, Dazhi, Siyuan, Chen, Xiaoqian, Yanzhou, and Meijun. I also want to thank Claudius and Mingjun for providing solid foundations for PhD candidates.

During the middle period of my PhD—in 2019, I worked in China. Everything was familiar and the study environment and friendships were excellent. I sped up my progress and finished three drafts based on the knowledge learned in Wageningen University & Research. I want to thank the accompaniment of Dali, Jiajia, Jinchuan, Xinpeng, Qinlei, Wencheng, Shaohui, Xiaoyong, Tengfei, Qian, RongRong, Lingling as well as the guidance of Shicheng and Shaojun. If I encounter problems in the future, I would hope to have discussions with them. They always give me new ideas and inspiration.

I want to thank my supervisors who worked overtime in revisions of my draft to meet the defense date. During these four years, Gerard was very patient with me and gave me careful detailed help in revising my draft. He is a great scientist and nice friend. He is humorous and always talks to me with a smile on his face. He provided me with solid knowledge about statistical modeling and uncertainty analysis. Zhanguo, as my co-promoter, was always available to answer my questions and help with difficulties in life. He could always explain complicated matters in a way that was easy to understand. Ping, as my promotor, provided constructive knowledge about nutrient management. She is an elegant lady. Talking with her is like admiring a beautiful painting. Her academic attitude is rigorous and in-depth. She allowed me to conduct my research freely but asked me to explain the results comprehensively and penetratingly. I want to thank her for financial support and academic help.

Moreover, I should thank my corridor mates—Bjorn, Viditha, Valerie, Ester, Stefan and Paulien. We had dinner, played cards and sang together when celebrating Christmas. I still remember that sweet time. In spirituality and faith, I would like to thank Almighty God. He always guides me and helps me in difficulties. I also thank my brothers and sisters in the Church fellowship—Ed, Margreet, Phil, EJ, Maaike, Christina and other members.

At last, I would like to thank my family. I feel regret for missing them during four New Year celebrations. My niece gave me great help in designing my thesis and chapter covers, which I hugely appreciate.

致谢

时光飞逝，转眼间已到毕业季，回忆博士求学生涯，一切历历在目。因为疫情原因我的博士学习部分在家完成，这使我的博士研究更加艰难。在求学中，我的导师、家人和朋友们一直鼓励着我，是他们的支持与鼓励让我拥有足够的毅力坚持读完博士。最初，因为口语的关系，在雅思考试中多次失败，所以过得很艰难。我的导师——白占国老师，他根据我的困境和自己丰富的经验帮我寻求有效的解决方案，这给予了我莫大的勇气。在他的建议下，我再次参加了荷兰的雅思考试。很庆幸，我通过了，让我开启了博士求学的篇章。我的中国导师何萍研究员在我求学中也多次鼓励我，帮助我重拾信心。在我雅思考试中同样给予了莫大的帮助和指导。我非常感谢她的信任，也很高兴有这样一位体贴的导师。几个月后，我意识到我的另一位导师 Gerard 也承担了很大的压力，因为他必须签署大量文件并保证我的英语可以熟练进行日常交流。不仅如此，Gerard 老师在我去荷兰时还开车去火车站接我，并把她女儿的自行车借给我，以确保我能够顺利通勤。另外，我发现荷兰人确实很高，因为那辆自行车对我来说太高了。我要感谢 Claudius 老师为我调整了自行车座椅，并帮助我提高了演讲技巧。他对我成为一名优秀的博士生给予了很多启示。这些优秀导师对我的默默帮助让我深受感动，感到自己更加幸运。

我的困难不仅来自语言。因为研究课题的原因，使我需要更多地了解地理信息系统（GIS）。我记得出国后参加学习的第一门课程是高级 GIS，因为我之前没有 GIS 的基础，不能很好的参加小组讨论，老师让我退出了这门课。那是我去荷兰的第一周，我的信心和动力受到了很大的冲击，也感到无助。Gerard 老师因为课程而对我感到失望。他借给我一本书和一些资料，建议我先自学。我的心里有些苦涩和痛苦。而且我在途中不幸生病了。我非常感谢我的好朋友王艳对我的照顾和鼓励，我们一起在外求学，她总是邀请我到她那吃晚饭。她让我很快重拾了破碎的心，并在后期顺利完成了培训和督导计划（TSP）中的大部分课程。2018 年 11 月雅思考试通过后，Gerard 老师对我的进步感到非常满意并给予了信心，给了我一个“GO” 。回顾 2018 年，对我来说是最艰难的一年，但这一年让我收获和成长了许多，我通过了英语考试，提交了 TSP 和博士研究计划，完成了大部分课程并学习了 R 编程，顺利通过了博士实习。

在瓦赫宁根的三年求学中，我同样与土壤地理景观课题组的同事们相处很融洽——Luc、Maricke、Simona、Cindy、Cynthia、Abbey、Anatol、Tijn、Selcuk、Alexandre、Marijn、Jasper、Arturo、Eric。他们给我紧张的博士生活带来了无数的快乐。此外，我要特别感谢 Dainius，他是我 R 编程的“启蒙老师“，也是我文章的合作者。我在荷兰也有很多中国朋友，比如艳、盼盼、笑笑、倩、丽娜、宇思、磊、登科、秀路、然、雪、静静、清博、景艳、大志、思远、陈、晓倩、烟舟，玫君。我很感谢中国农业科学院研究生院与荷兰瓦赫宁根大学合作举办博士学位教育项目，在项目中 Claudius 老师和张明军老师的辛苦付出为中国博士生顺利求学奠定了坚实的基础。

在博士中期的 2019 年，我回到了中国。一切都很熟悉，学习环境和氛围都很好。在国内我加快了学习进度，根据我在瓦赫宁根大学和研究中心学到的知识完成了三篇文章的草稿。在此，感谢大利、佳佳、进川、新朋、勤雷、文成、少辉、晓永、腾飞、倩、蓉蓉、玲玲等同学和师兄、师姐的陪伴，以及赵士诚老师和仇少君老师的指导。当我遇到问题时，他们总能给我新的想法和灵感。

此外，我要感谢我的导师加班修改我的论文以满足答辩要求。在这四年里，Gerard 老师对我非常有耐心，并且非常仔细地修改我的文章草稿。他是一位伟大的科学家，也是我的好朋友。他很幽默，说话总是面带微笑。他为我提供了统计建模和不确定性分析方面的帮助。白占国老师作为我的共同导师，经常解答我的疑惑，在学习和生活中给予了我许多帮助。他总能用通俗易懂的方式向我解释复杂的事情。作为我的中国农科院导师，何萍老师提供了许多有关养分管理专业方面的知识。她是一位优雅的女士，和她说话就像欣赏一幅美丽的画卷。她的学术态度严谨而深入，她允许我自由地进行研究，但要求我全面而深入地解释结果。我要感谢她的经济支持和学术帮助。

此外，我要感谢我的廊友们——Bjorn、Viditha、Valerie、Ester、Stefan 和 Paulien。我们共进晚餐，一起打牌唱歌庆祝圣诞节。我此刻依然清晰地记得那个温馨的场景。

最后，我要感谢我的家人。很遗憾，四个新年没有与他们一同庆祝。此外，我的外甥女在论文和章节封面绘画中提供了巨大的帮助，非常感谢！最后，我要对以上给予过帮助的老师、家人和同学们致以衷心的感谢和深深的祝福。

List of publications

Liu, Y., Heuvelink, G.B.M., Bai, Z., He, P., Jiang, R., Huang, S., Xu, X. (2022, accepted) Statistical analysis of nitrogen use efficiency in Northeast China using multiple linear regression and random forest. *Journal of Integrative Agriculture*.

Xu, X., He, P., Chuan, L., Liu, X., **Liu, Y.**, Zhang, J., Huang, X., Qiu, S. & Wei, Z. (2021). Regional distribution of wheat yield and chemical fertilizer requirements in China. *Journal of Integrative Agriculture*, *20*(10), 2772-2780.

Liu, Y., Heuvelink, G.B.M., Bai, Z., He, P., Xu, X., Ding, W., & Huang, S. (2021). Analysis of spatio-temporal variation of crop yield in China using stepwise multiple linear regression. *Field Crops Research*, *264*, 108098.

Huang, S., Yang, W., Ding, W., Jia, L., Jiang, L., **Liu, Y.**, Xu, X., Yang, Y., He, P., & Yang, J. (2021). Estimation of Nitrogen Supply for Summer Maize Production through a Long-Term Field Trial in China. *Agronomy*, *11*(7), 1358.

Xu, R., Xu, X. P., Hou, Y. P., Zhang, J. J., Huang, S. H., Ding, W. C., **Liu, Y.** & He, P. (2020). Increasing yield and nitrogen use efficiency of spring maize in Northeast China through ecological intensification management. *Journal of Plant Nutrition and Fertilizers*, *26*(3), 461-471 (*in Chinese with English abstract*).

Liu, Y., Heuvelink, G.B.M., Bai, Z., He, P., Xu, X., Ma, J., & Masiliūnas, D. (2020). Space-time statistical analysis and modelling of nitrogen use efficiency indicators at provincial scale in China. *European Journal of Agronomy*, *115*, 126032.

Huang, S., Ding, W., Yang, J., Zhang, J., Ullah, S., Xu, X., **Liu, Y.**, Yang, Y., Liu, M., He, P. & Jia, L. (2020). Estimation of nitrogen supply for winter wheat production through a long-term field trial in China. *Journal of Environmental Management*, *270*, 110929.

Ding, W., He, P., Zhang, J., **Liu, Y.**, Xu, X., Ullah, S., Cui, Z. & Zhou, W. (2020). Optimizing rates and sources of nutrient input to mitigate nitrogen, phosphorus, and carbon losses from rice paddies. *Journal of Cleaner Production*, *256*, 120603.

Ma, J., **Liu, Y.**, He, W., He, P., Haygarth, P.M., Surridge, B.W.J., Lei., Q. & Zhou, W. (2018). The long - term soil phosphorus balance across Chinese arable land. *Soil Use and Management*, *34*(3), 306-315.

He, W., Jiang, R., He, P., Yang, J., Zhou, W., Ma, J., & **Liu, Y.** (2018). Estimating soil nitrogen balance at regional scale in China's croplands from 1984 to 2014. *Agricultural Systems*, *167*, 125-135.

Liu, Y., Yang, J., He, W., Ma, J., Gao, Q., Lei, Q., He, P., Wu, H., Ullah, S. & Yang, F. (2017). Provincial potassium balance of farmland in China between 1980 and 2010. *Nutrient Cycling in Agroecosystems*, *107*(2), 247-264.

Liu, Y., Ma, J., Ding, W., He, W., Lei, Q., Gao, Q., & He, P. (2017). Temporal and spatial variation of potassium balance in agricultural land at national and regional levels in China. *PloS one*, *12*(9), e0184156.

Ma, J., He, P., Xu, X., He, W., **Liu, Y.**, Yang, F., Chen, F., Li, S., Tu, S., Jin, J., Johnston, A.M. & Zhou, W. (2016). Temporal and spatial changes in soil available phosphorus in China (1990–2012). *Field Crops Research*, *192*, 13-20.

About the author

Yingxia Liu was born on 18 March 1991 in Shandong province, south of Beijing-capital city of China. She lived in a small village and understood agriculture since a little girl. After high school, she chose to learn horticulture in Liaocheng University. In 2014, she finished her Bachelor thesis entitled "Effects of different nitrogen fertilizer management on growth and nutrient uptake of cucumber and Chinese cabbage". The same year she started to study plant nutrition in Jilin Agricultural University as a master student. After four months, she was honored to do her research in the Chinese Academy of Agricultural Sciences (CAAS) under the supervision of Professor Ping He. She did her master thesis on "Temporal and spatial variation in potassium nutrient surplus and deficit balance in farmland of China". During her master study, she developed a keen interest on nutrient management and statistical models. In 2017, she participated in the joint program between CAAS and Wageningen University and Research (WUR) as a sandwich PhD student. She stayed in the Netherlands for 8 months and finished most of the courses of the Training and Supervision Plan and had a pleasant experience with nice and kind colleagues at Soil Geography and Landscape group of WUR. It provided great encouragement and motion to finish her thesis. After one year data collection in China, she got a 2-year scholarship from the Chinese Scholarship Council and went to the Netherlands for a second time, from 2019 November to 2021 November. Partly due to COVID-19 her work was delayed by about a year. By the end of her PhD, Yingxia will return to China to continue working at CAAS.

Her contact email address:

Work email: yingxia.liu@wur.nl;

Private email: liuyingxia91@163.com;

Address: Department of Environmental Sciences, Wageningen University, P.O. Box 47, 6700 AA Wageningen, The Netherlands.

PE&RC Training and Education Statement

With the training and education activities listed below the PhD candidate has complied with the requirements set by the C.T. de Wit Graduate School for Production Ecology and Resource Conservation (PE&RC) which comprises of a minimum total of 32 ECTS (= 22 weeks of activities)

Review of literature (4.5 ECTS)
- Measurement and modelling of nitrogen use efficiency of crops

Writing of project proposal (4.5 ECTS)
- Space-time monitoring and mapping of nitrogen use efficiency in China

Post-graduate courses (9.3 ECTS)
- Hands on digital soil mapping; PE&RC (2018)
- Fundamentals of crops physiology in a changing world; PE&RC (2018)
- Introduction to R for statistical analysis; PE&RC (2018)
- Design of experiments; PE&RC (2018)
- Multivariate analysis; PE&RC (2018)
- Cleaning data in R; online; WUR(2020)
- Data manipulation with dplyr; online; WUR (2020)
- Data visualization with ggplot2; online; WUR (2020)
- Error and uncertainty in spreadsheets; online; WUR (2020)
- Introduction to data visualization with ggplot2; online; WUR (2020)
- Introduction to Python; online; WUR (2020)
- Introduction to R; online; WUR (2020)
- Introduction to writing functions in R; online; WUR (2020)
- Spatial analysis with sf and raster in R; online; WUR (2020)

Deficiency, refresh, brush-up courses (18 ECTS)
- Spatial modelling and statistics; PE&RC (2018)
- Plant, vegetation and systems ecology; WUR(2018)
- Geo scripting; WUR (2020)

Competence strengthening / skills courses (6.3 ECTS)
- Effective academic development including academic writing and presenting in English; WGS (2017)
- Speaking skills; Wageningen In'to Languages (2018)
- Scientific writing; Wageningen In'to Languages (2018)

Scientific integrity / ethics in science activity (0.6 ECTS)
- Ethics for social sciences research; WGS (2018)
- Ethics in Plant and Environmental Sciences; WGS (2020)

PE&RC Annual meetings, seminars and the PE&RC weekend (1.5 ECTS)
- PE&RC First year weekend (2018)
- PE&RC Midterm weekend (2020)

Discussion groups / local seminars or scientific meetings (4.5 ECTS)
- Time series (2018)
- R Users (2018-2020)

International symposia, workshops and conferences (4.1 ECTS)
- International symposium on soil and plant analysis; Wageningen, the Netherlands (2019)
- Netherlands annual ecology meeting; online; the Netherlands (2021)

Lecturing / supervision of practicals / tutorials (1.2 ECTS)
- Environmental data collection and analysis (2020)

This project received funding from the National Key Research and Development Program (2016YFD0200101 and 2016YFD0200109), Fundamental Research Funds for Central Non-profit Scientific Institution (1610132019047), National Natural Science Foundation of China (31972515), the Earmarked fund for the China Agriculture Research System (No. CARS-09-P31). I am especially acknowledgeable to financial support from the China Scholarship Council (CSC) [201903250115].

Colophon

Cover and chapter figures design in cooperation with Mengyu Li (author’s niece).