
Implementing an Agro-Environmental Information System (AEIS) Based on GIS, Remote Sensing, and Modelling

- A case study for rice in the Sanjiang Plain, NE-China

I n a u g u r a l - D i s s e r t a t i o n

zur

Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität zu Köln

vorgelegt von

Quanying Zhao

aus Jilin, China

Köln, 2015

Berichterstatter: Prof. Dr. Georg Bareth

Prof. Dr. Karl Schneider

Tag der mündlichen Prüfung: 15.01.2016

Abstract

Information on agro-ecosystems is crucial for understanding the agricultural production and its impacts on the environment, especially over large agricultural areas. The Sanjiang Plain (SJP), covering an area of 108 829 km², is a critical food base located in NE-China. Rice, soya bean and maize are the major crops in the SJP which are sold as commercial grain throughout China. The aim of this study is to set up an Agro-Environmental Information System (AEIS) for the SJP by employing the technologies of geographic information systems (GIS), remote sensing (RS), and agro-ecosystem modelling.

As the starting step, data carrying interdisciplinary information from multiple sources are organized and processed. For an AEIS, geospatial data have to be acquired, organized, operated, and even regenerated with good positioning conditions. Georeferencing of the multi-source data is mandatory. In this thesis, high spatial accuracy TerraSAR-X imagery was used as a reference for georeferencing raster satellite data and vector GIS topographic data.

For the second step, the georeferenced multi-source data with high spatial accuracy were integrated and categorized using a knowledge-based classifier. Rice was analysed as an example crop. A rice area map was delineated based on a time series of three high resolution FORMOSAT-2 (FS-2) images and field observed GIS topographic data. Information on rice characteristics (i.e., biomass, leaf area index, plant nitrogen concentration and plant nitrogen uptake) was derived from the multi-temporal FS-2 images. Spatial variability of rice growing status on a within-field level was well detected.

As the core part of the AEIS, an agro-ecosystem modelling was then applied and subsequently crops and the environmental factors (e.g., climate, soil, field management) are linked together through a series of biochemical functions inherent in the modelling. Consequently, the interactions between agriculture and the environment are better interpreted. In the AEIS for the SJP, the site-specific mode of the DeNitrification-DeComposition (DNDC) model was adapted on regional scales by a technical improvement for the source code. By running for each pixel of the model input raster files, the regional model assimilates raster data as model inputs automatically.

In this study, detailed soil data, as well as the accurate field management data in terms of crop cultivation area (i.e. rice) were used as model inputs to drive the regional model. Based on the scenario optimized from field observation, rice yields over the Qixing Farm were estimated and the spatial variability was well detected. For comparison, rice yields were de-

rived from multi-temporal FS-2 images and the spatial patterns were analysed. As representative environmental effects, greenhouse gas of nitrous oxide (N₂O) and carbon dioxide (CO₂) emitted from the paddy rice fields were estimated by the regional model.

This research demonstrated that the AEIS is effective in providing information about (i) agriculture on the region, (ii) the impacts of agricultural practices on the environment, and (iii) simulation scenarios for sustainable strategies, especially for the regional areas (e.g. the SJP) that is lacking of geospatial data.

Zusammenfassung

Informationen zu Agro-Ökosystemen sind entscheidend für das Verständnis der landwirtschaftlichen Produktion und ihrer Auswirkungen auf die Umwelt, vor allem auf großen landwirtschaftlichen Flächen. Die Sanjiang Ebene (SJP), mit einer Fläche von 108 829 km², ist eine kritische Lebensmittelbasis in Nordost-China. Reis, Sojabohnen und Mais sind die wichtigsten Kulturpflanzen in der SJP, welche als Getreide in ganz China verkauft werden. Das Ziel dieser Studie ist es, ein Agro-Umweltinformationssystem (AEIS) für die SJP einzurichten, mit Hilfe des Einsatzes der Technologien der geographischen Informationssysteme (GIS), Fernerkundung (RS) und der Agrarökosystem-Modellierung.

Als Ausgangsschritt werden Daten mit interdisziplinären Informationen aus mehreren Quellen prozessiert. Für ein AEIS müssen Geodaten aufgenommen, organisiert, gehandhabt, und sogar mit guter Positionierung umgewandelt werden. Die Georeferenzierung von Daten aus unterschiedlichen Quellen ist obligatorisch. In dieser Arbeit wurden TerraSAR-X Bilder mit hoher räumlicher Genauigkeit als Referenz verwendet, um Raster-Satellitendaten und topographische Vektor-GIS-Daten zu georeferenzieren.

Für den zweiten Schritt wurden die georeferenzierten Daten aus unterschiedlichen Quellen mit hoher räumlicher Genauigkeit mit Hilfe eines wissensbasierteren Klassifikators integriert und klassifiziert. Als Beispiel-Feldfrucht wurde Reis analysiert. Eine Anbaukarte für Reis wurde erstellt, auf der Grundlage einer Zeitreihe von drei hochauflösenden FORMOSAT-2 (FS-2) Bildern und im Feld aufgenommenen topographischen GIS-Daten. Informationen zu den Charakteristika des Reis (d.h. Biomasse, Blattflächenindex, Stickstoffkonzentration und Stickstoffaufnahme) wurden aus den multitemporalen FS-2 Bildern abgeleitet. Die räumliche Variabilität des Reisanbaustatus innerhalb der Felder konnte zufriedenstellend detektiert werden.

Als Kernstück des AEIS wurde dann eine Agrarökosystem Modellierung angewendet und damit die Feldfrüchte und deren Umweltfaktoren (z.B. Klima, Boden, Bewirtschaftung) miteinander verbunden, durch eine Reihe von biochemischen Funktionen aus der Modellierung. Infolgedessen können die Wechselwirkungen zwischen Landwirtschaft und Umwelt besser interpretiert werden. In dem AEIS für die SJP wurde der ortsspezifische Modus des DNDC Modell auf regionaler Ebene durch eine technische Verbesserung des Quellcodes angepasst. Durch Anwenden für jeden Bildpunkt der Modelleingangsrasterdateien, assimiliert das regionale Modell die Rasterdaten automatisch als Modelleingaben.

Datenunsicherheiten über Böden und Bewirtschaftungsinformationen stellten sich als die wichtigsten Faktoren bei der regionalen Modellierungsanwendung heraus. Deshalb werden

in dieser Studie detaillierte Bodendaten, sowie die genauen Bewirtschaftungsdaten in Bezug auf die Pflanzenanbaufläche (z.B. von Reis) als Modelleingaben verwendet, um das regionale Modell zu betreiben. Basierend auf einem durch Feldbeobachtungen optimierten Szenario, wurden Reiserträge auf regionaler Ebene geschätzt und die räumliche Variabilität zufriedenstellend detektiert. Zum Vergleich wurden die Reiserträge aus multitemporalen FS-2 Bildern abgeleitet und die räumlichen Muster analysiert. Als beispielhafte Umweltwirkungen wurden die Treibhausgase Distickstoffoxid (N_2O) und Kohlendioxid (CO_2), die aus den Reisfeldern emittiert werden, von dem regionalen Modell geschätzt.

Diese Untersuchung zeigt, dass ein AEIS effektiv ist, vor allem für eine Region (z.B. SJP) in der Geodaten fehlen, indem Informationen über (i) die Landwirtschaft in der Region, (ii) die Auswirkungen landwirtschaftlicher Praktiken auf die Umwelt und (iii) Simulationsszenarien für nachhaltige Strategien bereitgestellt werden.

Acknowledgements

I am sincerely acknowledge my supervisor Prof. Dr. Georg Bareth, for his patient guidance, encouragement and advices through all my study time as his student. I have been so lucky to have a supervisor who is always available to assist me not only for scientific research but also for other aspects mentally and financially. Many thanks to him, for suggesting me many great training courses and providing the opportunities to attend conferences. His curiosity, determination and rigorous academic attitude will guide me deeply in all my future life. I would also like to thank all the members of staff at AG-Bareth and Universität zu Köln who helped me in my supervisor's absence. I know that I must made a lot of troublesome to them especially at the beginning of my study, but they are always warm and never complain. In particular, I would like to thank Dr. Victoria Lenz-Wiedemann for all her great suggestions and assist all through my Ph.D. study time, and many great thanks to her for taking care of me as a sister all the time. Many thanks to Christoph Hütt and Sebastian Brocks for their great contributions to the cooperated publications. Special thanks to Dr. Juliane Bendig, a beautiful, and kind vegetarian, for her great proof reading and assists in many of the official procedures for the thesis submission.

I must express my acknowledgements to all my Chinese supervisors and the Chinese Scholarship Council. Many thanks for their trusts, recommendations and supports to offer me such great opportunity to study further oversea. Specially, I must acknowledge my corresponding supervisor Prof. Dr. Fusuo Zhang, for all his assist during my master and Ph.D. study. He is always enthusiastic and full of energy which greatly incentivize me to work harder and contribute as much as possible to agriculture.

Particularly, I have to acknowledge Dr. Fei Yuan, for her suggestions for my study and great help in revising my publications.

I must express my gratitude to all my family and the close relatives, for their continued support and encouragement. I am so grateful for them to accompany me with all the ups and downs of my research.

Table of Contents

ABSTRACT	I
ZUSAMMENFASSUNG	III
ACKNOWLEDGEMENTS	V
TABLE OF CONTENTS	VI
1 INTRODUCTION	10
1.1 PREFACE.....	10
1.2 RESEARCH PROBLEMS AND OBJECTIVES.....	12
1.2.1 Research problems	12
1.2.2 Research objectives	14
1.3 OUTLINE	15
2 STUDY AREA AND DATA	17
2.1 THE SANJIANG PLAIN AND ITS AGRO-ECOSYSTEM PROBLEMS.....	17
2.1.1 Geographic conditions	17
2.1.2 Climate status	18
2.1.3 Soil conditions	19
2.1.4 Land use change	20
2.1.5 Water resource	22
2.1.6 Ecosystem service.....	24
2.1.7 Rice cultivation	25
2.2 DATA	25
2.2.1 Multi-source data	25
2.2.2 Data organization	26
3 METHODOLOGY	29
3.1 GEOGRAPHIC INFORMATION SYSTEM	29
3.2 SATELLITE REMOTE SENSING	30
3.3 KNOWLEDGE-BASED SYSTEMS	31
3.4 AGRO-ECOSYSTEM MODELLING	33
3.5 CONCEPT OF AN AGRO-ENVIRONMENTAL INFORMATION SYSTEM (AEIS)	33
4 GEOREFERENCING MULTI-SOURCE GEOSPATIAL DATA USING MULTI-TEMPORAL TERRASAR-X IMAGERY: A CASE STUDY IN QIXING FARM, NORTHEAST CHINA	35
SUMMARY	35
ZUSAMMENFASSUNG	36
4.1 INTRODUCTION.....	37
4.2 STUDY AREA AND DATA.....	38
4.2.1 Study area.....	38

4.2.2	Data description	40
4.3	METHODS	41
4.3.1	Workflow of georeferencing multi-source datasets	41
4.3.2	Creation of the reference image from TSX stripmap acquisitions	42
4.3.3	Georeferencing of topographic vector data	43
4.3.4	Georeferencing of optical RS data	44
4.4	RESULTS	45
4.4.1	Georeferencing results of topographic vector data	45
4.4.2	Georeferencing results of optical RS data	45
4.4.3	Spatial accuracies of the georeferenced optical RS data	46
4.5	DISCUSSION	47
4.5.1	Analysis of the anticipated spatial error in the processed TSX reference image	47
4.5.2	Quantified spatial accuracy of the georeferenced datasets	48
4.5.3	Feasibility of the approach	49
4.6	CONCLUSIONS	50
	ACKNOWLEDGEMENTS	50
	REFERENCES	52
5	INVESTIGATING WITHIN-FIELD VARIABILITY OF RICE FROM HIGH RESOLUTION SATELLITE IMAGERY IN QIXING FARM COUNTY, NORTHEAST CHINA	56
	ABSTRACT	56
5.1	INTRODUCTION	57
5.2	STUDY AREA	61
5.3	DATA	61
5.3.1	Satellite RS images and GIS data	61
5.3.2	Ground truth data collection	62
5.4	METHODS	64
5.4.1	Satellite image pre-processing	64
5.4.2	Mapping rice cultivation areas	65
5.4.3	Ground truth data interpolation	66
5.4.4	Development of regression models for deriving agronomic variables	66
5.4.5	Validation of the regression models	67
5.5	RESULTS	68
5.5.1	Accuracy of rice area classification	68
5.5.2	Empirical regression models	68
5.5.3	Regional application of the regression models	72
5.6	DISCUSSION	77
5.6.1	Band selection for different growth stages	77
5.6.2	Background effects in the early stage	77
5.7	CONCLUSIONS	79
	ACKNOWLEDGEMENTS	80
	AUTHOR CONTRIBUTIONS	80

CONFLICTS OF INTEREST	80
REFERENCES.....	81
6 DETECTING SPATIAL VARIABILITY OF PADDY RICE YIELD BY COMBINING THE DNDC MODEL WITH HIGH RESOLUTION SATELLITE IMAGES.....	88
ABSTRACT	88
6.1 INTRODUCTION.....	89
6.2 MATERIALS AND METHODS.....	90
6.2.1 Study area.....	90
6.2.2 Field data	91
6.2.3 The DNDC agro-ecosystem model	92
6.2.4 Remote sensing approach.....	97
6.2.5 Statistical analysis	98
6.3 RESULTS	99
6.3.1 Site-specific model application and site validation	99
6.3.2 Soil-specific validation.....	101
6.3.3 Validation of RS-derived rice yield.....	102
6.3.4 Comparison of modelled and RS-derived yields	102
6.4 DISCUSSION.....	103
6.4.1 Model regionalization	103
6.4.2 Soil effects on DNDC modelled rice yield	104
6.4.3 Sources of uncertainty for regional model application	105
6.4.4 RS-derive rice yield	106
6.5 CONCLUSIONS	106
ACKNOWLEDGEMENTS.....	107
REFERENCES.....	108
7 DISCUSSION	113
7.1 MULTI-SOURCE DATA ANALYSIS	113
7.2 ACCURACIES OF RS-DERIVED CROP INFORMATION	116
7.3 SPATIAL SCALES FOR AN AEIS	117
7.4 MODEL UNCERTAINTY.....	118
7.5 ENVIRONMENTAL EFFECTS	120
7.6 SUMMARY	121
8 FUTURE CHALLENGES AND OUTLOOK.....	122
8.1 REQUIREMENT OF A STRONG NETWORK FOR GATHERING SUFFICIENT HIGH-QUALITY, IN- TIME DATA	122
8.2 IMPROVE MODEL CAPABILITY	122
8.3 ASSIMILATE RS-DERIVED CROP INFORMATION INTO PROCESS-BASED AGRO-ECOSYSTEM MODELLING ON LARGE EXTENT.....	122
8.4 DESCRIBING REGIONAL LONG-TERM STORIES USING THE AEIS.....	123
REFERENCES*CHAPTERS 1,2,3,7,8	124

LIST OF FIGURES AND TABLES.....	142
APPENDIX A: EIGENANTEIL ZU KAPITEL 4	145
APPENDIX B: EIGENANTEIL ZU KAPITEL 5	146
APPENDIX C: EIGENANTEIL ZU KAPITEL 6	147
APPENDIX D: ERKLÄRUNG.....	148
APPENDIX E: CURRICULUM VITAE	149
APPENDIX F: PUBLICATION LIST.....	150

1 Introduction

1.1 Preface

The global food demand is continuously increasing under the pressure of the increasing world population. Simultaneously, the total area of agricultural land in the world is decreasing due to urbanisation, desertification, water scarcity, and climate change (Lambin and Meyfroidt, 2011; Foley et al., 2005; Vörösmarty et al., 2000). Consequently, the gap between the increased food needs and the decreased agricultural areas results in a more intensified management in agriculture, which severely affects the agro-environmental resources in a non-sustainable manner (Tilman et al., 2011; Poudel et al., 2013; Macary et al., 2013; Popp et al., 2013; Naeem et al., 2012; Ronald, 2011; Georghiou, 2012). Arguably, there is a great potential to boost crop yield, as well as spare resources and reduce environmental consequences through optimizing management (Seufert et al., 2012; Stafford, 2000; Schaller, 1993), particularly in the developing countries, such as China, the loss of agricultural area is severe and food production is highly depending on intensive management (Ju et al., 2009; Zhang et al., 2013). Especially in China, where only 8% of the world's agricultural area has to feed up to 20 % of the world population (Smil, 1999), the pressure on the (agro-)environment increased significantly in the last three decades (Jiang et al., 2013; Siciliano, 2012). Therefore, it is of key importance in China to optimize inputs e.g. for crop production, to increase crop yield and to reduce environmental effects (Zhang et al., 2004; Zhu and Chen, 2002).

As one of the major food crops, rice feeds over half of the world population (FAO, 2014) and is especially important in China due to its long cultivation history (> 7000 years) (Cao et al., 2006) and its potential for high yield (Yuan, 1997). Particularly, in the areas that are rich of water resource and suitable climate, i.e. the north eastern part of China, *japonica* rice cultivation areas were increased in the past decades because of better economic profits, good quality, and excellent taste (Kako and Zhang, 2000; Wang et al., 2013). Regarding food security issue, improving rice yield is still the major concern of China for a long time (Peng et al., 2009). However, paddy rice under intensive management affects the environment significantly due to overuse of fertilizer and pesticide, consumptions of large amount of water, and greenhouse gas emissions (Zhang et al., 2012; Pingali and Roger, 2012; Zwart and Bastiaanssen, 2004). Thus, efforts of improving rice yield paralleled with reducing environment effects becomes the focus of agronomists in recent decades. For instance, to improve soil fertility, optimize field management strategies, and reduce greenhouse gas emissions etc., are common suggestions that have been continuously addressed (e.g., Schmidt et al., 2011;

Huang et al., 2010; Li, 2010). However, the agro-ecosystems are complex and cross-relationships exist among the agro-environmental factors (i.e., crop, climate, soil, water, management strategies), thus an integrated framework based on multi-disciplines is needed to provide agro-environmental information that may contribute to sustainable agriculture (van Ittersum et al., 2008; Schaller, 1993; van Cauwenbergh et al., 2007).

To understand and manage the complex agro-ecosystem for a balance between the competing needs on the increase of food productivity and on the maintenance environment/resources, agro-environmental resource management systems have to be implemented (McCloy, 2005). In this study, an Agro-Environmental Information System (AEIS), in the content of a spatial environmental information system (SEIS) (Bareth, 2009), was implemented to investigate and support the management of the agro-ecosystem resources. Specifically, the AEIS in this research refers to a spatial data infrastructure (SDI) that is organized to drive regional agro-ecosystem models. An SDI categorize and integrate geospatial data from multiple sources, to maximize the potential value of the available geospatial data, especially for a data poor region (Bareth, 2009).

This work was conducted in the Sanjiang Plain (SJP) in north eastern China, a significant commercial food base in China. The SJP is also one of the few areas that are managed under modern agricultural mechanization in China (Wu et al., 2007). During the past 60 years, the landscapes of the SJP changed dramatically from wetlands to farmlands, accompanied by sharp changes in the agro-environmental factors (e.g., soil, water, climate) (e.g., Yan et al., 2002; Huang et al., 2010; Wang et al., 2009; Zhang et al., 2007). Thus an AEIS is highly required in the SJP to guarantee food production, and to describe and manage the agro-environment.

To implement an AEIS in the SJP, we follow the logic pyramid of ‘data-information-knowledge-understanding-decision’ system (Rowley, 2007) by acquiring, organizing, building, operating and regenerating geospatial data. As a remote area in China, the SJP is poor in spatial data availability in the past. Nonetheless, managers including local officers, regional decision makers, and even farmers, are showing increasing interests in using geospatial data to analyze the agriculture status and to make decisions (e.g., Liang et al., 2013; Ma, 2015; Lu et al., 2002). For the AEIS in the SJP, it is a critical task to deal with the geospatial data regarding data availability, coverage, classification, accuracy and inconsistency.

Research of agricultural science is characterized by a high fragmentation in research methods and tools, and the integration of the diverse methodologies and techniques is required in the agro-ecosystem research (van Ittersum et al., 2008). Technically, the AEIS is implemented

based on geographical information systems (GIS), remote sensing (RS) and modelling. GIS technology provides a flexible environment for storing, analyzing, and displaying digital data necessary for an agro-environmental information database (Devogele et al., 1998; Güting, 1994). Satellite RS provides cost-effective multi-spectral and multi-temporal data, and turns them into information valuable for monitoring and understanding crop growth status (Moran et al., 1997). Agro-ecosystem models describe the exchange process and matter fluxes in the soil-vegetation-atmosphere system by calculating the crop growth as functions of various environmental factors, such as soil, climate, water and fertilizer (Baldocchi et al., 2002; Wang et al., 2014; Chen et al., 2014).

Specially, in the agriculture system research, the difficulties in the cross-scale issue, such as model application from one scale hierarchy to another, are recognized in previous studies (e.g., Dalgaard et al., 2003; Hansen and Jones, 2000; Resop et al., 2012). Information of the farmers' response on farm level/scale are vital in the research of an agro-ecosystems, for instance the climate change effects on agriculture (Reidsma et al., 2010). The AEIS in this study has to provide not only regional agro-environmental information for the managers, but also precise field-level information and management suggestions particularly for the farmers. Thus, the site-specific agro-ecosystem models have to be generalized onto a regional scale by e.g. creating additional computer scripts. Integrated by the geospatial data of crop information and environmental information, the AEIS provides and predicts agro-environment information at fine resolutions.

1.2 Research problems and objectives

1.2.1 Research problems

Spatial data in a well-organized structure are necessary to provide effective mechanisms for data storage, investigation, transfer and archiving (Mückschel et al., 2008; Curdt, 2014). However, problems of spatial inconsistencies inherent in the multi-source data are inevitable (Verburg et al., 2011; Li, 2010). Due to technical or political reasons (Bareth and Yu, 2002), a lot of spatial data for certain regions may not be available, especially the high resolution data. For instance, in the SJP, soil data of 1:1 000 000 scale are available for the total area, whereas the 1:200 000 scale is not available for every county. In this study, soil data of 1:200 000 are available only in 12 of the 23 counties in the SJP. In some cases, accurate information tends to be deficient because the available data generally have to be produced by many different people, with varying interests and perspectives, archived in different formats, and generally collected at different times (McCloy, 2005). Previous studies explored methods for pre-pro-

cessing and organizing multi-discipline data for soil-vegetation-atmosphere interactions research (e.g., Curdt, 2014; Reinartz et al., 2011). Expert systems are effective methods in processing data of multiple disciplines which have been implemented to integrate remotely sensed data with other data, to get land use classifications in higher accuracy (Stefanov et al., 2001; Wentz et al., 2008). To date, there is still a lack of research focusing on the transferable methods for georeferencing spatial data of different formats (e.g., raster, vector).

The second problem for an AEIS is to generate the accurate spatial information including soil, climate, water, as well as field management. Field management is critical in agro-ecosystem model applications on regional scale (Kersebaum et al., 2007). Accurate crop cultivation area classification from high resolution satellite images can be used to link crop-specific field management information to regional models (Waldhoff and Bareth, 2009). Besides, identifying crop cultivation areas accurately from high resolution images is also critical for crop monitoring (Jin et al., 2015). For a specific crop area delineation, Waldhoff (2014) demonstrated that the implementation of multi-data (i.e., RS raster data, GIS vector data) can significantly improve the crop area classification, although there might be a lack of GIS data in certain regions. In the AEIS for the data-poor SJP, methods of integrating GIS and RS data have to be explored to improve the usage of the available data. Spatial soil data are important not only crop for yield production but also for agro-ecosystem models (Eitzinger et al., 2004; Li et al., 2004; Kersebaum et al., 2007). Soil data of fine resolution and high (spatial) quality are most difficult to obtain because of the ‘costly’ data collection, and some ‘sensitive’ local policies (Bareth and Yu, 2002). Therefore, to merge, calculate, assign soil properties based on the available soil information, for instance to generate the soil hydraulic properties based on soil type and soil organic matter information, are one of the key steps to implement an AEIS in a data poor region.

Third, to get the knowledge of the development status and appearance of the crops through a year(s) is fundamental to the AEIS in the SJP. As an advanced technology, satellite RS imagery with coarse and medium resolution are widely used in rice cultivation research (Kuenzer and Knauer, 2013; Jin et al., 2015; Zhang, et al., 2015). However, studies conducted on rice using high resolution RS images were limited in the past two decades (Kim and Yeom, 2012; Chang et al., 2013). There is still a lack of investigations for detecting rice spatial variability on a within-field level with reliable accuracy.

Last but not least, cross-scale issue is an obstacle with endless complex for modelling application in multiple spatial scales (Ostrom, 2007; Sayer et al., 2013). Process-based agro-ecosystem models can be implemented in interdisciplinary projects to investigate the interac-

tions between complex components by assimilating data from diverse sources and represented in diverse formats (Câmara, 1996; Schreinemachers and Berger, 2011; An, 2012). However, most of the models are designed to simulate at site or field scales. Detailed information from the site model is also required in large geographic areas. The DeNitrification-DeComposition (DNDC) model is a process-based geochemical agro-ecosystem model which simulates the C and N cycle in agro-environment ecosystems (Li et al., 1992; Giltrap et al., 2012). Although the DNDC model was developed with both a site-specific mode and a regional mode, the site-specific mode was assumed to be more flexible and transparent (Perlman et al., 2013). In the regional mode, the research region is divided into small sub-units based on the assumption that the attributes in each unit are uniform. The model merges the results from all units to obtain a regional result. Whereas in the site-specific mode, more site-specific model parameters can be adjusted to decrease the model uncertainties. To transfer the site models onto a regional scale without losing detailed simulation information is beneficial for an agro-environmental study. To implement the AEIS in the SJP, generalisation of the site-specific model has to be innovated.

1.2.2 Research objectives

The overall purpose of this study is to investigate the potential of available spatial data to serve regional agro-ecosystem modelling for a data poor environment. Agro-ecosystem modelling facilitates to explore and to better understand the interactions among the factors for agricultural production (e.g., soil, water, crop, climate, etc.) which might support decision making for an optimized regional planning and management of crop production. Consequently, a key task is the acquiring, organizing, operating and generalizing of geospatial data, or the set-up of a SDI which fulfills the demands of the modelling purpose. The SDI and the interfaced model (models or the modelling scenarios) together form an Agricultural Environmental Information System (AEIS) (Bareth, 2009).

In this study, a state-owned farm, Qixing Farm, which is located in the SJP, NE-China, is selected a study area. In terms of available and accessible (geo)data, the study area can be considered as a data poor environment region. Therefore, (i) the set-up of a SDI for a regional modelling is first task. This includes to 1) acquire and organize data from multiple sources; 2) georeference the multi-source geospatial data to overcome the inherent spatial heterogeneity; 3) derive accurate rice maps by combining GIS and RS data for monitoring rice growing status and linking field management information to the accurate rice map in agro-ecosystem modelling; and 4) detect the temporal rice status with high spatial resolution satellite imagery. The second task is to (ii) implement agro-ecosystem models (i.e. DNDC)

for estimating paddy rice yield at a regional scale. Then the AEIS can be used to (iii) investigate the effects of environmental factors (i.e., soil) on rice yield. Finally (iv) the impacts of agricultural production on the environment can be analyzed from model scenarios and thus to provide information for reducing environmental effects.

1.3 Outline

The introduction chapter (**chapter 1**) presents the background of this study and the research aims. The SJP and its agro-environmental problems are addressed in **chapter 2**. The methodology for an AIES is introduced in **chapter 3**, which is followed by the main parts (**chapters 4 – 6**) that are based on three published or submitted journal papers.

In **chapter 4** (Zhao et al., 2015a), a georeferencing method of using TerraSAR-X imagery as ground control information is described to pre-process the geospatial data from multiple sources. To eliminate positional discrepancies in different geospatial datasets from multiple sources, multi-temporal TerraSAR-X imagery was processed as a referencing image. Using this method, both topographic data and raster satellite images data are georeferenced with high spatial accuracies. This approach promotes the implementation of an AEIS that is based on the integration of multiple data.

In **chapter 5** (Zhao et al., 2015b), an accurate rice cultivation area map for the Qixing Farm was derived using a Multi-Data Approach (MDA). An expert classifier was applied in the MDA to integrate GIS boundary data and multi-temporal (FORMOSAT) FS-2 images. The final classification accuracy for the rice map is >91 %. Besides, rice growing status was monitored on a within-field level using FS-2 images. As a data preparation procedure in **Chapter 5**, atmospheric correction and georeferencing of the FS-2 images are presented. Then empirical multiple linear regression (MLR) models are constructed to relate the satellite reflectances to rice parameters including weight of biomass, LAI values, plant nitrogen (N) concentrations and plant N uptake. Maps of rice status representing within-field variability are retrieved based on the strong relationships between the image reflectance information and the rice parameters.

In **chapter 6** (Zhao et al., submitted), spatial variability in rice yields at the Qixing Farm are estimated from process-based geochemical modelling DNDC and the FS-2 imagery. First the site-specific mode of the DNDC is applied and assessed using the detailed field measured data. Then the site-specific model is generalized onto a regional scale by creating additional scripts in Python environment. Two additional scripts are created for the model to process raster files automatically. Detailed spatial soil data (100 m × 100 m) are prepared as model input data. Based on the detailed soil data and the accurate rice cultivation map (described

in **chapter 5**), spatial variability in modelled rice yield was detected. Additionally, rice yield is derived from multiple FS-2 images using empirical MLR model. Finally, the modelled and RS-derived rice yields are compared and assessed. Advantages and drawbacks of both methods are discussed regarding the process-based DNDC model and the empirical MLR model.

In **chapter 7**, key problems regarding the AEIS are discussed. The advantages and limitations in the application of multi-source data are documented. Specially, as important results from the regional modelling, greenhouse gas (e.g., CH₄) emissions are analyzed under different field management strategies and environmental conditions. Drawbacks inherent in the ‘process-based’ agro-ecosystem model are pointed out.

Chapter 8 outlines the forthcoming research tasks and the future opportunities to promote the AEIS for the SJP.

2 Study Area and Data

2.1 The Sanjiang Plain and its agro-ecosystem problems

2.1.1 *Geographic conditions*

The SJP (**Figure 2-1**) locates (129°11' – 135°05'E, 43°49' – 48°28'N) in the northeast part of the Heilongjiang province, NE-China. It is an alluvial plain formed by the Songhua river, Heilong river and Wusuli river. The Heilong river forms the international border with Russia to the north, and the Wusuli river forms another boundary with Russia to the east.

The SJP consists of two sub alluvial plains separated by the Wanda mountain. The plain located in the north of the Wanda mountain, named 'the low Sanjiang plain', is formed by Songhua river, Heilong river and Wusuli river. The other plain, named 'Muling-Xingkai plain', located in the south of the Wanda mountain is formed by Xingkai lake and the south part of Wusuli river and its branches. The total area of the SJP is 108 829 km² with a plain area of 51 300 km². The area of The SJP exceeds the area of the Netherlands by 2.6 times and approaches one third of the total area of Germany.

The elevation of the SJP in the southwest is higher than in the northeast due to the low hills and mountains in the south and west part. The average elevation is about 45-80 m above sea level. Most of the rivers in the area have slight gradients and large channel curve coefficients.

The Qixing Farm locates in the central part of the SJP. It is a typical nation-owned farm which is a technique-pilot farm in rice cultivation in the SJP.

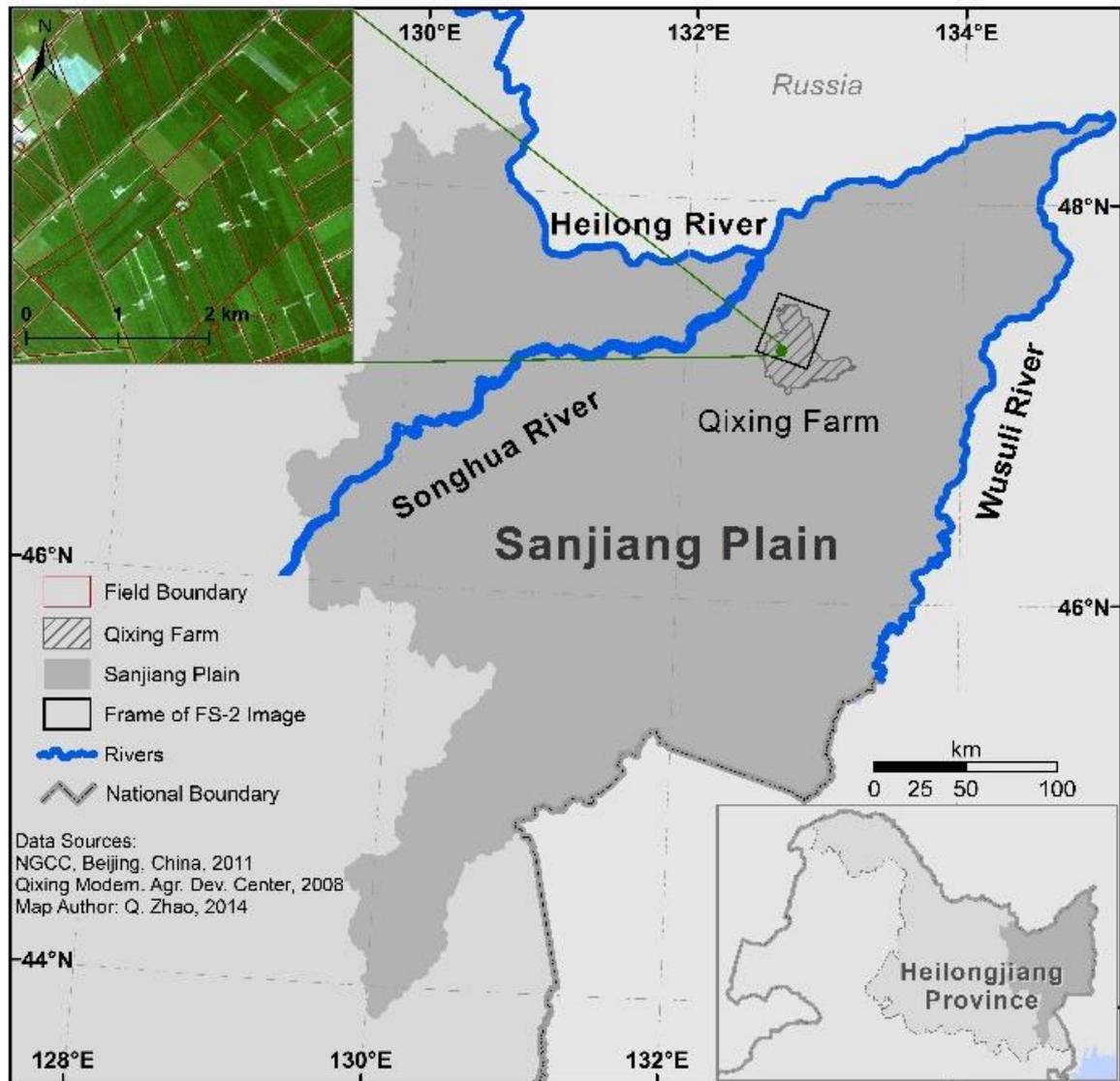


Figure 2-1: The study area Sanjiang Plain in Heilongjiang province, NE-China (Zhao et al., 2015b).

2.1.2 Climate status

The climate in the SJP belongs to the temperate humid or sub-humid continental monsoon climate. The mean annual temperature ranges from 1.4 to 4.3 °C, with an average maximum of 21 – 22 °C in July and an average minimum of -18 °C in January. The mean annual precipitation is 500-650 mm and 80% of the rainfall occurs between May and September. The frost-free period is 120 – 140 days.

During the last decades, the annual temperature in the SJP experienced a regional risen which was greatly contributed to warming up in winter and spring, paralleled with a decrease in annual precipitation (Yan et al., 2002; Luan et al., 2007).

The SJP is divided into three accumulated temperature zones according to the annual accumulated temperature of ≥ 10 °C. One zone locates in the northwest part of the SJP, with an annual accumulated temperature of about 2400 °C. The second zone locates in the south part of the SJP, with an annual accumulated temperature of about 2500 °C. The third zone locates in the east part of the SJP, with an annual accumulated temperature of about 2300 °C (Zou et al., 2010). The annual accumulated temperature increased in all zones (Zou et al., 2010). The frost-free days per year are 120 – 140 days (Huang et al., 2010). Under the global warming trend, there is climate warming up in the SJP, which benefited the rice cultivation in this area. Gong et al. (2015) investigated the cool injury for rice during the past 50 years from 1961 to 2010 in the Heilongjiang province. They found the center of the cool injury area has been decreased and moved northward. On the other hand, agriculture may affect the climate as well. Studies (Huang et al., 2010; Liu et al., 2013; Xu et al., 2015) show interactions between the agricultural land use and the climate in the SJP. Paddy rice decreases the greenhouse gas emission, compared to dry land cultivation (Huang et al., 2010).

2.1.3 Soil conditions

There are five main soil types in the SJP: black soil, meadow soil, albic soil, bog soil, and dark brown soil. These five soils occupy more than 95% of the whole area (Nachtergaele et al., 2008). The black soils are mainly distributed in areas with moderate slopes in Fujin, Baoqing, Jixian, and Jiamusi. More than 80 % of the black soils have been reclaimed. The thickness of the black soils are around 75 – 95 cm. They are favorable for agriculture because of the physical and chemical characteristics. The meadow soils are the azonally distributed soils in the SJP. They are distributed in the flat areas in the counties of Fujin, Jixian, Baoqing, and the national farms of Youyi, Wujiuqi and Erjiuyi. More than 55 % of the meadow soils have been reclaimed. The black soil layers of the meadow soils reach up to 50 – 100 cm. However, improvement of agricultural use for meadow soils are needed because the heavy clay soil characteristics make them prone to waterlogging. Albic soils mainly distribute in the Muling-Xingkai plain and the Fuyuan delta. They are the main arable land soil. The black soil layers of the albic soils are normally within a thickness of 10 – 20 cm. An area of more than 43 % has been reclaimed. Bog soils distribute in the low areas of water pools, covered by water seasonally or annually. The surface of the bog soils are peat layers of various thickness. Currently, 15 % of the bog soils are reclaimed. The dark brown soils distribute in the mountainous and hilly areas. They can only be used as forest land due to the terrain topography (Song et al., 2010).

Soil quality in the SJP was degrading regarding many of the soil elements such as N, phosphorus (P), and potassium (K) (Wang et al., 2009; Wang, 2005). Soil organic carbon (SOC) has significantly decreased due to the tillage cultivation of agriculture (Zhang et al., 2007). Soil erosion is another important issue that has received widely investigation (Zhang et al., 2008; Liu and Ma, 2000).

In order to select appropriate sustainable strategies for preventing the detrimental effects of agriculture on soils, i.e. soil erosion, desertification, salinization, compaction, pollution, etc., research should focus on development of an accurate soil quality monitoring and evaluation system at multiple scales (Wang et al., 2009).

2.1.4 Land use change

The SJP was a virgin area before the 1950s. After the foundation of the People's Republic of China, as Wang et al. (2006) mentioned, there were three important events that facilitate the reclamation in the SJP (**Figure 2-2**). The first event was the 'Great Leap Forward' movement. From 1956 to 1960, about 81 500 veterans were encouraged to go to the SJP to reclaim wetlands aiming to enlarge the agricultural area and improve food production. The second event was the 'Cultural Revolution'. From 1970 to 1972, about 450 000 educated young people entered this region for agriculture in response to the 'going to the countryside and settling in the communes' movement. The third event was the 'reform and open' policy of 1979. Due to the policy of the agricultural modernization, some farms were selected for pilot projects of modern agricultural farming. The application of advanced agricultural machinery improved the grain production as well as the reclamation efficiency (Wang et al., 2006; Luan and Zhou, 2013; Hohlr, 1991). The wetland has been fragmented seriously due to its significant loss (Wang et al., 2011).

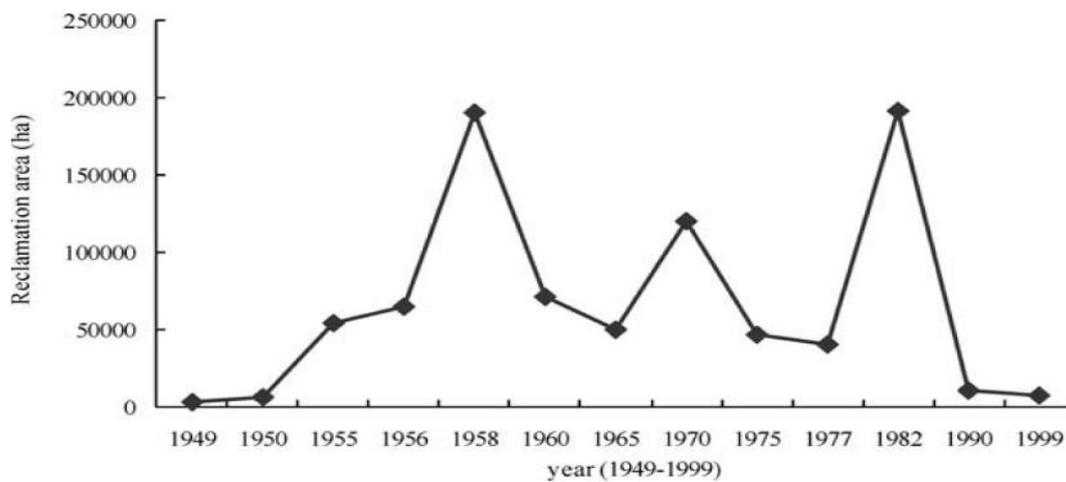


Figure 2-2: The reclamation area in the SJP from 1949 to 1999 (Wang et al., 2006).

Table 2-1 shows an overview of the land use change in the SJP during the 1980s until 2000. Specially, paddy rice area was increased by more than 1.5 times from 1995 to 2000.

Table 2-1: Land use change (%) for selected classes from the 1980s until 2000 in the Sanjiang Plain for the representative counties (Zhao et al., 2012).

Land use change (%) in the Sanjiang Plain										
	Paddy field		Dryland		Grassland		Wetland		Others	
Time	1980s – 1995	1995 – 2000	1980s – 1995	1995 – 2000	1980s – 1995	1995 – 2000	1980s – 1995	1995 – 2000	1980s – 1995	1995 – 2000
Baoqing	-88.5	1097.7	73.8	-40.3	-18.3	-17.0	-24.9	-11.2	-1.1	-3.8
Luobei	216.9	550.7	15.4	-28.6	-47.5	15.5	-27.3	-8.3	0.3	-3.1
Mishan	-56.6	185.4	24.3	-13.5	-2.9	-38.2	-7.4	-14.9	3.5	-6.6
Fuyuan	6.5	161.9	61.5	9.5	-89.3	-87.6	-7.3	-7.4	10.7	-1.4
Hulin	18.6	145.7	45.0	-4.9	-70.3	33.5	-3.0	-22.2	7.5	-2.2
Yilan	-14.9	80.5	11.6	-10.0	-54.9	9.9	4.2	-11.3	-0.6	-1.0
Huachuan	61.0	73.2	-7.3	-12.7	280.6	-6.0	4.6	-56.3	-16.1	-4.5
Tangyuan	-13.5	62.4	8.9	-14.8	16.6	-37.7	70.9	-41.5	-9.8	9.2
Fujin	99.1	43.3	3.6	-0.2	-29.8	23.6	-8.3	-3.8	1.3	-0.7
Youyi	1204.4	42.4	2.6	3.1	33.9	-25.3	-45.0	-44.3	5.8	-0.5
Muling	-22.3	28.7	-3.2	10.4	3.6	-2.5	no wet- land	no wet- land	1.8	-3.7
Total change	-28.3	157.2	14.6	-7.7	-45.0	2.4	-11.3	-10.6	1.6	-3.9

The groundwater table in the SJP is normally within a depth of ≤ 5 m. The hydraulic gradient of the groundwater is about 1/1000 in the hilly region, 1/5000–1/10000 in the central plain, and 1/500–1/150 near the river. The groundwater discharges into the rivers and lakes. The flow direction of groundwater is from southwest to northeast in the low Sangjiang Plain and the groundwater flows from northwest to southeast in the Muling-Xingkai plain (Song et al., 2010).

In the year 2009, the total agricultural water use was 75.9×10^8 m³, accounting for 97.1% of the total water consumption of industry, agriculture and citizens. 97 % of the irrigation area was cultivated with paddy rice and 69 % of the total paddy rice area use groundwater (Song et al., 2010).

Jiang et al. (2011) investigated the water resource carrying capacity in the SJP by creating a water evaluated index system and criteria to integrate a water resource system, a social system, an economic system, and an ecosystem. They found that in the middle part of the SJP, along the Wanda mountainous area, the water carrying capacity is high. Whereas in the south and north part of the SJP, where the two sub-plains are located, the water carrying capacity is moderate and the water resources exploitation and socioeconomic development are compatible.

Taking the Jiansanjiang area as an example, the groundwater table decreased in the past decades, coupling with an increase of the rice cultivation area (**Figure 2-4**).

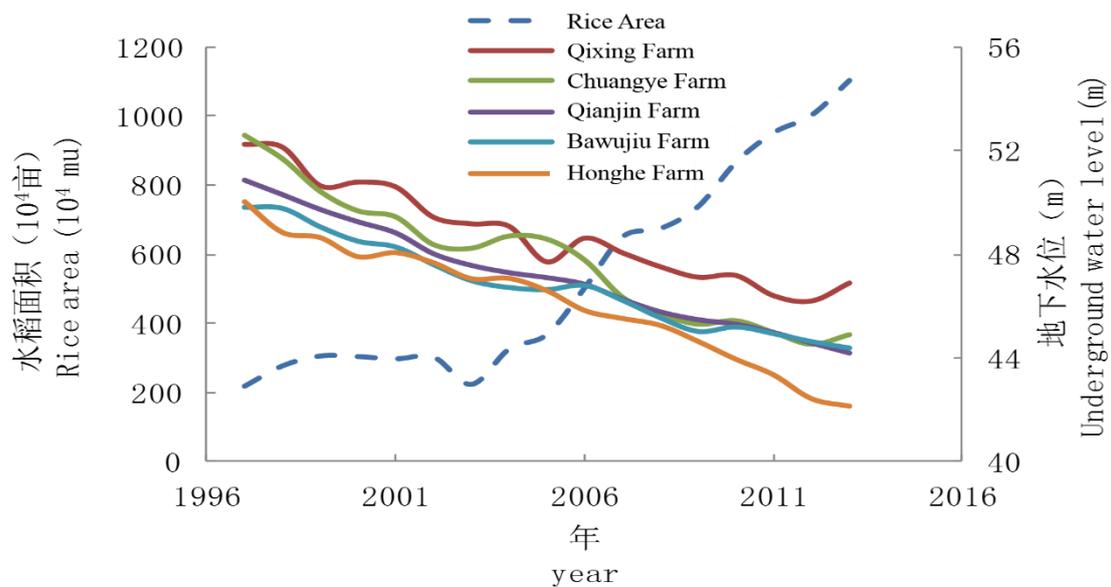


Figure 2-4: Change of groundwater in the Jiansanjiang area, coupled with the increased rice cultivation area (Ren 2014).

2.1.6 Ecosystem service

The pressure of agriculture and ecosystem services are especially severe in the developing world where human-environmental interaction patterns are very dynamic and under stress due to the rapid economic, socio-demographic and technological changes (Fegraus et al., 2012).

Wetlands have multiple functions in ecosystems as they can improve water quality, provide flood control, mitigate climate change, and assist in groundwater recharge. (Mitsch and Wilson, 1996; Munyati, 2000, An et al., 2007). National reserves are important ‘storage of natural genes’ (Zhou et al., 2009) as it is a habitat for various plants and wild animals.

Before the agricultural reclamation of the SJP, the upland areas were covered by the natural vegetation of dense and mixed coniferous forests and broad-leaf forests. The vegetation in the river floodplains and lowlands included an immense expanse of freshwater wetlands, reed marshes, wet sedge meadows, grass meadows, lakes, and riparian willows and other wet forests (Zhou and Liu, 2005). Due to its noteworthy rich biodiversity, the SJP wetlands are ranked as globally important in the *Directory of Asian Wetlands*. There are about 1000 species of plants, 37 ecosystem types and 528 species of vertebrate fauna in this area (Ni and Li, 1999). Some wildlife species ranked by the World Conservation Union as globally threatened are found in these wetlands. However, according to Wang et al. (2006), the ecosystem service values in the SJP have declined by 40 % between 1980 and 2000, attributing to the 53.4 % loss of wetland.

Human activities, especially agricultural reclamation, are considered as major threats to the wetland ecosystems in the SJP. Since the late 1990s, the government has been fully aware of the seriousness of environmental problems resulting from damages to natural ecosystems. They changed the ‘food first’ agricultural policy to an ‘environmental friendly agriculture’ policy. The idea of ‘farmland back to wetland’ and ‘construction of an ecological province for Heilongjiang province’, etc., were adopted. Studies tried to find proper sites for turning agricultural fields back to wetland according to the wetness index and agricultural productivity (Huang et al., 2010). Various measures have been applied to prevent and cure the severely disturbed environment and a conservation area has been built such as the Xingkai Lake, Honghe, and Sanjiang national nature reserves.

2.1.7 Rice cultivation

There are 23 counties and 52 nation-owned farms in the area, with a population of 8.66 million in 2010, among which 41 % were engaged in farming (HSB and HSTNB, 2011). 91% of the cultivation area in the Jiansanjiang area are rice (HSB and HSTNB, 2011).

The agriculture in the SJP faces the problem of serious waterlogging because of the flat geographical topography and clayed soil conditions. In the past, rice was used to offset the field waterlogging. Since 1992, when the market-directed economic system replaced the former planned economic system, paddy rice areas were increased dramatically. Driven by the higher economic efficiency and the availability of technologies for rice cultivation, many drylands were converted into paddy rice fields.

In the more than 30-years rice cultivation history in the SJP, people try their best to investigate rice cultivation mechanisms. To produce rice with higher yield, higher quality, while using less fertilizer inputs, series of rice cultivation regimes have been developed in the past decades. Wang et al. (2011) proposed an optimum rice cultivation regime for the albic soil area in the SJP. Cao et al. (2005) investigated field measures for high yield rice. Liu (2010) introduced an optimum rice cultivation regime for NE-China which emphasized the knowledge of pest and disease control. Water saving irrigation coupled with rice cultivation skills were also studied (Lv et al., 2014). These research studies emphasized several steps of paddy rice cultivation including the seed pre-germination procedure, the seedlings greenhouse nursery in early spring, key skills in the transplanting process, fertilization time and amount, irrigation strategies, pest-weed-disease control and harvest time. People in the SJP have to explore skill-intensive strategies to avoid wasting solar energy, since the growth season is short (130 – 135 days) and the frost-free period is even shorter (120 – 140 days).

2.2 Data

2.2.1 Multi-source data

To implement and AEIS, data have to be organized from various governmental bureaus or non-governmental organizations such as local research institutions or special research groups. In this study, datasets from different sources, each characterized by their unique attributes and properties, are referred to as multi-source data.

2.2.2 Data organization

- **Topographic data**

In this study, detailed topographic datasets for the whole SJP on a scale of 1:250 000 were purchased from the National Geomatics Center of China (NGCC). In which ground features of land use, elevation, cities, roads, waters, etc., are represented. Since the AEIS in this study aim to provide information also for farmers at the field level, more detailed topographic data at the field level are needed. Thus a set of field boundary data of Qixing Farm in an independent coordinate system was provided by the Qixing Modern Agriculture Research Center. Unfortunately, the scale of this data is not clearly defined. This field boundary data provides information on field borders, building areas, water pools (for storing irrigation water), as well as field ownership and crop types of the fields.

- **Climate data**

Climate data were organized for the SJP. For the model application, daily weather data collected from the nearest weather station-Jiansanjiang weather station (located in the northern part of Qixing Farm), was used. This data includes maximum and minimum temperature, precipitation, wind, humidity, visibility etc., which satisfies the DNDC model requirements.

- **Soil data**

Soil parameters are very often the most sensitive input parameters in agro-ecosystem models (Li et al., 1992; Kersebaum et al., 2007). Spatial information on soils is necessary to target disaggregated agro-ecosystem modeling. A soil information system is essential for providing agro-ecosystem model input parameters. In this study, a spatial soil type map with a scale of 1:100 000 was provided by the Chinese Academy of Agricultural Sciences. This soil data provide soil types in the genetic soil classification of China's soil classification system.

Soil organic carbon (SOC) and soil pH values of the top soil layer (0 – 20 cm) for the entire Qixing Farm were provided by the Qixing Modern Agriculture Research Center. These data were measured from 1156 measured samples, which were evenly distributed over the Qixing Farm.

Soil samples of the top soil layer (0 – 20 cm) at the field experiment sites were collected and analyzed for a site-specific analysis. This soil data provide information of the soil nutrients N, P, K, SOC and the pH value.

- **Remotely sensed data**

FORMOSAT-2 (FS-2) satellite imagery was used to derive rice vegetation information. The FS-2 collects multispectral images with a ground pixel resolution of $8\text{ m} \times 8\text{ m}$ over a swath of 24 km. The FS-2 images used in this study are optical images with 4 bands of blue (450–520 nm), green (520–600 nm), red (630–690 nm), and near-infrared (760–900 nm). Three tiles of high quality images covering the main arable land area ($\sim 56,000\text{ ha}$) of the Qixing Farm were captured on 24 June, 6 July, and 9 August in 2009. Thus, both the vegetative phase (24 June and 6 July) and the reproductive phase (August 9) of rice are well represented in these images. Integrated with the GIS boundary data for the Qixing Farm, the FS-2 images were also used to delineate accurate rice areas.

- **Field observation data**

Ground truth of agronomic data were measured in field campaigns. Samples at 42 sites were collected during the entire rice growing season in 2009. All these 42 sites were located in seven farmers' fields being spatially separated. Each site was represented by one plot covering approximately 0.1 – 0.3 ha. The final plant samples collected from each site were a mixture of three or four spatially separated samples which were taken from the same plot. As ground truth data, the areas of the sample sites were mapped using a Trimble™ Global Positioning System (GPS) receiver.

Field management calendars of transplanting, N topdressing, irrigation, application of insecticides, and harvest dates, were recorded. Several field campaigns were carried out to collect samples from the tillering stage, booting stage, heading stage, 20 days after heading, and the harvest stage. For each site, biomass, LAI and plant N concentration were measured and plant N uptake was calculated as well.

The plant samples were processed in the following steps: After the field sampling, the plants were first cleaned, and then separated into different organs (leaves, stems, panicles) to measure the biomass values. LAI was measured using a sub-sample of the leaf biomass. One sub-sample consisted of 10 – 20 leaves, randomly selected among the youngest fully developed leaves. All fresh samples were processed in the oven at $105\text{ }^{\circ}\text{C}$ for half an hour to stop enzyme activity. After that, they were dried at $75\text{ }^{\circ}\text{C}$ for at least 72 h until a constant weight was reached before they were finally weighted. N concentration was measured using the Kjeldahl-N method. The plant N uptake was calculated as the aboveground dry mass multiplied by the N concentration. Grain samples were collected at harvest time. The grain sample for each site was combined of three or more sub-samples that were collected from an area of 1 m^2 . All samples were processed by grain threshing, drying, weighting.

35 sites from different farmers were used to validate the model regional application. These 35 sites are mainly clustered in three areas in the Qixing Farm. The yields of these 35 sites were recorded by the farmers when they were selling their grain after harvest. It was calculated as the total yield of the total field area including the field ridges, water pools and channels. The building area of living houses might also be calculated into the field area.

Another field dataset was collected from 22 selected farmers. These farmers were assumed to be high yield farmers regarding to their rice yield records in the past 3 to 10 years. The 22 sites were spatially distributed over the entire Farm. Grain yield in this dataset were measured using the same method as in field dataset I. This dataset was used to validate the model regional application.

- **Field management data**

Field management data of the 42 experimental sites were recorded during the growing season in 2009. Farm management data such as sowing/transplanting/fertilizing date, amount of nitrogen fertilizer input, irrigation management (especially for rice), plant protection management, harvest time, or special regime of management are provided in the farmer's survey conducted in 2009.

In the modelling regional application, one optimized field management data were applied. These optimized field dataset was processed based on the records of the 42 experimental sites and another field survey of 79 farmers' field management data in 2009. This field management data were assigned to the delineated rice areas to drive the regional model.

3 Methodology

With the development of computer science, geographical information systems, satellite RS and modelling have been widely applied for agro-environmental research over large geographic areas (Wang et al., 2015; Shao et al., 2014; McCoy et al., 2011; Wahid et al., 2008).

Earth observatory and monitoring networks currently provide unprecedented volumes of complex, multi-scale, long-term multidisciplinary explicit data (e.g., NEON, 2015; OOI, 2015; SAEON, 2015; TERN, 2015; ILTER, 2015; SIGEO, 2015). Simultaneously, data pre-processing such as georeferencing, co-registration, and categorization are mandatory especially for an agro-environmental analysis which highly depends on multi-source data in different resolutions, formats, and quality (Li, 2010; Reinartz et al., 2011; Gessner et al., 2015). A useful technology, the expert system, can be applied to categorize and analyze the multi-source research data regarding to different research issues (Liao, 2005).

Since the mid-1960s, decision support systems have been developed with communication-driven, data-driven, document-driven, knowledge-driven and model-driven, paralleled with the development of computer sciences (Power, 2008). Agricultural decision support systems receive increasing interests by scientists and stakeholders. As an example, a framework of SEAMLESS (System for Environmental and Agriculture Modelling, Linking European Science and Society) is recently developed to contribute to the sustainable agriculture by designing and assessing the integration of agricultural technologies and agro-environmental and rural development policies in Europe (van Ittersum et al., 2008). The SEAMLESS explores links between computer models in multiple dimensions including organisational levels (i.e. point or field scale, farm, region, EU and the world), temporal scales and spatial scales (Ewert et al., 2011). Based on several backbone models (such as SEAMCAP [an agricultural sector model], CAPRI [common agricultural policy regionalized impact], FSSIM [farm system simulator], etc.), the SEAMLESS provides systematic analysis to design a sustainable agriculture. In this study, the three technologies of GIS, RS and modelling are integrated as the backbone tools to implement the AEIS in the SJP.

3.1 Geographic information system

The GIS is a computer application designed to capture, store, display, communicate, transform, analyze, and archive geographic information, that is, information tied to specific locations on or near the earth's surface (Goodchild, 2009). GIS were originally used isolated in land resource management, automated cartography, and transportation. GIS were rapidly developed in civilian applications as well as by the military since the 1950s. Especially after

the first deployed civilian satellites in the early 1970s, GIS software was quickly generated, largely raster-based, that found very useful applications in agriculture (Goodchild, 2009).

GIS represents ground surface features in two distinct ways. The first way is for discrete objects. The objects on the earth are assumed to be countable features that may overlap, but between them is emptiness. The features can be represented as, i.e. points, lines and areas. The second way is for continuous objects. In this way the earth's surface is described by a series of functions based on location, the functions are used to represent elevation or temperature, or a class or name. The earth's surface is represented by raster and vector structures in a GIS. In a raster structure the set of possible locations is finite, being defined by a grid. In a vector structure, every feature is located using an appropriate number of coordinates. Areas and lines are normally represented as ordered sets of coordinates connected by straight lines of polygons and polylines, respectively (Goodchild, 2009). These raster or vector data have to be referenced to the earth's surface using a certain form of coordinate system.

A wide range of GIS software products are available, ranging from versions for hand-held devices through desktop systems to server-side GIS. In recent years, open-source GIS are growing steadily, and a number of low cost options have appeared.

GIS combine layers of data with specific locations that provide powerful and critical information that can be used for decision-making tools (Estes and Star, 1990). GIS is not only a critical tool for balancing the anthropogenic influence and the nature carrying capacity, but also a promising tool for bridging these two issues, to achieve something that is bigger than the sum of the two in the future (Dangermond, 2009). In this study, GIS is used as a basic tool to prepare, process, integrate and analyze multi-source data.

3.2 Satellite remote sensing

By the late 1950s aerial photography had been institutionalized in applications in the government and civil society as a source of cartography information (Campbell and Wynne, 2011). In 1972, the launch of Landsat 1, the first of many earth-orbiting satellites designed for observation of the earth's land areas, marked another milestone. With the development of new instruments of collecting satellite images, by the 1990s, commercial capabilities for acquiring fine-resolution satellite imagery have been available for civil applications. During the first decade of the 21st century, the rapid development of the internet enabled public access to RS imagery. Therefore software based on RS products are promoted for the use of the broader public, for instance the widely applied Google Earth.

In the past years, satellite RS as an advanced technology has been used extensively in wetland research, forest planning, water erosion assessment, agriculture production, to obtain spatial and temporal information (Ozesmi and Bauer, 2002; Holmgren and Thuresson, 1998; Vrieling, 2006; Gitelson et al., 2014). Because of its operational and economical uses over large areas, satellite RS technology has been widely used to conduct in-season crop monitoring and yield forecasting for decision making on in-time field management and marketing intervention and policy support on regional or global scales (MacDonald and Hall, 1980; Duveiller et al., 2012b).

Image classification is an important process to interpret image information by assigning objects, features or areas to classes based on their appearance in the imagery. Another method to interpret the image information is using vegetation indices (VIs). The VIs are formulas of combinations of different spectral bands, which show better sensitivity for vegetation. They are used to quantitatively measure the status of the vegetation (Bannari et al., 1995). Numerous investigations have related the VIs to several vegetation phenomena ranging from vegetation seasonal dynamics at global and continental scales to tropical forest clearance, leaf area index measurement, biomass estimation, percentage ground cover determination, and photosynthetically active radiation estimation (Lillesand et al., 2014).

Using satellite RS techniques to monitor paddy rice areas has been proposed in the past decades (Frolking et al., 2002; Xiao et al., 2005; Kuenzer and Knauer, 2013). Satellite RS data with coarse and medium resolution are widely used in rice cultivation research (Kuenzer and Knauer, 2013; Wang et al., 2010; van Niel and McVicar, 2004). However, the number of conducted studies on rice using high resolution RS images was limited in the past two decades (Kim and Yeom, 2012; Chang et al., 2013; Kim and Yeom, 2014). Identification of rice cultivation areas and estimation of agronomic parameters from high resolution images are valuable for improving rice production.

In this study FORMOSAT-2 (FS-2) images and GIS data were combined to delineate the paddy rice areas. FS-2 images are used for rice status monitoring in the growing season and rice yield estimation.

3.3 Knowledge-based systems

The computer user dictionary defines knowledge-based systems (KBS) as computer systems that are programmed to imitate human problem-solving by means of artificial intelligence and reference to a database of knowledge on a particular subject. The components of knowledge-based systems are a knowledge base, and inference/reasoning mechanisms. The knowledge base is the core part of a KBS (Curtis and Cobham, 2008). Since the 1980s,

approaches have been reported to utilize both numerical and logical information in KBS with an integrated procedure for image analysis in application of agriculture (Kontoes et al., 1993; Mckeown, 1987; Goodenough et al., 1987).

To construct a spatial rice information database in Qixing Farm, an expert knowledge base, has been designed and implemented. This knowledge base is used to integrate the multi-source data based on logical rules. It consists of a rice area mask, a RS base, a soil base, a hydrology base, a topographic base, and a farmer management base.

- **Knowledge engineer**

In this study, the knowledge base was constructed using the Knowledge Engineer™ program in the expert classifier module in ERDAS IMAGINE software. The ERDAS Expert Classifier interface is designed to handle the process that an expert in a particular field of expertise would use to analyse spatial data and infer information within a given location. This process can then be repeated by someone without expertise in either the application field or in the use of software tools (Lei, 2008). This program can collect spectral information or classes from raster images and semi-automatically transfer them into a knowledge database system. The expert classification program consists of three components: hypothesis, rule, and condition. A condition compares a pixel (input data) to a real value. A rule assigns a meaning to one or more conditions. A hypothesis forms a classification (output class) on the basis of one or more rules. An expert classification system is a hierarchy of rules, or a decision tree (**Figure 3-1**). The decision tree grows in depth when the hypothesis of one rule is referred to by a condition of another rule. The terminal hypothesis of the tree represents the final class area of the interest.

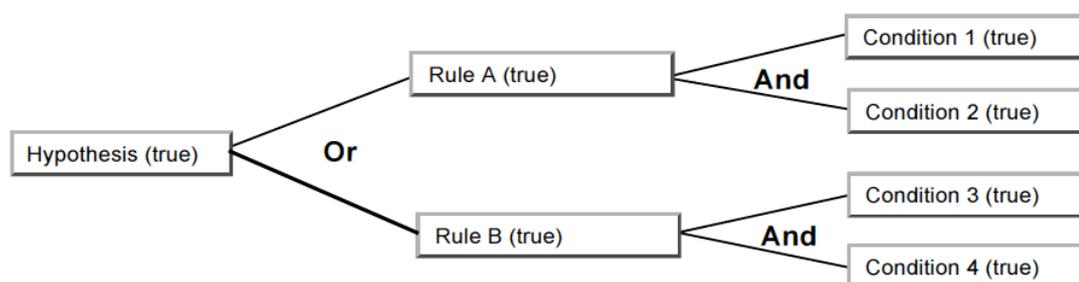


Figure 3-1: Components of a knowledge-based system.

- **Rules for each knowledge base**

A rule is a conditional statement, or list of conditional statements, about the variable's data values and/or attributes. Rules consist of a number of variables, of which the data values and/or attributes are assigned into a certain class constrained to conditional statements. All

of the conditional statements must be true in order to make the rule true. The mathematical operators in conditional statements in the rules include: =, \neq , >, \geq , <, and \leq . The confidence of the variables depend on the extent of knowledge of the variable values.

3.4 Agro-ecosystem modelling

Crop modeling is an effective means in investigating the interactions between crops and the environment due to its capability in the ensemble of varying factors (model inputs) temporally, and subsequently to simulate crop production and its environmental effects under different scenarios (Rosenzweig et al., 2014; Zhang et al., 2014). With the rapid development of computational technologies, process-based crop modelling are widely applied. These crop modelling are developed based on a series of bio-physical & -chemical equations regarding crop growth. In the 1960s, crop modelling were developed to investigate crop growth in response to abiotic environmental factors (de Wit, 1965; Duncan et al., 1967). Models are currently being used in support of theoretical research, yield predictions, and decision making in agriculture (Long et al., 2006; Lobell and Burke, 2010; Schreinemachers and Berger, 2011).

For this study, the DeNitrification-DeComposition (DNDC) model (version 9.5, developed by Li et al., 1992, modified by Li et al., 2000, 2007) was chosen in order to test the model capability for detecting within-field variability in paddy rice yield for a study area in North-east-China. The overall aim is to analyze agro-environmental patterns of spatial variability in the context of precision agriculture.

The DNDC model simulates the carbon and nitrogen biogeochemical cycles and is composed of the following six interacting sub-models: soil-climate, plant growth, decomposition, nitrification, denitrification and fermentation. In several studies, the DNDC model was applied for paddy rice fields in China (e.g., Li et al., 2002; Zou et al., 2009; Zhang et al., 2014).

3.5 Concept of an agro-environmental information system (AEIS)

Since the last Century, famers have been expected to pay attention not only to economic profits but also to environmental impacts in the agricultural system (e.g., Sigrimis et al., 1999; Bareth, 2009). The agriculture systems are complex. It is difficult, and often even impossible, to characterize the functioning of such a complex systems by means of direct measurements. Many studies based on empirical regression of the single factors to the environment, i.e., pesticide pressure (Vernier et al., 2013), or innovate some indices to evaluate the environmental effects (agro-environmental indicators) (Girardin et al., 1999). Indicators are a compromise between scientific results and the need for concise information. However,

simple indicators ignore the dynamics of the actual status, may provide considerable errors especially in terms of amplifications.

To understand the interactions between the agriculture systems and the environmental impacts, a comprehensive and dynamic policy approach covering a range of spatial and temporal scales and issues is required. A crucial component of this approach is the implementation of information systems that are relevant, robust, and easily operated by all stakeholders, practitioners, policymakers, and scientists. For instance such information system can help to adjust practices, processes and capital in response to the actuality or threat of greenhouse gas emissions or climate change (Howden et al., 2007). Ewert et al. (2011) implemented a series of model in a ‘integrated assessment and modelling’ system which integrated series of models to investigate the components in the agricultural system on different scales.

With the development of information acquisition technologies of GIS and RS, it is possible to get agricultural information on multiple spatial and temporal scales and thus to merge these multi-source data. Process-based geochemical models are effective means to link data from the environment factors to biotic crops. A prospect way to achieve the knowledge of the interactions between the agricultural systems and its environmental impacts is to combine the three means of GIS, RS, and crop modelling. And subsequently, by linking multi-source data using the crop modelling, information about the agro-environmental system on multiple spatial and temporal scales can be obtained.

In this study, the AEIS means an information system that provides information for agriculture, environment, and simulation scenarios through regional agro-ecosystem models which are driven by the SDIs, especially for a data-poor region. The core idea of the AEIS for the SJP is to explore the potential use of the available geospatial data and mainly based on the technologies of GIS, RS, and modelling. The AEIS is capable in providing temporally and spatially dynamic agro-environmental information at multiple scales, which could contribute greatly to field management and regional planning. By conducting a case study in the Qixing Farm, this research aims to implement an AEIS to investigate rice production and its environmental effects based on multi-source data.

4 Georeferencing Multi-Source Geospatial Data Using Multi-Temporal TerraSAR-X Imagery: a Case Study in Qixing Farm, Northeast China

Quanying Zhao ^{1,*}, Christoph Hütt ¹, Victoria I.S. Lenz-Wiedemann ¹, Yuxin Miao ², Fei Yuan ³, Fusuo Zhang ², Georg Bareth ¹

Published in: *Photogrammetrie - Fernerkundung - Geoinformation* 2015, 2, 173–185.

doi:10.1127/pfg/2015/0262

Original manuscript is embedded in dissertation format.

- 1 Institute of Geography, GIS & RS Group, University of Cologne, Albertus-Magnus-Platz, 50923 Cologne
- 2 College of Resources and Environmental Sciences, China Agricultural University, Beijing China, 100193 Beijing
- 3 Department of Geography, Minnesota State University, Minnesota, MN 56001, USA

* Corresponding author, Tel.: +49-221-470-1951; E-Mail: zhaquanying@gmail.com

Summary

Geodata, including optical remote sensing (RS) images and topographic vector data, can be collected from multiple sources such as surveying and mapping agencies, commercial data acquisition companies, and local research institutes. These multi-source data have been widely used in past RS and geographic information system (GIS) studies in various applications. However, spatial inconsistencies inherent in the multi-source data require accurate georeferencing to be applied. This is challenging for study sites with limited accessibility and few reference maps. To address this challenge, this paper proposes an approach for generating ground control points (GCPs) using TerraSAR-X (TSX) data. In a case study, TSX images were used to georeference multi-source data covering the Qixing Farm in Northeast China. First, a stack of five multi-temporal TSX images were processed into one reference image to retrieve GCPs. These were then used to georeference the other datasets including Huanjing (HJ), Landsat 5 (LS 5), FORMOSAT-2 (FS-2), and RapidEye (RE) satellite images, as well as topographic vector datasets. Identifying tie points in the multi-source datasets and the corresponding GCPs in the TSX reference image enables georeferencing without field measurements. Finally the georeferencing accuracies for the optical RS images were assessed by independent check points. Good results were obtained for the HJ, LS 5, FS-2 and RE

images, with an absolute error of 7.15 m, 6.97 m, 8.94 m and 10.52 m, respectively. For the topographic vector datasets, ideal visual results were achieved, attributable to the rubber sheeting algorithm. These results demonstrate that the TSX reference image is suitable for georeferencing multi-source data accurately and cost-efficiently. The developed procedure can be applied in other study regions and is especially valuable for data-poor environments.

Zusammenfassung

Georeferenzierung von Raster- und Vektordaten aus unterschiedlichen Quellen mit Hilfe von multitemporalen TerraSAR-X-Aufnahmen – eine Fallstudie der Qixing-Farm im Nordosten Chinas. Für räumliche Analysen kommen Geodaten wie Fernerkundungsdaten und topographische Vektordaten zum Einsatz, die von diversen Einrichtungen, u.a. Vermessungsämtern, kommerziellen Geoinformations-Dienstleistern und Forschungsinstituten bereitgestellt bzw. Bezogen werden. Diese aus unterschiedlichen Quellen stammenden Daten (Multidaten) werden für zahlreiche Anwendungen in Fernerkundungs- und GIS-Studien genutzt. Jedoch beinhalten diese Daten räumliche Ungenauigkeiten, die zunächst eine präzise Georeferenzierung erforderlich machen. Dieses stellt vor allem für Untersuchungsgebiete mit eingeschränkter Zugänglichkeit und nicht verfügbaren Referenzdaten eine Herausforderung dar. Dieser Artikel erklärt, wie Passpunkte aus Daten des Radarsatelliten TerraSAR-X (TSX) für die Georeferenzierung von Multidaten generiert werden können. In einer Fallstudie der Qixing-Farm im Nordosten Chinas wurden fünf multitemporale TSX-Radarbilder zu einem Referenzbild zusammengefügt, um mit hoher Genauigkeit Passpunkte

abzuleiten. Diese Passpunkte dienen der Georeferenzierung mehrerer Multidaten aus diversen Quellen, welche sowohl Huanjing (HJ)-, Landsat 5 (LS 5)-, FORMOSAT-2 (FS-2), und RapidEye (RE)-Satellitenbilder als auch topographische Vektordaten umfassen. Die Identifizierung derselben Passpunkte in dem TSX-Referenzbild und in den Multidaten diverser Quellen ermöglicht eine genaue Georeferenzierung ohne im Gelände aufgenommene Messdaten. Die Genauigkeit der Georeferenzierung für die optischen Satellitenbilder wurde durch unabhängige Kontrollpunkte bewertet. Es wurden gute Ergebnisse für die HJ-, LS 5-, FS-2 und RE-Satellitenbilder mit absoluten Fehlern von 7,15 m, 6,97 m, 8,94 m bzw. 10,52 m erzielt. Für die Georeferenzierung der topographischen Vektordaten wurden optimale visuelle Resultate erzielt, welches dem eingesetzten „Rubber Sheeting Algorithm“ zuzuschreiben ist. Diese Ergebnisse demonstrieren die Eignung der aus TSX-Daten abgeleiteten Passpunkte, um Multidaten verschiedener Quellen genau und kosteneffizient zu georeferenzieren. Das entwickelte Verfahren kann auf andere Untersuchungsregionen übertragen werden und ist besonders wertvoll für Gegenden mit schlechter Verfügbarkeit von Referenzdaten.

Keywords: georeferencing; spatial inconsistency; multi-source data; TerraSAR-X;

Topographic vector data; optical remote sensing imagery

4.1 Introduction

Data quality plays a critical role in geodata related research (Bareth, 2009). To ensure data quality, georeferencing becomes a mandatory and crucial task. In this paper, datasets from different sources, each characterized by their unique attributes and properties, are referred to as multi-source data. Compared to single-source data, multi-source data can provide adequate information with different spatial and temporal resolutions, map scales, and spectral properties (Li, 2010; Waldhoff et al., 2012). Multi-source data provided by various governmental bureaus or non-governmental organizations such as local research institutions or special research groups may vary in many interpretation aspects and in terms of (spatial) data quality. Both Geographic Information System (GIS) and remote sensing (RS) data carry plenty of geospatial information but with different nature and content and with different semantics (Weis et al., 2005). The integration of RS and GIS is emerging as a new research field (Zhang, 2010). Gómez-Candón et al. (2012) indicated that the locational errors in high resolution images, e.g. GeoEye-1 images, affect the delineation of the input prescription map which is a core problem for the implementation of site-specific agricultural management strategies. Weber et al. (2008) confirmed that coregistration errors between imagery and field sites led to remarkable errors in landscape classification, particularly when the size of the target site was similar to the image pixel size. Moreover, in some cases, such as in China, detailed topographic data (1:5,000 – 1:25,000) with high spatial accuracy may not be accessible due to data sharing and management policies or lack of surveying and mapping activities (Bareth and Yu, 2004). Because of heterogeneous qualities, the integration and georeferencing processes for multi-source data are indispensable, complex and highly dependent on the purpose of the study.

A variety of methods for multi-source data integration and georeferencing have been developed in the past decades to eliminate spatial inconsistencies in multi-source datasets. For example, a Markov random field model was applied to merge images from multiple sensors for a land use classification (Solberg et al., 1996). A statistical approach to match relational features was introduced by Walter and Fritsch (1999). An iterative closest point algorithm was implemented to match features using a spatially precise map as the reference (von Gosseln and Sester, 2004). Empirical and theoretical methods were implemented by Usery et al. (2009) for integrating the national maps of the United States with different scales and

resolutions in vector and raster datasets. In addition, several automatic approaches have been developed to compute the imagery-to-vector conversion (Wu et al., 2007), identify control point pairs from images using vector datasets as the glue layers (Chen et al., 2006), conflate vector maps to high resolution imagery (Song et al., 2009), or georeference image sequences in real-time (Choi and Lee, 2012).

In recent studies, Synthetic Aperture Radar (SAR) imagery has been used to quantify the spatial in-consistencies of geodata and to collect ground control points (GCPs) for georeferencing. SAR sensors are all-weather and day-night active microwave sensors that collect information of the targets according to the signal transport time between the sensor position and the terrain height. They have the potential to provide images with very high geometric accuracy (Ager and Bresnahan, 2009; Rodríguez et al., 2006). In particular, the German TerraSAR-X (TSX) satellite launched in 2007 is equipped with a highly flexible phased array antenna for SAR Stripmap, ScanSAR, and Spotlight operations (Mittermayer and Runge, 2003). An overall ground accuracy of ≤ 1 m has been demonstrated when the images are projected to a precise terrain height (Ager and Bresnahan, 2009; Koppe et al., 2010; Nonaka et al., 2008). Therefore, the TSX products can be used to generate topographic maps and create accurate orthoimagery products (Badurska, 2011; Reinartz et al., 2011; Schneider et al., 2009).

To further explore the potential capability of TSX imagery as a source for locating GCPs and subsequently to georeference multi-source data characterized by varying properties and accuracies over a large area, a feasible and robust method which takes the advantage of the high spatial resolution and high geometric accuracy of TSX imagery is introduced. The main specific objectives are (i) to georeference topographic vector data from multiple sources; (ii) to improve the georeferencing results of Huanjing (HJ), Landsat 5 (LS 5), FORMOSAT-2 (FS-2), and RapidEye (RE) satellite images; and (iii) to assess the accuracy of georeferenced datasets and to evaluate if the results are highly dependent on the spatial accuracy of the TSX imagery.

4.2 Study area and data

4.2.1 Study area

The Sanjiang Plain (SJP), located in Northeast China, is an alluvial plain formed by the Songhua River, the Heilong River and the Wusuli River. The topography is fairly flat with a slope of $< 0.012^\circ$. With an area of approximately 11 million ha, it is an important wetland area and ecosystem in China. Some wetland sites in this area have been designated for the

list of wetlands of international importance (Wang et al., 2006). In addition, the SJP is the largest food base of China, where 52 national-owned farms are located. The climate is temperate sub-humid, with a mean annual precipitation of 500 mm – 600 mm (80 % of it occurring between May and September), and an average temperature of 21 °C – 22 °C in July and ~ -18 °C in January. Nowadays, single season crops of paddy rice, soybean and maize are mainly planted in this area.

The study site Qixing Farm (47.2 °N, 132.8 °E), which covers an area of approximately 120,000 ha, is located in the central part of the SJP (Figure 4-1). As of 2010, 62 % of the study site was arable, dominated by three quarters of paddy rice and one quarter of dryland (Zhang, W., Qixing Farm, personal communication, June 2012). In the paddy rice fields, rainfed and irrigation systems simultaneously exist. To improve the growing conditions of agricultural crops, shelter forests were planted in the late 1980s, primarily to reduce the speed of ground wind (Liu and Zhao, 1996).

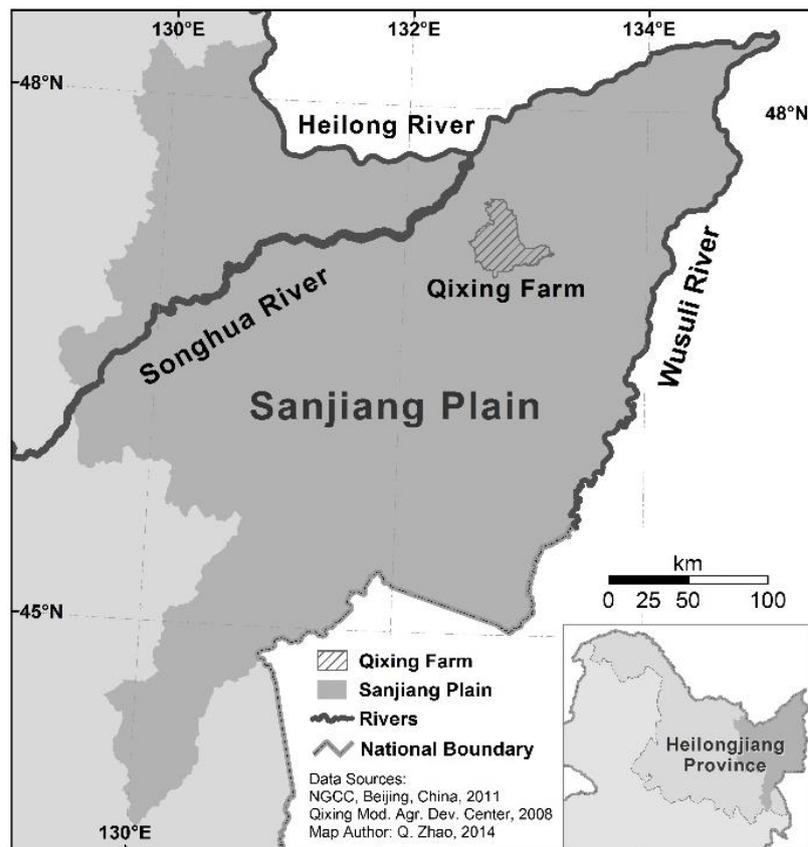


Figure 4-1: Location of the study area Qixing Farm in Northeast China.

4.2.2 Data description

A time series of five TSX images (stripmap, VV-polarisation, incidence angle $\sim 35^\circ$, relative orbit 88, descending) was taken within 44 days from June 24 to August 7 of 2009 (see **Table 4-1**). These five stripmap images in the basic Single Look Slant Range Complex (SSC) form with intensity and phase information for each pixel in slant range geometry were used to create a TSX reference image. The orbit precision was set to ‘science’, which means that the satellite position during image acquisition is calculated with an error of ≤ 20 cm in a post processing step (Fritz and Eineder, 2013). This post processing dramatically increases the positional accuracy and thereby the image potential for generating GCPs (Koppe et al., 2010).

The Qixing Farm field boundary file was produced by the Qixing Modern Agriculture Development Center. This GIS layer was given in Universal Transverse Mercator (UTM) coordinate reference system, zone 53 N. It provides the information on crop field boundaries, irrigation wells, water drainages, and shelter forests edges at a fine field unit scale. However, this dataset did not line up with any of the other datasets in our project. The inconsistency was nonsystematic in distance or directions (**Figure 4-3**). An offset of more than 200 m between this dataset and the TSX images was identified in the northwest part, whereas in the southeast part the shift was more than 300 m in the opposite direction.

The public version of the 1:250,000 topographic vector dataset was produced by the National Geomatics Center of China (NGCC). This dataset includes multiple layers of administration boundaries, settlements, railways, roads, hydrological information, and landscapes. However, as Bareth and Yu (2004) indicated, the spatial accuracy is not as high as expected. Therefore, a refined georeferencing of the public version is needed in this study.

The HJ, LS 5, FS-2, and RE satellite images were acquired from 2009 to 2012 in the growing season. The agricultural constructions, e.g. irrigation channels and raised ridges, for paddy rice in the study area are the same year by year and the field boundaries are mostly stable. Therefore, one TSX reference image can be used in multiple years. The detailed information of the RS data is listed in **Table 4-1**.

Table 4-1: Characteristics of the RS images.

Satellite	Pixel Spacing (m)	Bands	Acquisition date	Projection	Cloud Cover (%)	Processing Level
TSX	1.89 (az) × 1.57 (rg)	-	June 24, July 5, 16, 27, Aug. 7, 2009	WGS 84 UTM 53 N	--	SSC
HJ(CCD2)	30 × 30	4	June 29, 2012	WGS 84 UTM 53 N	0 (subset)	2
LS 5	30 × 30	7	Aug. 26, 2011	WGS 84 UTM 52 N	0 (subset)	1T
FS-2	2 × 2 (PAN)	5	July 6, 2009	Geographic (Lat/Lon)	0	1A
RE	5 × 5	5	May 19, 2012	WGS 84 UTM 53 N	0	3A

4.3 Methods

4.3.1 Workflow of georeferencing multi-source datasets

The schematic workflow of multi-source data georeferencing is shown in **Figure 4-2**. There are mainly four steps involved: (1) pre-processing of the multi-temporal TSX images to generate one single reference image; (2) selection of GCPs from the processed TSX imagery and corresponding tie points from optical RS images or topographic vector maps; (3) reducing locational errors by recursively reselecting GCPs and corresponding tie points until achieving low positional error (PE) values or satisfactory visual results; (4) generating georeferenced datasets by image resampling or GIS data matching. Steps 1 and 2 are the key steps of this approach, which highly affect the quality of the GCP interpretation and consequently the final results.

We decided to use the PE because it is implemented in the software that we used and because the documentation of the individual error of every point proves best the efficiency of the proposed method. The standard deviation (Std.) that characterizes the overall error is also given as a comparison.

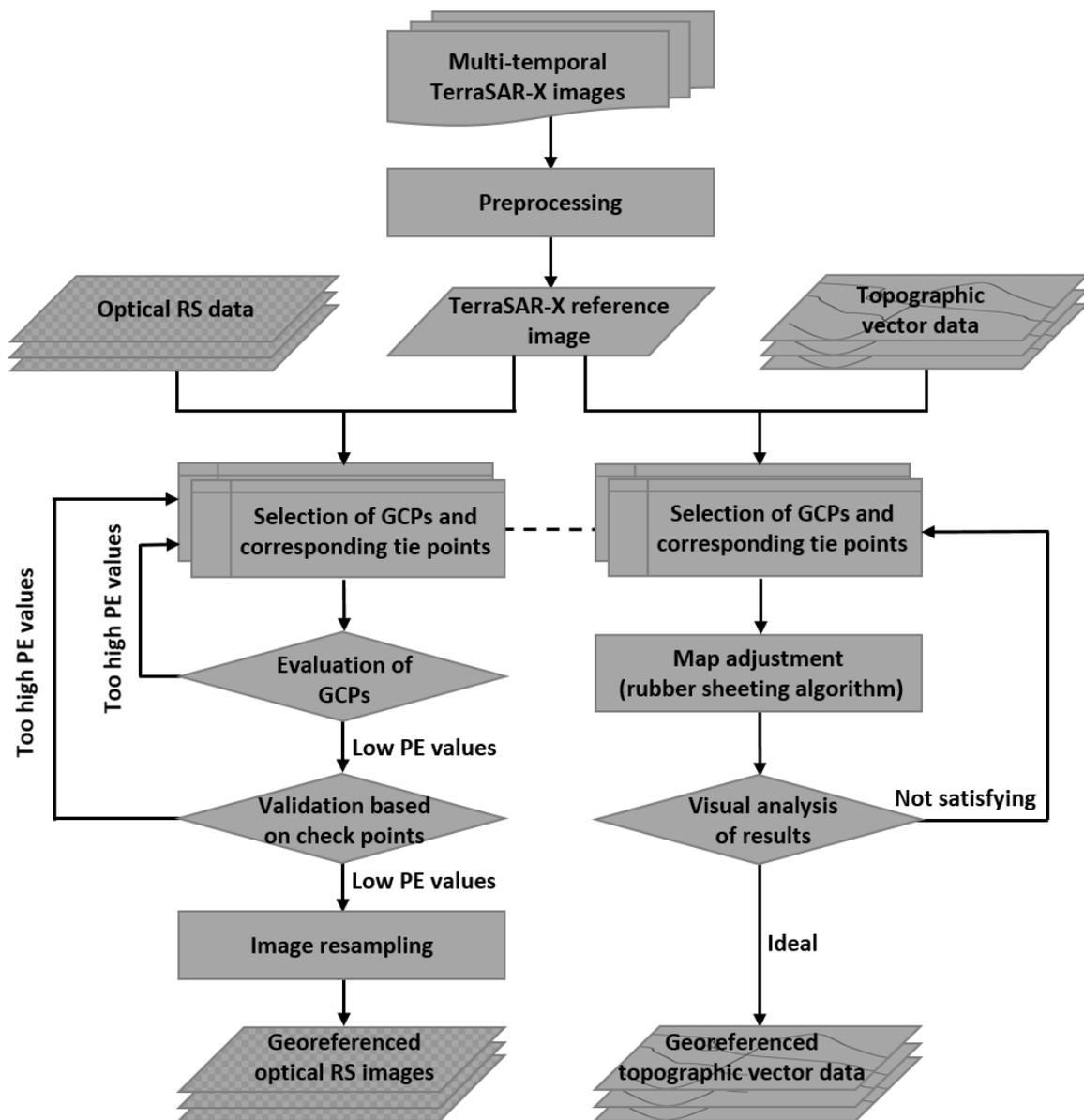


Figure 4-2: Georeferencing workflow of the multi-source geospatial data, PE = positional error.

4.3.2 Creation of the reference image from TSX stripmap acquisitions

A stack of five TSX stripmap images was used to create the reference image. Radar image processing was performed using the Next ESA SAR Toolbox (NEST) distributed under the GNU General Public License. To meet the requirement of a geocoded image in which the precise outlines of objects are identifiable, certain pre-processing techniques were applied. First the “complex pixel values” were used to calculate an amplitude image representing the strength of radar backscatter for each radar pixel cell. During the following “range Doppler

terrain correction”, the elevation data from the Space Shuttle Topography Mission (SRTM) in a spatial resolution of 3 arc-seconds served to transform the radar images from slant range geometry into the UTM coordinate reference system. Pixel spacing of the resultant geocoded product was set to 2 m to minimize spatial information loss, and to meet the file requirements of a manageable product. The main drawback of the SAR image with regard to the visible interpretation is the speckle effect which is an inherent noise of all radar images, often called grainy salt and pepper noise. To reduce this effect, a mean image of the five geocoded images was calculated and a 3x3 mean speckle filter applied. The radiometric resolution was reduced from 16 bits to 8 bits. Therefore, the data size was considerably reduced. Likewise, the image representation speed was dramatically increased. In spite of a radiometric information loss during this procedure, the processed TSX reference image provides sufficient information for human interpreters to clearly define unambiguous GCPs with a high spatial resolution. Absolute radiometric calibration was not needed in this process as all five images have the same calibration constants, and moreover, the quantitative analysis of the backscattered signal was not the focus of this study. The resultant grayscale radar image was almost speckle-free and the shapes of all objects necessary in this research could be identified.

4.3.3 Georeferencing of topographic vector data

Georeferencing of the topographic vector data was based on a rubber sheeting algorithm. The rubber sheeting, alternatively called rubber sheet, algorithm is one of the earliest and the most common computer cartogram algorithms (Tobler, 2004). This technique derives its name from the logical analogy of stretching a piece of rubber to fit over some objects (Cobb et al., 1998). During the process, map areas are subdivided into triangular-shaped regions and local adjustments are applied on each single region. After that, each triangle either enlarges or shrinks iteratively toward its ideal size without changing the topology of the map (Gillman, 1985; Dougenik et al., 1985). An iterative math-physical cartogram algorithm for continuous area was proposed by Dougenik et al. in 1985. This algorithm was recently improved by implementing an auxiliary quadtree structure in the process (Sun, 2013a; 2013b).

In this study, the rubber sheeting tool of ArcGIS 10.1 was used to transfer the topographic vector data. Approximately 600 reference points, evenly distributed over the entire area of Qixing Farm, were selected as georeferencing points from the TSX reference image. As Rinartz et al. (2009) proposed, the selection of reference points from the TSX image is not always a straightforward procedure. Based on our experience, corresponding points were selected according to following rules: (1) Select points in the TSX reference image that are

located at the intersection of the paddy field ridges, rural road edges, canopy crossings of different crops, or corners of artificial waters, which are in all cases clearly identifiable and unchanged during the 3 year-period from 2009 to 2012. (2) Avoid elevated objects such as forest edges or tall buildings due to their systematic locational errors such as foreshortening, layover, and shadowing, induced by the radar imagery acquisition procedure. (3) Select only points that have a corresponding (tie) point in the vector dataset, e.g. the Qixing Farm boundary data with line intersections and corners. A similar process was applied to the topographic GIS data provided by the NGCC.

4.3.4 Georeferencing of optical RS data

Multiple optical RS aforementioned data were also selected to demonstrate the georeferencing process based on the TSX reference image. In particular, image subsets covering the Qixing Farm were created for the HJ and LS 5 satellite data. All optical satellite images were georeferenced according to these main steps: First, all satellite images were reprojected into the UTM WGS 84 system to obtain an over-view of the data inconsistencies. Second, a set of control points was selected from the TSX reference image based on the aforementioned rules. Consequently, the corresponding points have to match the objects which can be clearly identified in the optical RS imagery in this case. Third, in order to improve the transformation model and to minimize the errors caused by the manual measurement, GCPs and corresponding tie points were updated iteratively by eliminating the points with highest PEs and selecting additional control points until the residual errors fell below the maximum allowed value. The decision if a PE value was too high depended on the spatial resolution of the image to be georeferenced. For every single GCP, the maximum allowed value was within the subpixel range. Finally, a certain number of independent points were defined as check points to evaluate the accuracy of the transformation. During the validation process, the GCPs were used to calculate the transformation model while the check points were used to evaluate the errors in the geometric transformation independently.

In our case, the PE is the horizontal distance between the input location of a GCP and the transformed location of the same GCP. The PE was calculated according to (4 – 1) (Congalton and Green, 2008).

$$PE = \sqrt{\Delta X^2 + \Delta Y^2} \quad (4 - 1)$$

Where ΔX and ΔY are the positional differences between the reference point and the corresponding image or map position in the X and Y directions, respectively.

4.4 Results

4.4.1 Georeferencing results of topographic vector data

After the georeferencing based on the rubber sheeting algorithm, the georeferenced vector data of the Qixing Farm field boundaries (cyan) sufficiently fit to the new field boundaries which are clearly detect-able in the TSX image. The problem of nonsystematic spatial inconsistency was well overcome and the shape of the vector graphics was preserved (**Figure 4-3**). Similar results were also obtained for the topographic data provided by the NGCC.

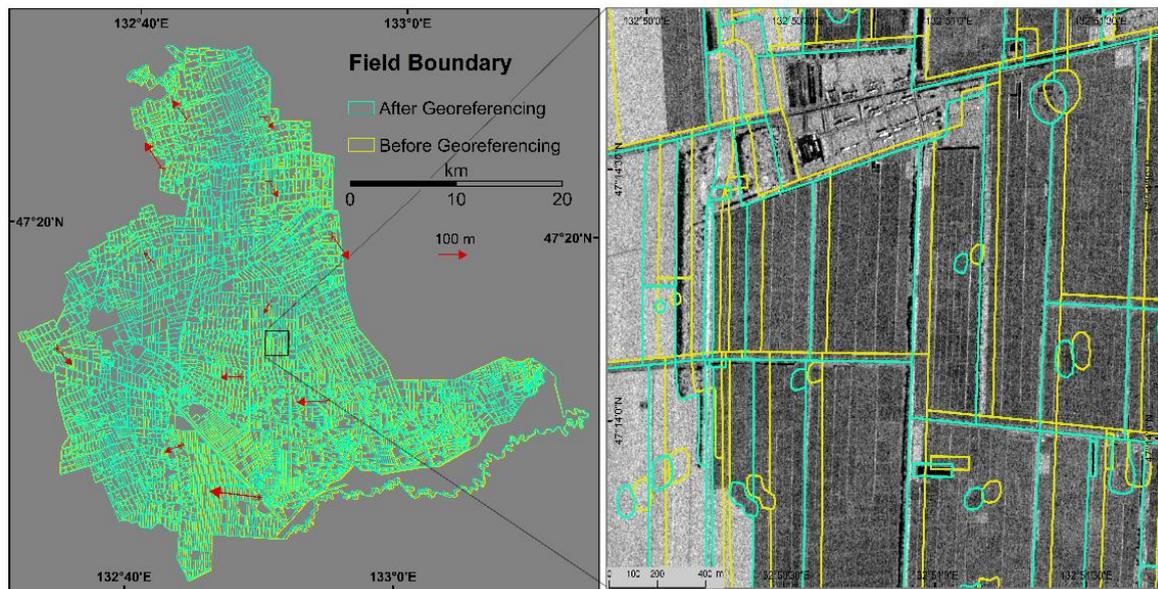


Figure 4-3: Field boundary data, before (yellow) and after (cyan) the georeferencing; red arrows in the left figure show the vector force of the rubber sheeting procedure. Background data in the right figure: TSX reference image.

4.4.2 Georeferencing results of optical RS data

Optical RS data were georeferenced according to the method described in the previous sections. **Table 4-2** shows the relevant information of the selected GCPs.

Table 4-2: Accuracy of the selected GCPs (PE = positional error, Std. = standard deviation).

Satellite	Pixel Size (resampled) (m)	Imagery/Subset Spatial extension (km)	Number of Control Points	PE (average) (m)	PE (max) (m)	PE (min) (m)	Std. (m)
HJ (CCD2)	30 × 30	55 × 55	100	12.66	27.39	1.87	6.70
LS 5	30 × 30	48 × 68	220	9.04	16.63	0.59	3.85
FS-2	2 × 2 (PAN)	30 × 28	143	3.43	5.91	0.3	1.35
RE	5 × 5	24 × 24	64	4.09	9.36	0.60	2.12

After recursively selecting control points, the final PEs were less than half a pixel for both the HJ (CCD2) and LS 5 images, and nearly one pixel for the FS-2 and RE images. **Figure 4-4** shows the georeferencing results visually. The ground features from each of the images fit well. The roads match properly in all images and the paddy field block boundaries are ideally aligned to each other in the higher resolution images (TSX, FS-2 and RE).

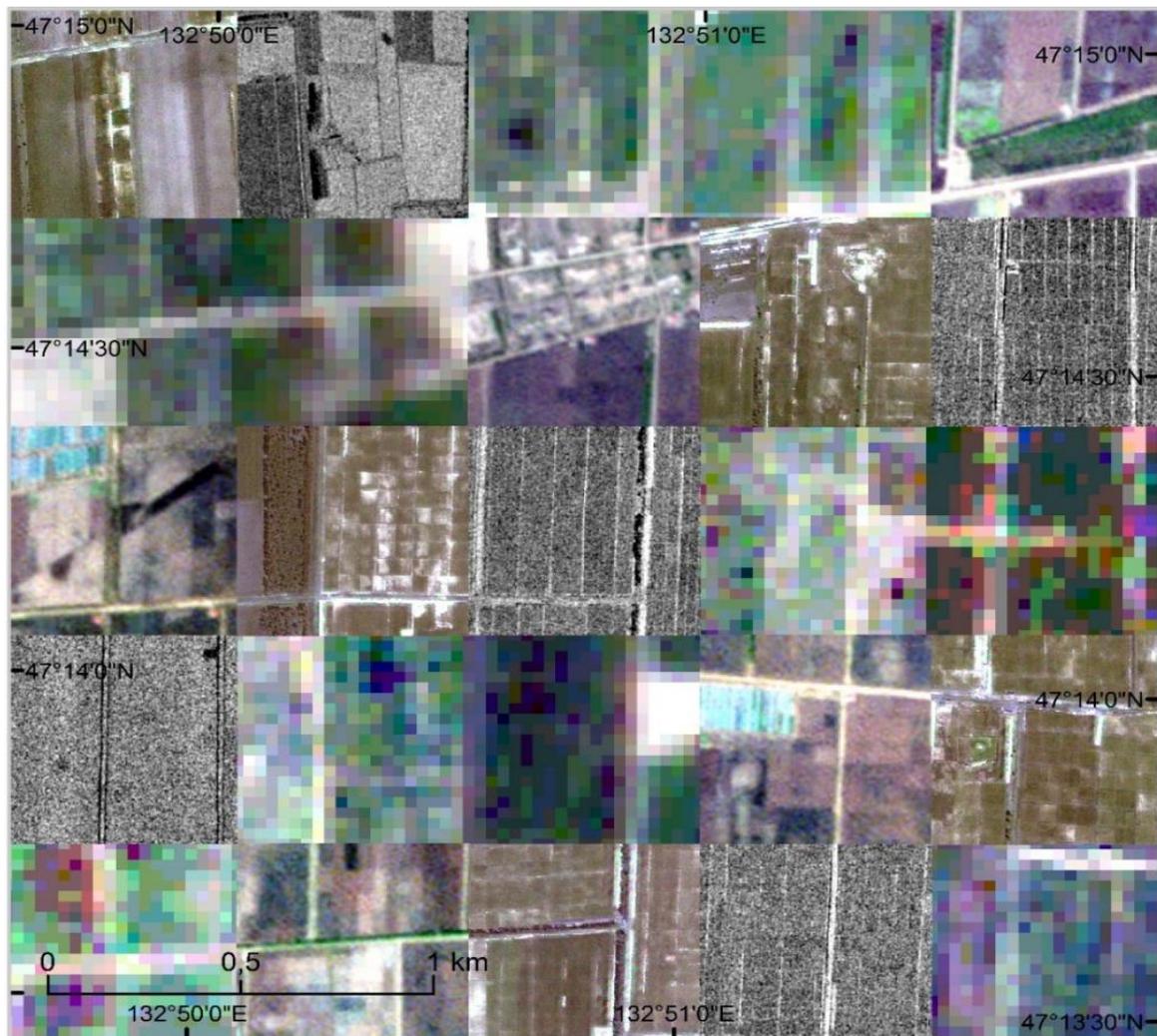


Figure 4-4: An example of georeferenced multi-source RS images in comparison to the TSX image. From left to right, 1st row: FS-2, TSX, LS 5, HJ, RE, 2nd row: LS 5, HJ, RE, FS-2, TSX, 3rd row: RE, FS-2, TSX, LS 5, HJ, 4th row: TSX, LS 5, HJ, RE, FS-2, 5th row: HJ, RE, FS-2, TSX, LS 5.

4.4.3 Spatial accuracies of the georeferenced optical RS data

To evaluate the spatial accuracies of the georeferenced optical RS data, independent check points covering the whole scene were created and their spatial parameters were analyzed. To capture the maximum PE results, the check points were located in the areas where the GCP

density was relatively low. The results were summarized in **Table 4-3**. The average PEs of the check points were at a sub-pixel value (slightly more than 0.1 pixel) in the HJ (CCD2) and LS 5 images. Accuracies of 2.5 pixels and 1.3 pixels were achieved for the FS-2 and RE imagery, respectively. The average PE values for all four types of satellite images ranged from 3.11 m to 6.66 m.

Table 4-3: Accuracy of the independent check points.

Satellite	Pixel Size (resampled) (m)	Imagery/ Sub-set Spatial extension (km)	Number of Check Points	PE (average) (m)	PE (max.) (m)	PE (min.) (m)	Std. (m)
HJ (CCD2)	30 × 30	55 × 55	20	3.29	8.05	1.81	1.55
LS 5	30 × 30	48 × 68	34	3.11	6.48	1.80	1.11
FS-2	2 × 2	30 × 28	30	5.08	7.44	1.07	1.89
RE	5 × 5	24 × 24	10	6.66	8.42	4.08	1.21

4.5 Discussion

4.5.1 Analysis of the anticipated spatial error in the processed TSX reference image

The geometric distortion of SAR imagery products can be caused by three components (Curlander and McDonough, 1991): (1) sensor/platform instability and signal propagation effects, (2) terrain height, and (3) processor induced errors. The uncertainties embedded in the SSC products comprise only the first type of error, which is less than 1 m (Nonaka et al., 2008; Fritz and Eineder, 2013). The second type of errors comes from the SRTM DEM dataset. Rodríguez et al. (2006) found that the absolute height error of the SRTM in Eurasia was less than 6.2 m; whereas in the SJP study site, where the topography is fairly flat, the absolute error was less than 2 m according to the SRTM THED (terrain height error data) product.

Hence, the target range location error (ΔR) determined by the terrain height estimation can be calculated using (4 – 2) (Curlander and McDonough, 1991):

$$\Delta R = \Delta h / \tan \eta \quad (4 - 2)$$

Where Δh is the height (DEM elevation) estimation error (2 m) and η is the location incidence angle (35° in this study). Therefore, the ΔR for this study was calculated as 2.86 m. The processor induced error is process dependent and is denoted as $\Delta \delta_i$. The overall absolute

spatial error of the projected TSX imagery can therefore be calculated by these three components, with the result of $(3.86 + \Delta\delta_i)$ m. Where the processor induced error $\Delta\delta_i$ during TSX image processing can be assumed to be infinitely small.

4.5.2 Quantified spatial accuracy of the georeferenced datasets

Considering all spatial inconsistent sources, the overall absolute error of the georeferenced datasets can be estimated. The overall errors of the georeferenced optical RS data, which are equal to the sum of the PE in **Table 4-3** and the geometric distortion of the TSX image (3.86 m), were 7.15 m, 6.97 m, 8.94 m, and 10.52 m for HJ, LS 5, FS-2, and RE satellite images, respectively. Dai and Khorram (1998) found that a registration error of less than one-fifth of a pixel should be achieved to detect 90 % of the true changes. Hence, the registration results for the HJ (CCD2) and LS 5 images can support a change detection analysis with a spatial error close to 10 %.

In surface area estimation, Ozdogan and Woodcock (2006) noted that spatial errors are dependent on both RS image resolution and the field size because of ‘the distribution of sub-pixel proportions’, especially when the field size is similar to or less than the RS data resolution. In this study, the results of the FS-2 and RE images processing are sufficient for field-unit level analysis since the size of each field block unit is typically larger than 5,000 m². The field block is the smallest area of a farm management unit and is considered as the primary scale for management decisions. The high accuracies for the HJ and LS 5 image processing are also beneficial for studies at the farmer-unit level, as a farmer’s crop field is generally larger than 20 ha. **Figure 4-5** provides a visual result of datasets from multiple sources over the entire area of the Qixing Farm.

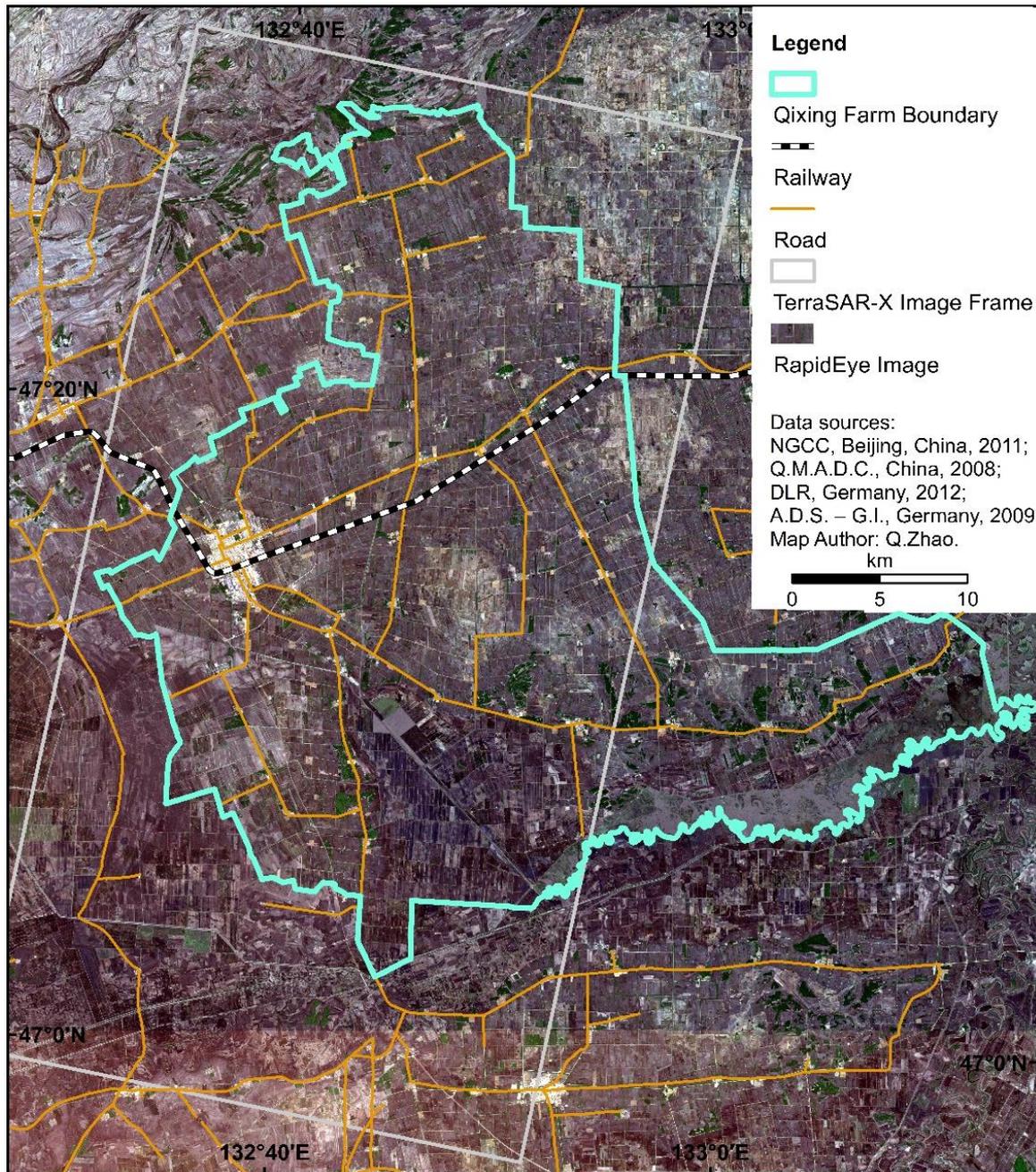


Figure 4-5: Georeferenced multi-source data for the study area of Qixing Farm.

4.5.3 Feasibility of the approach

In this study, topographic vector datasets and optical RS images from multiple sources were georeferenced using GCPs derived from the TSX reference image without the need for labour intensive field work. The creation of the TSX reference image and its use to locate

accurate GCPs is critical, because it not only determines the precision of the results but also the feasibility of this method.

Although many studies (Sowmya and Trinder, 2000; Sohn and Dowman, 2007; Reinartz et al., 2009) have attempted to extract geometric features, e.g. GCPs, automatically from satellite images, there is a lack of reports on automatic methods for georeferencing multi-source data. Automatic feature extraction methods have limited applicability due to their complex parameterization and strict condition requirements (Cobb et al., 1998). Moreover, automatic methods for integrating GIS data and satellite imagery are rare.

The strategy proposed in this study showed that for each dataset, different GCPs were required due to the diverse characteristics of the multi-source data. Manual procedure meets this requirement and ensures the spatial accuracy. Although the high resolution TSX imagery supplies a sufficient number of GCPs, the selection of the GCPs and their corresponding tie points is never straightforward. There is still a need to establish the criteria for selecting reference points systematically. Another drawback of this method is its inefficiency in processing a large number of datasets. However, the proposed method still is especially valuable for data-poor environments lacking reference data.

4.6 Conclusions

This study provides an applicable and cost-effective approach for georeferencing multi-source data with different characteristics and non-systematic spatial inconsistencies. It is an especially beneficial technique for large study sites with limited accessibility and reference maps. The results demonstrated the feasibility of using TSX imagery to accurately georeference multi-source datasets without in-situ GCP data collection. By using the mean of five TSX images and the mean filter, a speckle-free reference image was generated. This proved to be critical for locating sufficient GCPs successfully. The PEs of the check points were less than 0.2 pixel for the 30 m resolution images (HJ and LS), approximately 2.5 pixels for the FS-2 images, and 1.3 pixels for the RE images. The overall positional errors were nearly less than 10 m for all four types of images. The discrepancies among each pair of the TSX and GIS data were only assessed visually, which demonstrates a need for further study.

Acknowledgements

This research was funded by the Natural Science Foundation of China (NSFC, project No. 31071859), the International Bureau of the German Federal Ministry of Education and Re-

search (BMBF, project No. 01DO12013), and the German Research Foundation (DFG, project No. BA 2062/8-1). The TerraSAR-X images were provided by the Airbus Defence and Space – Geo-Intelligence. The GIS data layer of the Qixing Farm field boundary was provided by the Qixing Modern Agriculture Development Center in Heilongjiang, China. The 1:250,000 scale topographic dataset with multiple layers for the study area was provided by the National Geomatics Center of China (NGCC) (<http://ngcc.sbsm.gov.cn/>). The Huanjing satellite images were acquired from the China Centre for Resources Satellite Data and Application (CCRS DA) (<http://www.cresda.com/n16/index.html>). The Landsat 5 TM images were downloaded from the website of the Earth Resources Observation and Science (EROS) Center of U.S. Geology Survey (USGS) (<http://eros.usgs.gov/>). The FORMOSAT-2 images were obtained from the National Space Organization (NSPO) of Taiwan. The RapidEye satellite images were kindly provided by the RE Science Archive (RESA), German Aerospace Center (Deutsch-es Zentrum für Luft- und Raumfahrt, DLR). The leading author would also like to acknowledge the support of the China Scholarship Council (Nov. 2010 – Oct. 2014), Beijing, China.

References

- Ager, T. P. and Bresnahan, P. C., 2009. Geometric precision in space radar imaging: results from TerraSAR-X. – ASPRS Annual Conference, Baltimore, MD, USA.
- Badurska, M., 2011. Orthorectification and Geometric Verification of High Resolution TerraSAR-X Images. *Geomatics and Environmental Engineering*, 5 (3), 13–25.
- Bareth, G., Yu, Z., 2004. Verfügbarkeit von digitalen Geodaten in China. *Petermanns Geographische Mitteilungen*, 148 (5), 78–85 (written in German with English abstract).
- Bareth, G., 2009. GIS- and RS-based spatial decision support: structure of a spatial environmental information system (SEIS). *International Journal of Digital Earth*, 2 (2), 134–154.
- Chen, C. C., Knoblock, C. A., Shahabi, C., 2006. Automatically conflating road vector data with orthoimagery. *GeoInformatica*, 10 (4), 495–530.
- Choi, K., Lee, I., 2012. A sequential aerial triangulation algorithm for real-time georeferencing of image sequences acquired by an airborne multi-sensor system. *Remote Sensing*, 5 (1), 57–82.
- Cobb, M. A., Chung, M. J., Foley, H., Petry, F. E., Shaw, K.B., 1998. A rule-based approach for the conflation of attributed vector data. *GeoInformatica*, 2 (1), 7–35.
- Curlander, J. C., McDonough, R. N., 1991. *Synthetic aperture radar: systems and signal processing*. 1st edition, 672 p., Wiley, New York, NY, USA.
- Dai, X., Khorram, S., 1998. The effects of image misregistration on the accuracy of remotely sensed change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 36 (5), 1566–1577.
- Dougenik, J. A., Chrisman, N. R., Niemeyer, D. R., 1985. An algorithm to construct continuous area cartograms. *The Professional Geographer*, 37, 75–81.
- Congalton, R. G., Green, K., 2008. *Assessing the accuracy of remotely sensed data: principles and practices*. 183p., CRC press, Boca Raton, FL, USA.
- Fritz, T., Eineder, M., 2013 (eds.). *TerraSAR-X Ground Segment. Basic Product Specification Document*. Issue 1.9, Doc., TX-GS-DD-3302, DLR. https://tandemx-science.dlr.de/pdfs/TX-GS-DD-3302_Basic-Products-Specification-Documents_V1.9.pdf (18.4.2014).

- Gillman, D. W., 1985. Triangulations for rubber-sheeting. 7th International Symposium on Auto-Carto, 191–199, Washington D.C., USA.
- Gómez-Candón, D., López-Granados, F., Caballero-Novella, J. J., PeñaBarragán, J. M., García-Torres, L., 2012. Understanding the errors in input prescription maps based on high spatial resolution remote sensing images. *Precision Agriculture*, 13 (5), 581–593
- Koppe, W., Kiefl, N., Hennig, S., Janoth, J., 2010. Validation of Pixel Location Accuracy of Orthorectified TerraSAR-X Products. 8th European Conference on Synthetic Aperture Radar (EUSAR), Aachen.
- Li, D., 2010. Remotely sensed images and GIS data fusion for automatic change detection. *International Journal of Image and Data Fusion*, 1 (1), 99–108.
- Liu, T., Zhao, Y., 1996. Ecological benefit analysis and estimation of the forest shelter in the Nongken area, Sanjiang Plain. – *Forest Science and Technology*, 5, 12–14 (written in Chinese).
- Mittermayer, J., Runge, H., 2003. Conceptual studies for exploiting the TerraSAR-X dual receive antenna. IGARSS, 2140–2142, Toulouse, France.
- Nonaka, T., Ishizuka, Y., Yamane, N., Shibayama, T., Takagishi, S., Sasagawa, T., 2008. Evaluation of the geometric accuracy of TerraSAR-X. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVII, Beijing, China.
- Ozdogan, M., Woodcock, C. E., 2006. Resolution dependent errors in remote sensing of cultivated areas. *Remote Sensing of Environment*, 103 (2), 203–217.
- Reinartz, P., Müller, R., Suri, S., Schneider, M., Schwind, P., Bamler, R., 2009. Using geometric accuracy of TerraSAR-X data for improvement of direct sensor orientation and ortho-rectification of optical satellite data. *IEEE International Geoscience & Remote Sensing, IGARSS*, 5, V-44–V-47, Cape Town, South Africa.
- Reinartz, P., Müller, R., Schwind, P., Suri, S., Bamler, R., 2011. Orthorectification of VHR optical satellite data exploiting the geometric accuracy of TerraSAR-X data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66 (1), 124–132.
- Rodríguez, E., Morris, C. S., Belz, J. E., 2006. A Global Assessment of the SRTM Performance. *Photogrammetric Engineering and Remote Sensing*, 72 (3), 249–260.

- Schneider, M., Müller, R., Krauss, T., Reinartz, P., Hörsch, B., Schmuck, S., 2009. Urban Atlas—DLR Processing Chain for Orthorectification of Prism and AVNIR-2 Images and TerraSAR-X as possible GCP Source. 3rd ALOS PI Symposium, 1–6, Kona, HI, USA. http://www.asf.alaska.edu/content/pi_symp/Schneider_M._urban-paper.pdf (20.4.2014).
- Sohn, G., Dowman, I., 2007. Data fusion of high-resolution satellite imagery and LiDAR data for automatic building extraction. *ISPRS Journal of Photogrammetry and Remote Sensing*, 62 (1), 43–63.
- Solberg, A. H. S., Taxt, T., Jain, A. K., 1996. A Markov random field model for classification of multisource satellite imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 34 (1), 100–113.
- Song, W., Keller, J. M., Haithcoat, T. L., Davis, C. H., 2009. Automated Geospatial Conflation of Vector Road Maps to High Resolution Imagery. *IEEE Transactions on Image Processing*, 18 (2), 388–400.
- Sowmya, A., Trinder, J., 2000. Modelling and representation issues in automated feature extraction from aerial and satellite images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 55 (1), 34–47.
- Sun, S., 2013a. An optimized rubber-sheet algorithm for continuous area cartograms. *Professional Geographer*, 65 (1), 16–30.
- Sun, S., 2013b. A fast, free-form rubber-sheet algorithm for contiguous area cartograms. *International Journal of Geographical Information Science* 27 (3): 567–593.
- Tobler, W., 2004. Thirty five years of computer cartograms. *Annals of the Association of American Geographers*, 94 (1), 58–73.
- Usery, E. L., Finn, M. P., Starbuck, M., 2009. Data layer integration for the national maps of the United States. *Cartographic Perspectives*, 62, 28–41.
- Von Gösseln, G. V., Sester, M., 2004. Integration of geoscientific data sets and the German digital map using a matching approach. XXth Congress of the International Society for Photogrammetry and Remote Sensing, Commission IV, XXXV-B4, 1247, Istanbul, Turkey.

- Waldhoff, G., Curdt, C., Hoffmeister, D., Bareth, G., 2012. Analysis of multitemporal and multisensor remote sensing data for crop rotation mapping. XXIIth Congress of the International Society for Photogrammetry and Remote Sensing, Melbourne, Australia.
- Walter, V., Fritsch, D., 1999. Matching spatial datasets: A statistical approach. *International Journal of Geographical Information Science*, 13 (5), 445–473.
- Wang, Z., Zhang, B., Zhang, S., Li, X., Liu, D., Song, K., Li, J., Li, F., Duan, H., 2006. Changes of land use and of ecosystem service values in Sanjiang Plain, northeast China. *Environmental Monitoring and Assessment*, 112 (1-3), 69–91.
- Weber, K.T., Théau, J., Serr, K., 2008. Effect of coregistration error on patchy target detection using high-resolution imagery. *Remote Sensing of Environment*, 112 (3), 845–850.
- Weis, M., Müller, S., Liedtke, C. E., Pahl, M., 2005. A framework for GIS and imagery data fusion in support of cartographic updating. *Information Fusion*, 6 (4), 311–317.
- Wu, X., Carceroni, R., Fang, H., Kirmse, A., 2007. Automatic alignment of large-scale aerial rasters to road-maps. 15th annual ACM international symposium on Advances in geographic information systems, Seattle, WA, USA.
- Zhang, J., 2010. Multi-source remote sensing data fusion: status and trends. *International Journal of Image and Data Fusion*, 1 (1), 5–24.

5 Investigating Within-Field Variability of Rice from High Resolution Satellite Imagery in Qixing Farm County, Northeast China

Quanying Zhao ^{1,2,*}, Victoria I.S. Lenz-Wiedemann ^{1,2}, Fei Yuan ^{2,3}, Rongfeng Jiang ⁴ Yuxin Miao ^{2,4}, Fusuo Zhang ⁴, Georg Bareth ¹

Published in: ISPRS International Journal of Geo-information 2015, 4, 236–261.

doi:10.3390/ijgi4010236

Original manuscript is embedded in dissertation format.

- 1 Institute of Geography, GIS & RS Group, University of Cologne, Albertus-Magnus-Platz, 50923 Cologne
- 2 ICASD-International Center for Agro-Informatics and Sustainable Development, www.icasd.org
- 3 Department of Geography, Minnesota State University, Minnesota, MN 56001, USA
- 4 College of Resources and Environmental Sciences, China Agricultural University, Beijing China, 100193 Beijing

* Corresponding author, Tel.: +49-221-470-1951; E-mail: zhaquanying@gmail.com

Abstract

Rice is a primary staple food for the world population and there is a strong need to map its cultivation area and monitor its crop status on regional scales. This study was conducted in the Qixing Farm County of the Sanjiang Plain, Northeast China. First, the rice cultivation areas were identified by integrating the remote sensing (RS) classification maps from three dates and the Geographic Information System (GIS) data obtained from a local agency. Specifically, three FORMOSAT-2 (FS-2) images captured during the growing season in 2009 and a GIS topographic map were combined using a knowledge-based classification method. A highly accurate classification map (overall accuracy = 91.6%) was generated based on this Multi-Data-Approach (MDA). Secondly, measured agronomic variables that include biomass, leaf area index (LAI), plant nitrogen (N) concentration and plant N uptake were correlated with the date-specific FS-2 image spectra using stepwise multiple linear regression models. The best model validation results with a relative error (RE) of 8.9% were found in the biomass regression model at the phenological stage of heading. The best index of agreement (IA) value of 0.85 with an RE of 13.6% was found in the LAI model, also at the heading

stage. For plant N uptake estimation, the most accurate model was again achieved at the heading stage with an RE of 11% and an IA value of 0.77; however, for plant N concentration estimation, the model performance was best at the booting stage. Finally, the regression models were applied to the identified rice areas to map the within-field variability of the four agronomic variables at different growth stages for the Qixing Farm County. The results provide detailed spatial information on the within-field variability on a regional scale, which is critical for effective field management in precision agriculture.

Keywords: rice; FORMOSAT-2; agronomic variable; expert classification;

multi-data-approach (MDA); within-field variability; Sanjiang Plain; Northeast China

5.1 Introduction

Rice (*Oryza sativa* L.) is one of the most important staple food crops, feeding over half of the world's population. In 2010, the global rice production was approximately 672 million tons from a cultivation area of around 154 million ha, with China contributing to 29% and 19% of the rice production and cultivation area, respectively (FAO, 2014). Paddy rice management and its irrigation strategy have significant effects on greenhouse gas emissions (Yan, et al., 2003; Smith et al., 2008). Globally, rice paddies contribute about 10% of the total methane flux to the atmosphere (Global Methane Initiative, 2014). In 2000, the soil nitrogen (N), phosphorus (P), and Potassium (K) nutrient deficit induced by rice production accounted for 42% of the global deficit amount (Tan et al., 2005). Paddy rice agriculture in China is therefore of national and global significance in terms of both food security and sustainable development.

In the past years, remote sensing (RS) as an advanced technology has been used extensively in agriculture to obtain spatial and temporal information about crops (Kumar and Monteith, 1981; Moulin et al., 1998; Hatfield et al., 2008; Gitelson et al., 2014). Paddy rice areas were well projected using RS techniques (Frolking et al., 2002; Xiao et al., 2005; Kuenzer and Knauer, 2013). Geographic Information System (GIS) data have been proved to be important to enhance the accuracy of land use and land cover classification (Waldhoff et al., 2012; Rozenstein and Karnieli, 2011). RS data with coarse and medium resolution are widely used in rice cultivation research (Kuenzer and Knauer, 2013; Chang et al., 2005; Wang et al., 2010; van Niel and McVicar, 2004; Duggin and Piwinski, 1984; McCloy et al., 1987; Okamoto et al., 1998; Che and Price, 1992; Fang et al., 1998; Kim and Yeom, 2012; Son et al., 2013; Huang et al., 2013; Peng et al., 2014). However, the number of conducted studies on rice using high resolution RS images was limited in the past two decades (Kim and Yeom 2012;

Chang et al., 2013; Kim and Yeom, 2014). Identification of rice cultivation areas and estimation of agronomic parameters from high resolution images are valuable for improving rice production.

Mapping rice cultivation areas accurately is fundamental for the assessment of agricultural and environmental productivities, the analysis of food security and therefore national and international food trade decisions (MacDonald and Hall, 1980; Xiao et al., 2006; Duveiller et al., 2012a). Many previous studies have mapped rice areas (Kuenzer and Knauer, 2013). Various classifiers, including maximum likelihood, artificial neural network, decision tree, and spatial reclassification kernel, have been used in vegetation mapping using RS. However, classification accuracies are mainly determined by the quality and quantity of RS data. No ideal image classifier is uniformly applicable to all tasks (Xie et al., 2008). Extensive field knowledge and auxiliary data help to improve the classification accuracy (Xie et al., 2008). In the Multi-Data-Approach (MDA) for land use and land cover classifications as well as crop rotation mapping, multi-temporal RS and GIS vector data are combined to derive as much information as possible. Studies (Wang et al., 2010; van Niel and McVicar, 2004) demonstrated that considering extensive field knowledge and ancillary GIS data in the post-classification process can improve the classification accuracy by up to 10%.

Agricultural RS refers to the method of non-contact measurements of electromagnetic radiation reflected or emitted from plant materials or soils in agricultural fields (Mulla, 2013). Different vegetative covers can be distinguished according to their unique spectral behavior in relation to overall ground elements (Tucker, 1979): Visible radiation in the red (630–690 nm) is absorbed by chlorophyll while radiation in the near infrared (760–900 nm) is strongly reflected by leaf cellular structures. Vegetation Indices (VIs) are developed to qualitatively and quantitatively evaluate vegetative characteristics by combining spectral measurements from different wavelength channels (Bannari, et al., 1995). Theoretically, the VIs should be particularly sensitive to vegetative covers, insensitive to non-vegetation factors such as soil properties, atmospheric effects and sensor viewing conditions (Jackson et al., 1983). In practice, factors of soil characteristics, atmosphere and sensor radiometry degradation, as well as differences in the spectral responses and bidirectional effects all have considerable effects on vegetation indices. Therefore, many VIs have been developed to enhance the vegetative cover signal while minimizing the background response (Bannari et al., 1995). Hansen and Schjoerting (2003) discussed the optimized NDVI from different band centers and band widths to represent wheat parameters such as biomass, leaf area index (LAI), chlorophyll, and N status. In their study, the partial least square regression algorithm was applied. Yao

et al. (2012) successfully explored empirical models based on ground-based RS techniques to estimate N status and improve N use efficiency in rice. Yu et al. (2013) investigated the potential of hyperspectral band combinations in interpreting canopy N status in rice. Despite the lack of interpretations of the physical mechanisms and interactions between the target properties and the measured signals, these empirical models are fundamental materials for further research in order to represent the physical reality.

Remotely sensed photosynthetically active radiation has been used to evaluate the primary production of crops (Hall et al., 1990; Baret et al., 2007). The most sophisticated method commonly combines RS data with dynamic crop growth models (Bouman, 1992; Prasad et al., 2006). However, these methods require numeric parameters and are sensitive to soil background conditions.

Precision agriculture (PA) emerged in the middle of the 1980s (Mulla, 2013). Later, in 1991, satellite data were firstly used in PA (Bhatti et al., 1991). Spatial and temporal variability of soil and crop factors within a field is the essential base of PA (Zhang et al., 2002). Measurement of various crop canopy variables during the growing season provides an opportunity for improving grain yield and quality by site-specific fertilizer applications. Due to the ability of providing high temporal, spatial and spectral resolution images, satellite RS has a significant potential in PA (Mulla, 2013; Thenkabail, 2003). The multi-temporal within-field information on crop status captured by satellite RS is invaluable in PA.

Because of its operational and economical uses over large areas, satellite RS technology has been widely used to conduct in-season crop yield forecasting for decision making on marketing intervention and policy support on regional or global scales (MacDonald and Hall, 1980; Duveiller et al., 2012b). Satellite RS is also an essential technique for agro-ecosystem studies on regional scales (Donlon et al., 2012; Duveiller et al., 2012a; Duveiller and Defourny, 2010).

The overall aim of this study is to investigate the possibilities and accuracies of within-field variability of rice status monitoring during the entire growing season on a county scale (Qixing Farm) using satellite RS data. The term “within-field variability” in this study refers to the spatial variability of agronomic variables (biomass, LAI, N concentration, N uptake etc.) within a rice crop field defined by enclosed boundaries. The size of a rice field in this study ranges from 0.5–100 ha. First, a map of rice cultivation areas was produced using multi-temporal RS FORMOSAT-2 (FS-2) images, coupled with auxiliary GIS data (MDA) to improve the classification accuracy. Then, empirical regression models were developed to relate RS spectra with field measurements of the four agronomic variables. One advantage of this

research is the construction of specific regression models for different growth stages. Based on a stepwise regression analysis, the optimized predictors from the satellite images were identified to investigate the dynamics of crop canopy characteristics at different stages. The specific objectives of this study are to (1) identify rice cultivation areas with high classification accuracies based on multi-temporal RS and GIS data; (2) develop regression models for deriving rice crop variables from the FS-2 imagery; (3) validate the ability of the regression models to estimate rice parameters; (4) apply the regression models to the entire Qixing Farm County study area.

The presented method of extracting high resolution within-field variability information from the FS-2 imagery will assist farmers in their site-specific rice crop management and strategy planning.

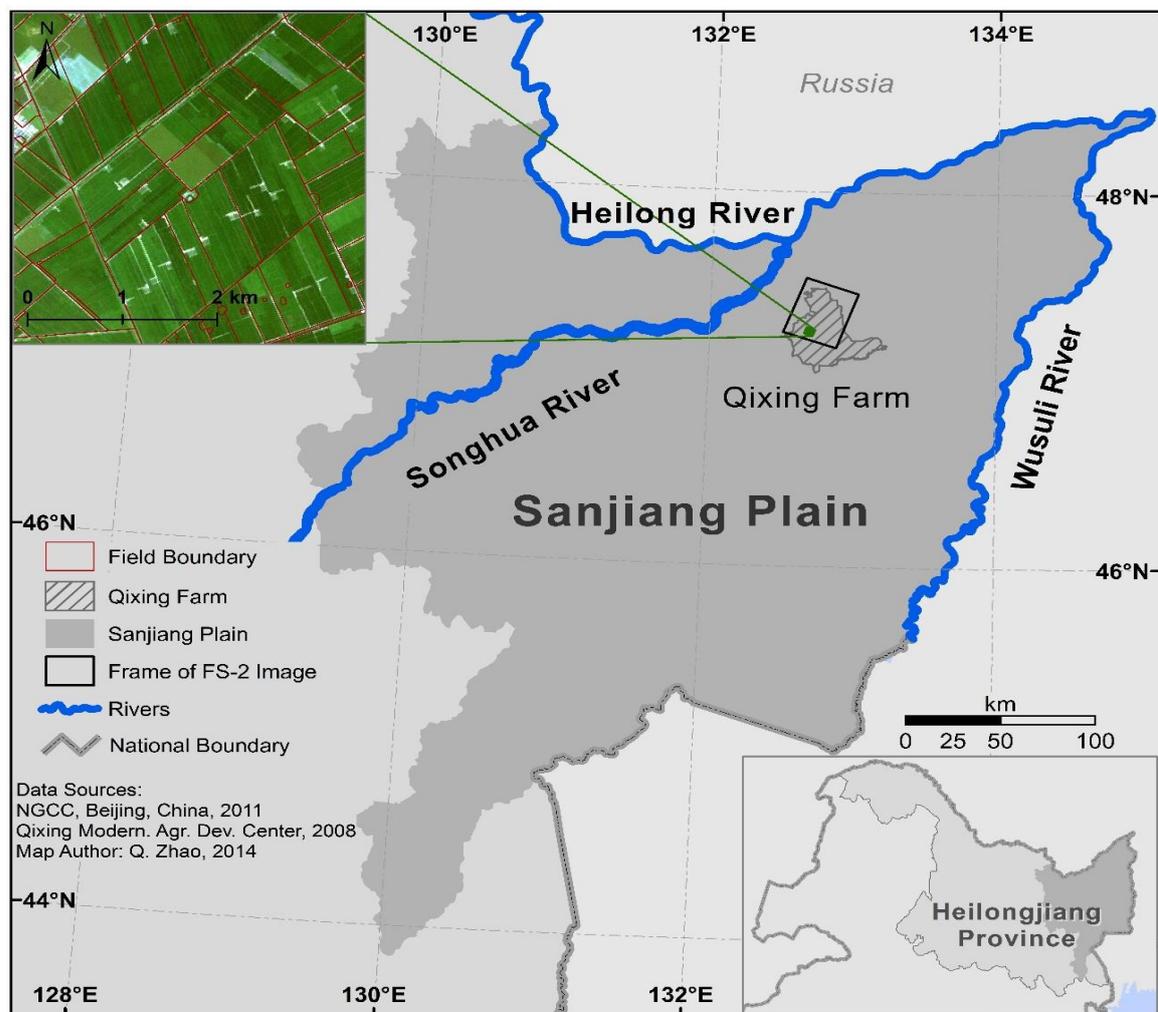


Figure 5-1: Location of the study area in Northeast China (the upper left corner shows a subset of the FS-2 image acquired on 9 August 2009).

5.2 Study area

The study site Qixing Farm (47.2°N, 132.8°E), is located in the Sanjiang Plain (SJP) in North-eastern China (see **Figure 5-1**). The SJP is an alluvial plain formed by the Songhua River, the Heilong River and the Wusuli River. The administrative area of the Qixing Farm County is about 120,000 ha. The climate in the SJP is temperate sub-humid, with a mean annual precipitation of 500–650 mm. Rainfall mainly occurs from May to September during the growing season of crops. The accumulated ≥ 10 °C temperature all across the year is about 2300–2500 °C and only single-season crops are planted. The topography is rather flat with an average elevation of 60 m and is characterized by broad alluvial plains and low terraces formed by the rivers. The SJP, which covers an area of more than 100,000 km², exceeding the size of the Netherlands almost by three times, is one of the major agricultural areas of China. In 2009, arable land in the SJP accounted for nearly 60% of the total land, being dominated by paddy rice (57%) (Ouyang et al., 2013). Since the irrigation constructions (water channels, raised ridges) for paddy rice are reusable year by year, the field boundaries are mostly stable. In recent years, for better economic profit, there has been moderate land use change from dryland to paddy rice. Compared with Europe's highly fragmented agricultural landscapes which prevent the utility of coarse resolution RS data for quantitative crop monitoring and yield forecasting (Xie et al., 2008), SJP's large homogenous landscapes provide an ideal site for monitoring crops using satellite RS.

5.3 Data

5.3.1 Satellite RS images and GIS data

FORMOSAT-2 (FS-2) collects multispectral images with a ground pixel resolution of 8×8 m² over a swath of 24 km. The FS-2 images used in this study are optical images with 4 bands of blue (450–520 nm), green (520–600 nm), red (630–690 nm), and near-infrared (760–900 nm). Three tiles of high quality images covering the main arable land area (~56,000 ha) of the Qixing Farm were captured on 24 June, 6 July, and 9 August, in 2009. Thus, both the vegetative phase (24 June and 6 July) and the reproductive phase (9 August) of rice are well represented in these images.

GIS vector data of Qixing Farm field boundaries were provided by the Qixing Modern Agriculture Research Center. Information on crop field boundaries, irrigation wells, water drainages, and shelter forests edges are given at a fine field unit scale. Additional information such as crop type of each field is given in the GIS attribute table.

5.3.2 Ground truth data collection

Field campaigns for the agronomic data collection were carried out during the entire rice growing season in 2009. In total, 42 sample sites covered by the FS-2 images were selected for this study. All these 42 sites were located in seven farmers' fields being spatially separated. Each site was represented by one plot covering approximately 0.1–0.3 ha. The final plant samples collected from each site were a mixture of three or four spatially separated samples, taken from the same plot. As ground truth data, the areas of sample sites were mapped using a Trimble™ Global Positioning System (GPS) receiver. Field management calendars of transplanting, N topdressing, irrigation, application of insecticides, and harvest dates, were recorded. Several field campaigns were carried out to collect samples from the tillering stage, booting stage, heading stage, 20 days after heading, and the harvest stage. For each site, biomass, LAI and plant N concentration were measured and plant N uptake was calculated as well.

After the field sampling, the plants were first cleaned, and then separated into different organs (leaves, stems, panicles) to measure the biomass values. LAI was measured using a sub-sample of the leaf biomass. One sub-sample consisted of 10–20 leaves, randomly selected among the youngest fully developed leaves. All fresh samples were processed in the oven at 105 °C for half an hour to stop enzyme activity. After that, they were dried at 75 °C for at least 72 h until a constant weight was reached before they were finally weighted. N concentration was measured using the Kjeldahl-N method. The plant N uptake was calculated as the aboveground dry mass multiplied by the N concentration. Detailed information is listed in **Table 5-1**.

In the study area, rice cultivation technologies in high latitude are relatively sophisticated and rice cultivation regulations developed by the government are applied by most of the farmers. The rice seedlings are first grown in greenhouses and are then transplanted into the paddy fields. In the regulations, a date window of 15–25 May is suggested for transplanting in order to capture the maximum accumulated temperature throughout the whole year. In this research, we divided our ground truth data sets into two groups according to the transplanting dates. Sample sites with seedlings transplanted from 15–25 May were used to construct empirical regression models between the agronomic parameters and the RS data. Sample sites with seedlings transplanted beyond those dates were used as validation data sets to evaluate the performance of the regression models.

Table 5-1: Agronomic parameters of the sample sites in 2009.

Field	Number of Sample Sites	Variety	Number of Foli-planting age	Trans-planting Date	Quantity of N Topdressing (kg/ha)	Plant Density (plants/m ²)	Soil Properties				
							pH	K (mg/kg)	P (mg/kg)	N (mg/kg)	SOM (g/kg)
Field 1	4	Longjing 21	12	May 10	104, 104, 104, 104	109	6.05	75.7	26.7	181.1	37.98
	3	Kongyu 131	22	May 17	0, 94, 141						
	1	Xixuan 1	13	May 15	132	101	6.22	134.5	36.6	177.4	35.47
Field 2	1	Longdun 104	12	May 16	118		6.07	142.5	30.1	173.6	46.85
	1	Longjing 21	12	May 17	118	133	6.08	82.4	28.3	175.8	37.78
	4	Longjing 21	12	May 13	84, 84, 84, 84	115	5.59	103.8	33.2	259.9	41.83
Field 3	3	Longjing 21	12	May 13	0, 67, 101	115					
	4	Longjing 21	12	May 12	67, 67, 67, 67	117	6.13	126.5	24.0	209.0	49.52
	4	Longjing 21	12	May 12	0, 54, 81, 67	117					
Field 4	4	Kendao 6	11	May 20	83, 83, 83, 83	158	6.29	372.5	21.5	249.8	66.03
	4	Longjing 24	11	May 16	95, 95, 95, 95	166	5.84	73.4	37.8	176.2	35.76
	1	Longjing 24	11	May 15	0	166					
Field 5	1	Longdun 249	12	May 20	102		5.78	126.5	30.6	169.7	37.08
	1	Chaoyou 949	11	May 18	106	144	6.08	98.4	32.0	187.5	39.91
	3	Jinxuan 1	11	May 20	0, 86, 129						
Field 7	3	Kongyu 131	11	May 17	0, 82, 123	122					

5.4 Methods

5.4.1 Satellite image pre-processing

Pre-processing of satellite images prior to vegetation information extraction is essential to remove noise and increase the interpretability of image data (Dhaliwal and Benbasat, 1996). Solar radiation reflected by the earth surface to the satellite sensors is significantly affected by its interaction with the atmosphere. Atmospheric correction is an important pre-processing step for satellite RS (Saastamoinen, 1972; Kaufman and Sendra, 1988). The uncertainties resulting from atmospheric effects in agriculture applications of satellite RS have been well discussed during the past decades (Duggin and Piwinski, 1984; Che and Price, 1992; Courault et al., 2003; Hadjimitsis et al., 2010). In this study, atmospheric correction was performed using ENVI FLAASH, version 5.1. **Table 5-2** lists the main parameters used in the atmospheric correction process. The sub-arctic summer model for rural region was selected.

Table 5-2: Main atmospheric correction parameters for the FS-2 images.

Date	Visibility (km)	Zenith Angle	Azimuth Angle
June 24	50	146°10'0.84"	-115°13'19.56"
July 6	50	152°15'58.34"	-45°34'58.34"
August 9	50	155°12'2.90"	-83°58'22.08"

Geometric distortion is another important factor affecting the results of image processing, especially when combining geospatial data from different dates or multiple sources. Dai and Siamak (Dai and Khorram, 1998) noted that the geometric error even on a sub-pixel level can significantly affect the accuracy of land use classification from satellite images. The precision of geometric correction depends on the number, distribution, and accuracy of the Ground Control Points (GCPs) (Moré and Pons, 2011). To avoid labor intensive work of in-situ GCP collection, Zhao et al. (2015) developed a geometric correction method for georeferencing multi-source geodata by using TerraSAR-X data as a reference. The same method was applied in this study. Specifically, a stacked TerraSAR-X image produced from five dates served as the reference to georectify the FS-2 images. The main georeferencing parameters are shown in **Table 5-3**. The positional error (PE) was less than 6 m for all three images. More details about the method can be found in Zhao et al. (2015).

Table 5-3: Main georeferencing parameters for the FS-2 images.

Image Capture Date	Number of Control Points	Number of Check Points	PE of Control Points (m)	PE of Check Points (m)
June 24	100	20	2.991	5.901
July 6	104	20	4.143	5.032
August 9	101	20	3.353	5.666

The mean reflectance and VIs for each sample plot were calculated using the “Zonal Analysis” tool in ERDAS IMAGINE 2013. For each of the three FS-2 images, correspondingly 42 RS samples were extracted and used to develop the empirical regression models for crop status monitoring.

5.4.2 Mapping rice cultivation areas

A rice area map of the Qixing Farm in 2009 was produced using aforementioned MDA. In particular, multi-temporal RS and GIS data were integrated to improve the accuracy of rice mapping. Based on the properties/attributes of the different datasets, a knowledge base was constructed and implemented, using the Knowledge Engineering tool mounted in ERDAS IMAGINE 2013. A series of logical rules were created in the Knowledge Engineering tool to integrate the RS and GIS data in order to achieve higher classification accuracy.

The general steps for delineating rice areas are summarized as follows: (1) a supervised classification based on the maximum likelihood algorithm was carried out. In order to extract more unique spectral signatures, more than five subclasses were classified firstly (for the image of June 24: nine subclasses in total with one of them being rice; for the image of July 6: 16 subclasses in total with three of them representing rice; for the image of August 9: 11 subclasses in total with three representing rice); (2) the resulting subclasses were further combined to five main classes, including rice, dryland, forests, residential areas, and “other areas” for each image; (3) isolated pixels were eliminated using the “clump and eliminate” functions embedded in ERDAS IMAGINE 2013; (4) an expert classification system based on knowledge rules was implemented to integrate and improve the rice classification results from multiple dates; and (5) the vector GIS data were used as auxiliary dataset in the post-classification process to further improve the accuracies of rice classification.

In a MDA study, Waldhoff et al. (2012) reported that the support vector machine (SVM) approach and the maximum likelihood classifier (MLC) yielded similar classification accuracies. They found although the SVM method performed slightly better (up to 3%) in three of the four cases, the MLC had a shorter processing time. Therefore, the MLC was selected in

this study to derive multi-temporal high resolution land cover classifications over a large study area. After the classification, the rice classes derived from the three RS images were “combined” by logical rules to take advantage of the specific spectral information corresponding to growth stages. Specifically, in the first step of the supervised classification, to improve the producer’s accuracy of rice classes, relatively “broad” spectral information was selected as rice spectral signatures. Thus, some other land covers such as “dryland” and “bare soils” may have been classified into rice classes as well. However, these misclassified areas from one single image could be different from those areas from the other dates, because spectral differences between rice and other land covers vary with growth stage (Chen et al., 2011). In the next step, using the knowledge base, the pixels classified into rice classes in all three RS images were categorized into a new “rice class 1”. The areas classified into rice classes in two RS images were categorized into a new “rice class 2”. After that, these refined RS rice areas were further “combined” with the GIS data (rice area masks). To assign a pixel into the final rice class, the following conditions had to be satisfied: (1) this pixel was in the “rice class 1”; or (2) this pixel was in the “rice class 2” and in the GIS rice mask.

5.4.3 Ground truth data interpolation

Ground truth data for each site were collected at the tillering stage (28–30 June), the jointing stage (10–12 July), the heading stage (2–8 August), 20 days after heading (22–28 August), and the harvest time. The FS-2 images were acquired on 24 June, 6 July and 9 August, at slightly different dates from the field campaigns. Therefore, the values of the agronomic variables of biomass, LAI, plant N concentration and N uptake were interpolated for these three FS-2 dates. First, a specific polynomial growing curve for each site was constructed based on the time series of ground truth data. The values of the agronomic variables for the three RS-2 dates were then interpolated. These interpolated values were used as the new ground truth data to explore their relationships with the satellite image reflectance and VIs.

5.4.4 Development of regression models for deriving agronomic variables

A multiple linear regression method was constructed for each agronomic variable based on reflectance values and the VI derived from the FS-2 images and the corresponding field measurement. The VI used in this method was treated as an equivalent factor of reflectance. Correlation and regression analyses were performed in SPSS 21 (SPSS, Inc., Chicago, IL, USA). The format of the multiple linear regression models was as follows:

$$Y_E = \alpha_0 + \alpha_{Blue} \times R_{Blue} + \alpha_{Green} \times R_{Green} + \alpha_{Red} \times R_{Red} + \alpha_{NIR} \times R_{NIR} + \alpha_{NDVI} \times NDVI \quad (5 - 1)$$

The NDVI is calculated by the following function:

$$NDVI = (R_{NIR} - R_{red}) / (R_{NIR} + R_{red}) \quad (5 - 2)$$

R stands for the reflectance value at the subscripted satellite band. In Equation (1), Y_E stands for the estimated agronomic variable; α_i is the coefficient of the reflectance at the corresponding band/vegetation index. The performance of the regression model was evaluated by the coefficient of determination (R^2).

Biomass and LAI are essential crop physiological variables which determine the crop yield. LAI refers to the ratio of leaf surface area to ground area. It is a fundamental canopy parameter in agronomy and RS since it drives absorption of solar radiation and evapotranspiration for carbon assimilation, and thus primary production. In this research, both biomass and LAI were related to the FS-2 band reflectance based on the aforementioned Equation (5 - 1).

N is one of the most remobilizable elements during the reproductive stage in rice plants (Tanaka, 1956). Since its remobilization causes leaf senescence, it is directly related to crop productivity (Mae and Ohira, 1981). Accurate plant N status detection to develop site specific N management strategies for rice in the SJP is of importance regarding both agricultural and environmental aspects (Yu et al., 2013). In this study, plant N concentration and N uptake were also derived from the FS-2 images using the regression model represented by Equation (5 - 1).

5.4.5 Validation of the regression models

The regression models were evaluated in a validation analysis. The feasibility of the model was quantified by the statistical measures of relative error (RE) and index of agreement (IA). In a further step, scatterplots were generated to assess the performance of the regression models. The RE is the ratio of the Root Mean Square Error (RMSE) to the mean of observed values, describing the differences between the predicted and the observed values relative to the mean of the ground truth values. The IA represents the degree of agreement between the model estimations and observed values (Willmott, 1981). It is calculated as:

$$IA = 1 - \frac{\sum_{i=1}^n (O_i - P_i)^2}{\sum_{i=1}^n (|P_i - \bar{O}| + |O_i - \bar{O}|)^2} \quad (5 - 3)$$

where O_i is the observed value, P_i is the model-simulated value, and \bar{O} is the mean of observed values. The denominator in Equation (5 - 3) was defined as a “potential

error” by Willmott (1981); therefore the IA represents the ratio between the mean square error and the “potential error”. Although the IA is sensitive to extreme values (Legates and McCabe, 1999), it can be interpreted straightforwardly since it ranges from 0–1.

5.5 Results

5.5.1 Accuracy of rice area classification

To compare the classification accuracies from different data sources, 800 random points were created in the study area. More than 80 random points were distributed in each class. Error matrices were generated to quantify the classification accuracies. The kappa coefficients of agreement and overall accuracies for the rice class are shown in **Table 5-4**.

Table 5-4: Accuracies of rice maps produced from different data sources.

Data Source	User’s Accuracy of Rice	Producer’s Accuracy of Rice	Kappa Coefficient	Overall Accuracy
RS single date (August 9)	94.1%	82.6%	0.733	80.8%
RS multiple dates (3 dates)	89.4%	91.7%	0.781	85.0%
RS and GIS data combined	94.2%	92.7%	0.881	91.6%

With auxiliary GIS data, both the Kappa and the overall accuracy values for the rice class were improved by 0.148 and 10.8%, respectively. While a similar user’s accuracy was obtained after combining the three RS classification maps and the GIS ancillary data, the producer’s accuracy increased from 82.6% to 92.7%, implying a decrease of 10% in omission errors. These results are in line with the conclusions made by Shrestha and Zinck (2001) and Rozenstein and Karnieli (2011).

5.5.2 Empirical regression models

5.5.2.1 Model development

The parameters of the best qualified regression models for all agronomic variables are shown in **Table 5-5**. The coefficient of determination R^2 was used to evaluate the regression models.

Table 5-5: Parameters of empirical regression models.

Agronomic Variable	Phenological Stage	Number of Model Construction Sites	Regression Model Parameters								R ²
			Model Coefficient								
			α_0	α_{Blue}	α_{Green}	α_{Red}	α_{NIR}	α_{NDVI}			
Biomass (kg/ha)	Tillering	17	-316.063	52440	-25250					0.551 **	
	Booting	23	7280.379	13480	-121530	21270	-12113.182			0.765 **	
	Heading	23	5791.901	-148910		31710				0.519 **	
LAI	Tillering	17	-1.04	140	-50					0.458 *	
	Booting	23	-0.451				4.378			0.431 **	
	Heading	23	-5.463	-50	380	30				0.650 **	
N Concentration (%)	Tillering	17	4.371	-101.548	71.492					0.401 #	
	Booting	23	4.895	50.798	-22.079	-23.972	-4.718			0.269 #	
	Heading	23	0.193	-35.742	-44.725	105.14	3.459			0.075 #	
N Uptake (kg N/ha)	Tillering	17	-21.022	-370	260	140				0.273 #	
	Booting	23	-18.834	810	-1420	280				0.460 *	
	Heading	23	167.251	-4270		550				0.483 **	

Significance test: ** $p < 0.01$, * $p < 0.05$, # $p > 0.05$.

A stepwise multiple linear regression was applied to select the best RS predictors for the regression models at different phenological stages.

For the biomass regression models, the best R^2 values, ranging from 0.519–0.765, were achieved. At the tillering stage, only green and red bands were retained applied to represent the biomass; while at the booting stage, all the four bands as well as the NDVI value were included in the model. At the heading stage, only green and NIR bands were used. Across all three dates, the model for the booting stage (second date), near the height of the growing season, gives the highest R^2 value of 0.765.

The stepwise regression analysis for the LAI estimation chose red and green bands as the predictors in the tillering stage, NDVI as the only predicting variable in the booting stage, and all the four bands as predictors in the heading stage. The LAI estimation model with the best performance occurred was found at the heading stage when panicles grow outside of the rice stem, with an R^2 of 0.65.

Variations in N concentration are more difficult to detect than biomass or LAI development. The N concentration is a weak absorber of spectral radiation and occurs in very small quantities in leaf tissue (0.5%–3%) (O'Neill et al., 2002). The red and green bands of the FS-2 image were included in the regression model in the tillering stage, while all four bands were used in both the booting and heading stages. At the tillering and booting stages, when the plant N concentration is at a high level, higher associations were found between the plant N concentration and the RS spectra.

Stepwise regression analysis for N uptake included all four bands as predictors in the tillering and booting stages, whereas only green and NIR were needed in the heading stage. The highest R^2 was 0.483 at the heading stage.

5.5.2.2 Validation of the regression models

Model validation was conducted using independent ground truth data sets as mentioned in section 4.4.3.2. **Table 5-6** shows the statistics of the model validation.

Table 5-6: Validation results for the regression models.

Agronomic variable	Phenological Stage	Number of Validation Sites	Model Validation Parameter	
			RE (%)	IA
Biomass (kg/ha)	Tillering	12	39.4	0.38
	Booting	19	27.0	0.32
	Heading	19	8.9	0.68
LAI	Tillering	12	31.4	0.38
	Booting	17	15.1	0.60
	Heading	17	13.6	0.85
N Concentration (%)	Tillering	12	17.7	0.41
	Booting	19	15.3	0.48
	Heading	19	27.7	0.48
N Uptake (kg N/ha)	Tillering	12	39.9	0.39
	Booting	19	32.5	0.37
	Heading	19	11.0	0.77

Validation analyses (**Table 5-6**) showed that the regression model for biomass reached the lowest RE of 8.9% and highest IA of 0.68 at the heading stage. Across all three dates, for the biomass, LAI and N uptake estimation models, the RE value decreases as rice grows. The LAI model in the heading stage showed the lowest RE and highest IA values of 13.6% and 0.85, respectively. These results demonstrated that the models based on the FS-2 images performed better at later growth stages.

The RE for the N concentration model was the lowest (15.3%) at the booting stage, while the IA values increase slightly from 0.41 at the tillering stage up to 0.48 at the later booting and heading stages. For the N uptake models, the lowest RE (11%) and the highest IA (0.77) also occurred at the heading stage.

Across all three dates, relationships between the FS-2 derived values and the interpolated ground truth data were analyzed by scatterplots (**Figure 5-2**). The R² value was the highest (0.98) for biomass modeling results, followed by the N uptake model (0.93), whereas a moderate R² value of 0.78 and 0.84 was achieved for LAI and plant N concentration estimation. These results showed that the RS derived values were highly related to the ground truth values, and the linear regression trend lines were ideally close to the 1:1 line (dashed diagonal lines in the figures).

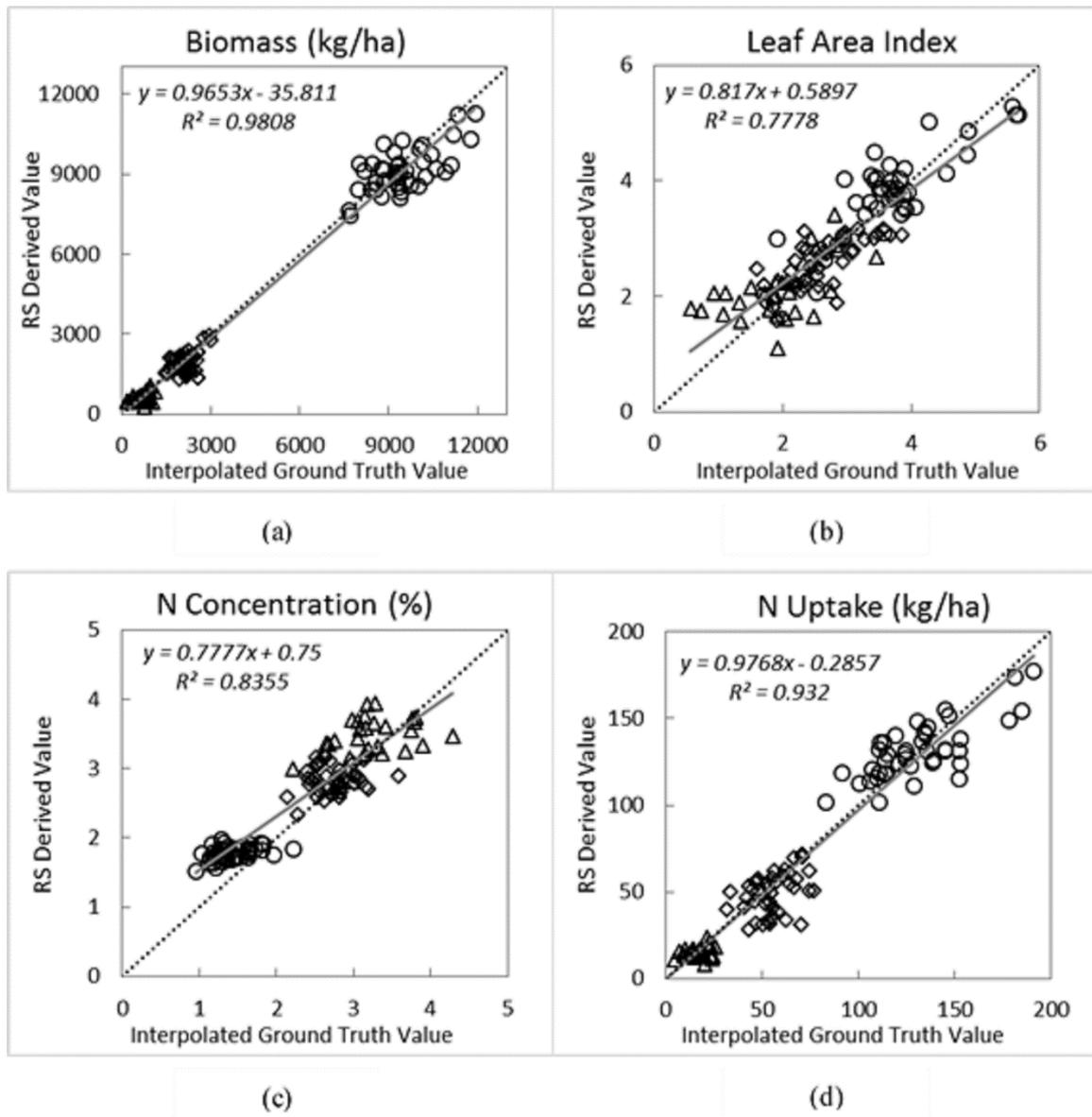


Figure 5-2: Relationships between RS derived values and ground truth values for (a) biomass; (b) LAI; (c) N concentration; and (d) N uptake at the tillering (Δ), booting (\diamond), and heading (\circ) stages.

5.5.3 Regional application of the regression models

Finally, the regression models were applied to the identified rice areas to map the within-field variability of the four agronomic variables for the area covered by the FS-2 images in the Qixing Farm County. As an example, the results for July 6 are shown in **Figures 5-3** to **5-6**. The within-field variability of the rice status can be easily detected in these maps. In the biomass map (**Figure 5-3**), unevenly distributed values and patterns can be identified in

most of the field blocks, revealing significant within-field variability of the growing status. Overall, a relatively higher biomass level is shown in the southern part.

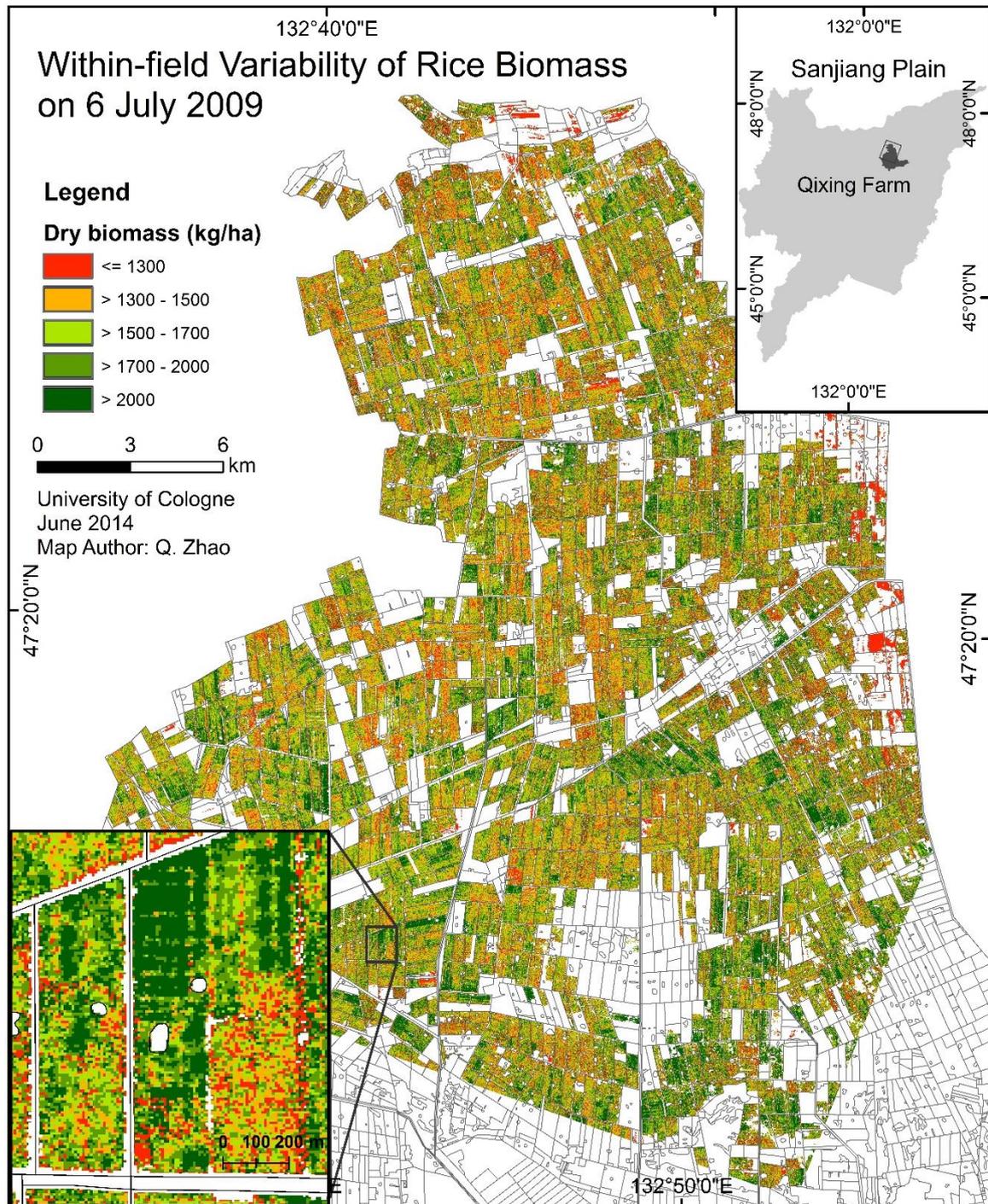


Figure 5-3: Within-field spatial variability of biomass derived from the FS-2 image of 6 July 2009.

The LAI map (Figure 5-4) reveals higher LAI values in the south, especially the southwest, region of the study area. Variations within the field blocks are also well detected in this map.

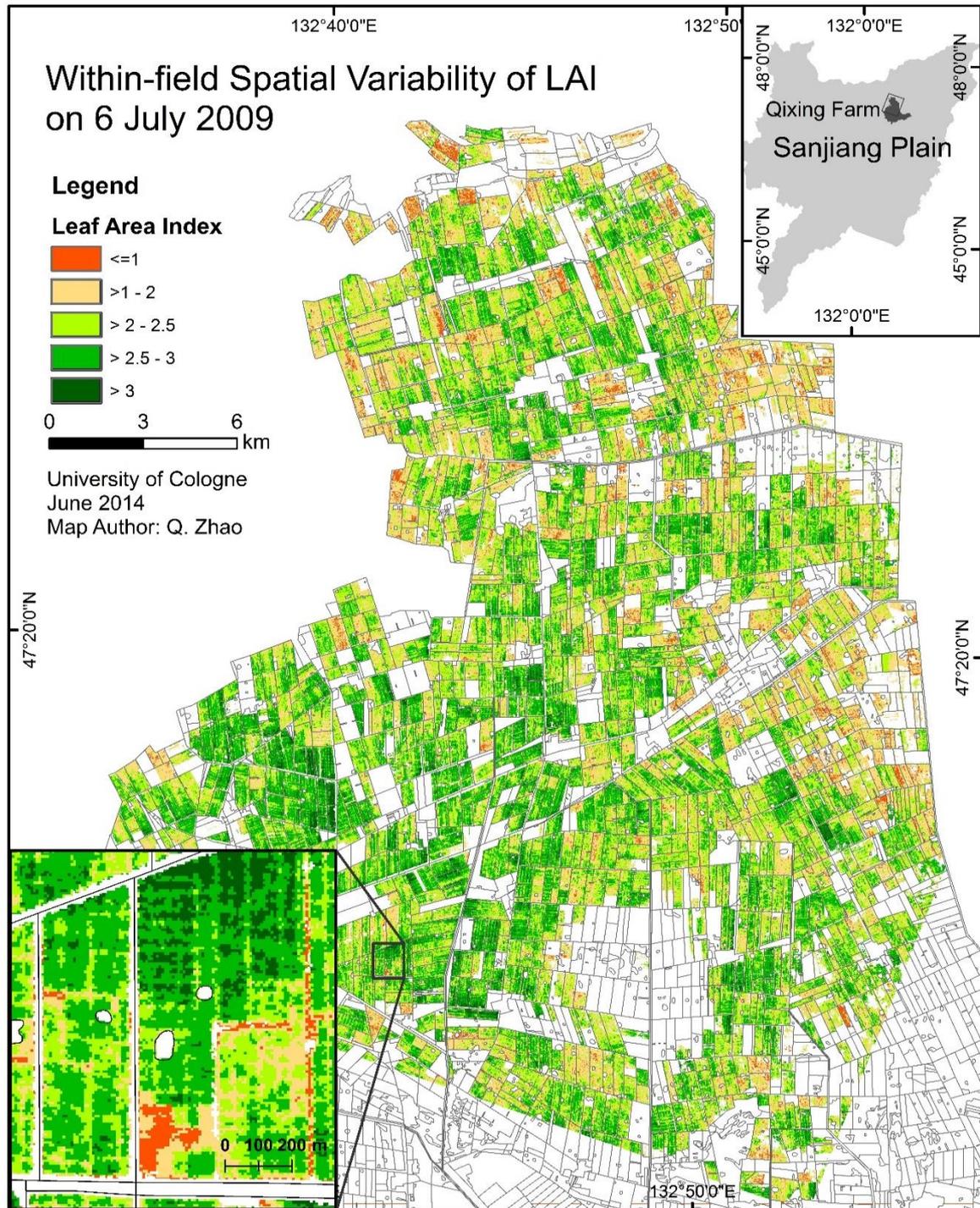


Figure 5-4: Within-field spatial variability of LAI derived from the FS-2 image of 6 July, 2009.

The N concentration distribution map (Figure 5-5) shows a general random pattern of spatial variation on the intra-field level. Most of the fields had N concentration values of 2.5–3.5%. However, the within-field variability is still presented.

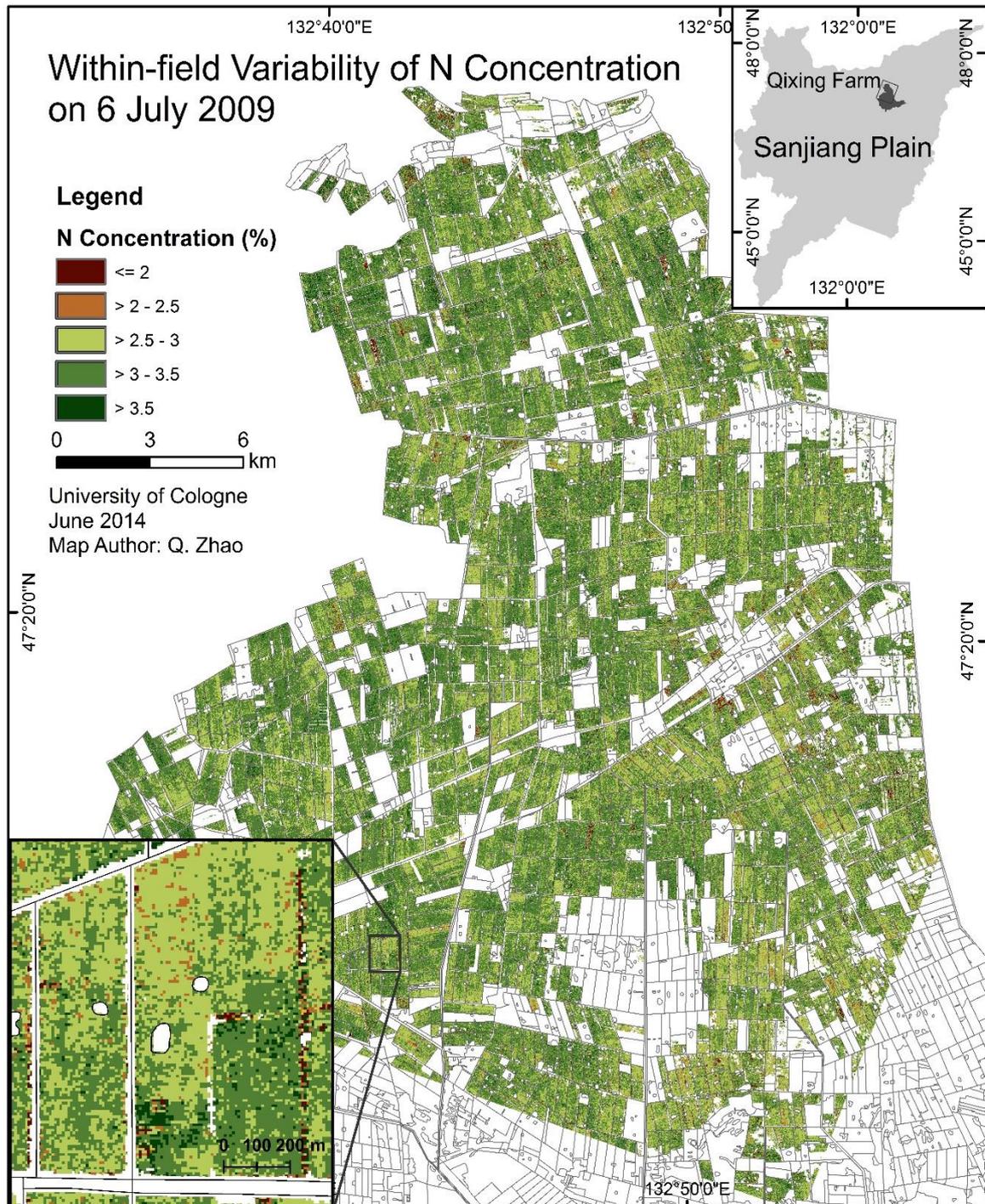


Figure 5-5: Within-field spatial variability of N concentration derived from the FS-2 image of 6 July 2009.

Likewise, the N uptake map (Figure 5-6) displays apparent within-field spatial variation. In the south and southeast part of the image, a relatively higher plant N uptake is detected.

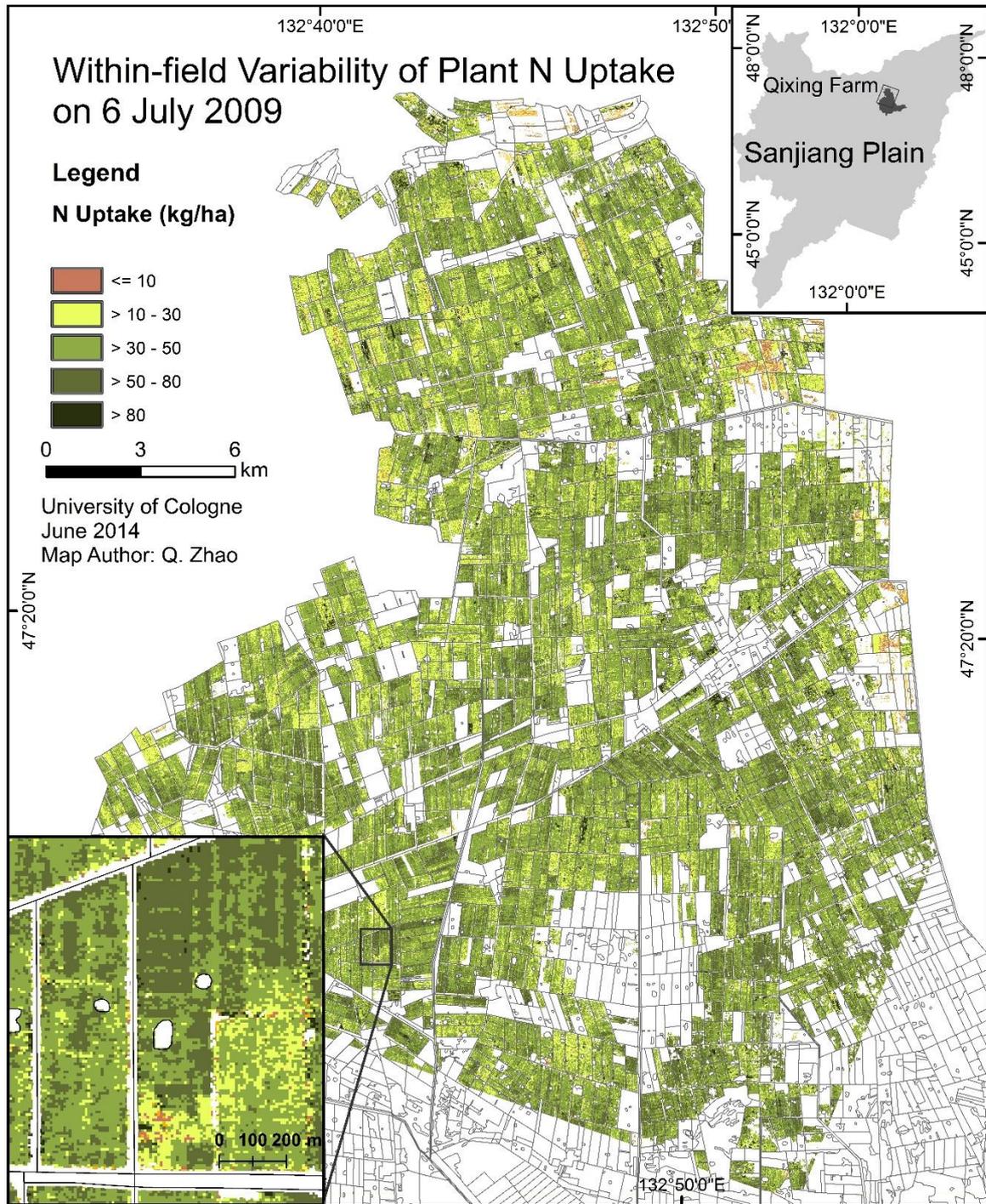


Figure 5-6: Within-field spatial variability of N uptake derived from the FS-2 image of 6 July 2009.

5.6 Discussion

5.6.1 *Band selection for different growth stages*

The recorded reflectance of optical imagery from a vegetated surface is a function of several physical properties such as vegetation structure, soil type, plant moisture content as well as sensor configuration. Plant pigments have distinct absorption spectra and some have an impact upon the reflectance spectra of leaves. This offers the opportunity of characterizing pigment concentration from reflectance spectra (Blackburn, 2007), and thus to interpret the plant status. Chlorophylls are one of the most important pigments because they control the absorption amount of solar radiation which determines the photosynthesis and, consequently, primary production. Several studies have demonstrated that the plant chlorophyll concentration indirectly indicates the plant nutrient status because the molecular structure of chlorophyll incorporates a large proportion of total leaf nitrogen (Filella et al., 1995; Moran et al., 2000; Yu et al., 2013).

In our results, different spectral bands were selected for the regression models at different growth stages. The red and green band combination performed well for most of the agronomic variables at the tillering stage, whereas green and NIR band combinations performed best for biomass and N uptake estimation at the heading stage. At the earlier growth (tillering) stage, the rice canopy coverage is lower (much less than 100%), resulting in a lower chlorophyll concentration on the pixel scale. In addition, the chlorophyll concentration on the leaf scale is lower compared to that at the heading stage (Chen et al., 2010; Cao et al., 2001). However, at the later (heading) stage, the canopy cover is higher, and the leaf level chlorophyll concentration is higher as well. The blue band did not frequently show up in the stepwise regression models. This may be attributed to the strong scattering effects of the atmosphere on blue radiation.

In addition, we found when the plant N concentration was relatively high, green (520–600 nm) and red (630–690 nm) bands were effective for the regression models. This result conforms to the finding of O'Neill et al. (2002) that the highest correlation between N concentration and reflectance occurred at 697.8 nm, followed by the region of 536.8–558.5 nm.

5.6.2 *Background effects in the early stage*

According to the model validation results of REs, it can be concluded that the regression models for the tillering stage did not perform as well as the ones for the other two stages, regardless of regarding all agronomic variables except for N concentration. The highest RE

of N concentration occurred at the heading stage, which mainly was caused by the low N concentration in leaf tissue at that stage, as mentioned in section 5.2.1.

According to the IA, the regression models performed best in the heading stage for all the agronomic variables, probably resulting from the higher biomass density and thus lower field soil and water effects compared to the early stages. This result is in conflict with the findings of Gnyp et al. (2014), who utilized hyperspectral field data for rice biomass estimation and achieved better results at the early growth stages. However, in the study of Gnyp et al. (2014), a specific vegetation index such as NDVI or RVI (ratio vegetation index) was used as the single independent variable in the regression models. They reported that NDVI became saturated at later growth stages before biomass reached 3 t/ha, which might have affected the model performance. They also noted that at the heading stage, the canopy reflectance signal became more complicated since panicles emerged from the sheath. In this study, the original multi-spectral reflectances in addition to NDVI were both used as the descriptive variables in the stepwise regression, which might help to remove the effects of the NDVI saturation problem on the model performance at later growth stages. Nevertheless, this needs to be further confirmed by analyzing the datasets from the two studies. Additionally, different measurement methods, RS instruments, physical conditions of the rice crop, sun angles, and sensor view angles, may also contribute to the discrepancy in these two studies.

In vegetation studies using satellite RS, several caveats have been noted by Myneni et al. (1995). These caveats include bidirectional effects, atmospheric effects, canopy structure effects, background or soil effects, nonlinear effects of scattering, effects of spectral heterogeneity, adjacent effects, nonlinear mixing, and topographic effects. In this study, the background soil and water effects and the rice canopy structure significantly affect the spectral response. Canopy conditions at different stages are clearly represented in the following photos (**Figure 5-7**). At the early stage, the rice canopy coverage was relatively low and therefore more soil and water background effects occurred. As one of the important canopy structure parameters, leaf orientation characterized by a smaller leaf angle (see **Figure 5-7**) at the early stage might result in the weak performance of the regression models as well.



Figure 5-7: Rice growth on 24 June (left), 6 July (middle) and 9 August (right).

5.7 Conclusions

Due to the integration of multi-temporal FS-2 imagery and GIS data, the rice cultivation areas in the Qixing Farm County were more clearly identified. The overall accuracy of the entire classified map was improved remarkably from 81.8%– to 91.6%. This highly accurate rice cultivation map provides an ideal basis for further analyses of rice crops in the study area.

This study showed that the performance of the regression models was significantly affected by rice growth stages. Thus, an optimized band selection for every growth stage is important due to the varying spectral reflectance properties. Based on the R^2 values, relatively higher goodness-of-fit values were found in the biomass and LAI estimation models than in the plant N uptake and plant N concentration models. In particular, for the estimation of plant N concentration, better model goodness-of-fit occurred at the earliest growth stage (tillering) when the N concentration was relatively high. The RS derived values and the interpolated ground truth values for all the three dates were highly correlated. The most accurate models with the lowest REs and the highest IA values were found at the heading stage for three of the four agronomic variables, except for the N concentration. In conclusion, this

study provides a framework and example of how high resolution satellite RS can support agricultural field management such as fertilizer, irrigation and pesticides management strategies by providing within-field agronomic information on a regional scales. The information derived from satellite RS could be further used to study the relations between crop growth and other phenomena such as carbon fixation, climate change, and sustainable management of natural resources.

Acknowledgements

This work was funded by the Natural Science Foundation of China (NSFC, project No. 31071859), the International Bureau of the German Federal Ministry of Education and Research (BMBF, project No. 1DO12013), and the German Research Foundation (DFG, project no. BA 2062/8-1). The authors gratefully acknowledge Yinkun Yao, Shanyu Huang, Minmin Su and Guangming Zhao from the China Agriculture University; Martin Gnyp and Christoph Hütt from the University of Cologne; Liqin Zhao from the Qixing Modern Agriculture Development Center for their supports in ground truth data collection. The authors are extremely thankful to the cooperating officers and farmers for their assistance in field experiments at the Qixing Farm. The Qixing Research and Development Center and the Jiansanjiang Agricultural Research Station, both in Heilongjiang, China, are gratefully acknowledged. The leading author is supported by the China Scholarship Council (November 2010–October 2014), Beijing, China.

Author Contributions

Quanying Zhao conducted the field campaigns, processed and analyzed the satellite images and the agronomic data. She also was responsible for writing the manuscript and incorporating the co-authors' contributions and comments. Victoria Lenz-Wiedemann and Fei Yuan intensively contributed to the writing of the manuscript and the data analysis. Rongfeng Jiang, Yuxin Miao and Fusuo Zhang designed and supervised the field campaigns, and contributed to the agronomic data analysis and those parts in the manuscripts. Georg Bareth intensively contributed to the structure of the manuscript, the GIS analysis and the satellite image classification. At different stages of the analysis and writing process, all co-authors contributed ideas for the preparation and improvement of the paper.

Conflicts of interest

The authors declare no conflict of interest.

References

- Bannari, A., Morin, D., Bonn, F., Huett, A. R., 1995. A review of vegetation indices. *Remote Sens. Rev.*, 13, 95–120.
- Baret, F., Hagolle, O., Geiger, B., Bicheron, P., Miras, B., Huc, M., Berthelot, B., Niño, F., Weiss, M., Samain, O., Roujean, J. L., Leroy, M., 2007. LAI, fAPAR and fCover CYCLOPES global products derived from VEGETATION: Part 1: Principles of the algorithm. *Remote Sens. Environ.*, 110, 275–286.
- Bhatti, A. U., Mulla, D. J., Frazier, B. E., 1991. Estimation of soil properties and wheat yields on complex eroded hills using geostatistics and thematic mapper images. *Remote Sens. Environ.*, 37, 181–191.
- Blackburn, G. A., 2007. Hyperspectral remote sensing of plant pigments. *J. Exp. Bot.*, 58, 855–867.
- Bouman, B. A. M., 1992. Linking physical remote sensing models with crop growth simulation models, applied for sugar beet. *Int. J. Remote Sens.*, 13, 2565–2581.
- Cao, S., Lu, W., Zhai, H., Sheng, S., Gong, H., Yang, T., Zhang, R., 2001. Research on the method to estimating flag leaf photosynthesis function duration at rice seedling stage by relative steady phase of chlorophyll content. *Chin. J. Rice Sci.*, 15, 309–313 (in Chinese).
- Chang, C. H., Liu, C. C., Tseng, P. Y., 2013. Emissions inventory for rice straw open burning in Taiwan based on burned area classification and mapping using FORMOSAT-2 satellite imagery. *Aerosol Air Qual. Res.*, 13, 474–487.
- Chang, K., Shen, Y., Lo, J., 2005. Predicting rice yield using canopy reflectance measured at booting stage. *Agron. J.*, 97, 872–878.
- Che, N., Price, J. C., 1992. Survey of radiometric calibration results and methods for visible and near infrared channels of NOAA-7, -9, and -11 AVHRRs. *Remote Sens. Environ.*, 41, 19–27.
- Chen, J., Huang, J., Hu, J., 2011. Mapping rice planting areas in southern China using the China Environment Satellite data. *Math. Comput. Model.*, 54, 1037–1043.

- Chen, X., Chen, C., Zhou, L., 2010. Measurements and relationship analyses of chlorophyll contents and phenological growing stages in rice. *Mod. Agricu. Sci. Technol.*, 17, 42–44.
- Courault, D., Seguin, B., Olioso, A., 2003. Review to estimate evapotranspiration from remote sensing data: Some examples from the simplified relationship to the use of mesoscale atmospheric models. In *Proceedings of the 2003 Workshop on Remote Sensing of ET for Large Regions*, Montpellier, France, 17 September 2003, pp.1–18.
- Dai, X., Khorram, S., 1998. The effects of image misregistration on the accuracy of remotely sensed change detection. *IEEE Trans. Geosci. Remote Sens.*, 36, 1566–1577.
- Dhaliwal, J. S., Benbasat, I., 1996. The use and effects of knowledge-based system explanations: Theoretical foundations and a framework for empirical evaluation. *Inf. Syst. Res.*, 7, 342–362.
- Donlon, C., Berruti, B., Buongiorno, A., Ferreira, M. H., Féménias, P., Frerick, J., Goryl, P., Klein, U., Laur, H., Mavrocordatos, C., Nieke, J., Rebhan, H., Seitz, B., Stroede, J., Sciarra, R., 2012. The global monitoring for environment and security (GMES) sentinel-3 mission. *Remote Sens. Environ.*, 120, 37–57.
- Duggin, M. J., Piwinski, D., 1984. Recorded radiance indices for vegetation monitoring using NOAA AVHRR data; atmospheric and other effects in multitemporal data sets. *Appl. Opt.*, 23, 2620–2623.
- Duveiller, G., Baret, F., Defourny, P., 2012a. Remotely sensed green area index for winter wheat crop monitoring: 10-Year assessment at regional scale over a fragmented landscape. *Agr. For. Meteorol.* 166, 156–168.
- Duveiller, G., Defourny, P., 2010. A conceptual framework to define the spatial resolution requirements for agricultural monitoring using remote sensing. *Remote Sens. Environ.*, 114, 2637–2650.
- Duveiller, G., López-Lozano, R., Seguni, L., Bojanowski, J. S., Baruth, B., 2012b. Optical remote sensing requirements for operational crop monitoring and yield forecasting in Europe. In *Proceedings of 2012 nel-3 OLCI/SLSTR and MERIS/(A) ATSR Workshop*, Frascati, Italy, 15–19 October 2012.
- Fang, H., Wu, B., Liu, H., Huang, X., 1998. Using NOAA's AVHRR and Landsat TM to estimate rice area year-by-year. *Int. J. Remote Sens.*, 19, 521–525.

- FAO (Food and Agriculture Organization of the United Nations), 2014. FAO Statistical Database. Available online: <http://www.fao.org/docrep/015/i2490e/i2490e03d.pdf> (accessed on 13 June 2014).
- Filella, I., Serrano, L., Serra, J., Peñuelas, J., 1995. Evaluating wheat nitrogen status with canopy reflectance indices and discriminant analysis. *Crop Sci.*, 35, 1400–1405.
- Frolking, S., Qiu, J., Boles, S., Xiao, X., Liu, J., Zhuang, Y., Li, C., Qin, X., 2002. Combining remote sensing and ground census data to develop new maps of the distribution of rice agriculture in China. *Global Biogeochem. Cycle*, 16, doi:10.1029/2001GB001425.
- Global Methane Initiative, 2014. Global Methane Emissions and Mitigation Opportunities. Available online: <http://www.globalmethane.org> (accessed on 26 August 2014).
- Gitelson, A. A., Peng, Y., Arkebauer, T. J., Schepers, J. 2014. Relationships between gross primary production, green LAI, and canopy chlorophyll content in maize: Implications for remote sensing of primary production. *Remote Sens. Environ.*, 144, 65–72.
- Gnyp, M. L., Miao, Y., Yuan, F., Ustin, S.L., Yu, K., Yao, Y., Huang, S., Bareth, G., 2014. Non-destructive estimation of paddy rice aboveground biomass at different growth stages with hyperspectral canopy remote sensing. *Field Crop Res.*, 155, 42–55.
- Hadjimitsis, D. G., Papadavid, G., Agapiou, A., Themistocleous, K., Hadjimitsis, M. G., Retalis, A., Michaelides, S., Chrysoulakis, N., Toullos, L., Clayton, C. R. I., 2010. Atmospheric correction for satellite remotely sensed data intended for agricultural applications: Impact on vegetation indices. *Nat. Hazard Earth Syst.*, 10, 89–95.
- Hall, F. G., Huemmrich, K. F., Goward, S. N., 1990. Use of narrow-band spectra to estimate the fraction of absorbed photosynthetically active radiation. *Remote Sens. Environ.*, 32, 47–54.
- Hansen, P. M., Schjoerring, J. K., 2003. Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. *Remote Sens. Environ.*, 86, 542–553.
- Hatfield, J. L., Gitelson, A. A., Schepers, J. S., Walthall, C. L., 2008. Application of spectral remote sensing for agronomic decisions. *Agron. J.*, 100, S117–S131.

- Huang, J., Wang, X., Li, X., Tian, H., Pan, Z., 2013. Remotely sensed rice yield prediction using multi-temporal NDVI data derived from NOAA's-AVHRR. *PLoS One*, 8, e70816.
- Jackson, R. D., Slater, P. N., Pinter Jr., P. J., 1983. Discrimination of growth and water stress in wheat by various vegetation indices through clear and turbid atmospheres. *Remote Sens. Environ.*, 13, 187–208.
- Kaufman, Y.J., Sendra, C., 1988. Algorithm for automatic atmospheric corrections to visible and near-IR satellite imagery. *Int. J. Remote Sens.*, 9, 1357–1381.
- Kim, H. O., Yeom, J. M., 2012. Multi-temporal spectral analysis of rice fields in South Korea using MODIS and RapidEye satellite imagery. *J. Astron. Space Sci.*, 29, 407–411.
- Kim, H. O., Yeom, J. M., 2014. Effect of red-edge and texture features for object-based paddy rice crop classification using RapidEye multi-spectral satellite image data. *Int. J. Remote Sens.*, 35, 7046–7068.
- Kuenzer, C., Knauer, K., 2013. Remote sensing of rice crop areas. *Int. J. Remote Sens.*, 34, 2101–2139.
- Kumar, M., Monteith, J. L., 1981. Remote sensing of crop growth. In *Plants and the Daylight Spectrum*, Smith, H., Ed., Academic Press: London, UK, pp. 133–144.
- Legates, D. R., McCabe, G. J., 1999. Evaluating the use of “goodness-of-fit” measures in hydrologic and hydroclimatic model validation. *Water Resour. Res.* **1999**, 35, 233–241.
- MacDonald, R. B., Hall, F. G., 1980. Global crop forecasting. *Science*, 208, 670–679.
- Mae, T., Ohira, K., 1981. The remobilization of nitrogen related to leaf growth and senescence in rice plants (*Oryza sativa* L.). *Plant Cell Physiol.*, 22, 1067–1074.
- McCloy, K. R., Smith, F. R., Robinson, M. R., 1987. Monitoring rice areas using LANDSAT MSS data. *Int. J. Remote Sens.*, 8, 741–749.
- Moran, J.A., Mitchell, A.K., Goodmanson, G., Stockburger, K. A., 2000. Differentiation among effects of nitrogen fertilization treatments on conifer seedlings by foliar reflectance: A comparison of methods. *Tree Physiol.*, 20, 1113–1120.
- Moré, J., Pons, X., 2011. Preliminary considerations about the assessment and visualisation of the quality on geometric corrections of satellite imagery depending on the number

- of ground control points. In Proceedings of the 7th International Symposium on Spatial Data Quality, Coimbra, Portugal, 12–14 October 2011, pp. 12–14.
- Moulin, S., Bondeau, A., Delecolle, R., 1998. Combining agricultural crop models and satellite observations: From field to regional scales. *Int. J. Remote Sens.*, 19, 1021–1036.
- Mulla, D. J., 2013. Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. *Biosyst. Eng.*, 114, 358–371.
- Tucker, C. J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.*, 8, 127–150.
- Myneni, R. B., Maggion, S., Jaquinta, J., Privette, J. L., Gobron, N., Pinty, B., Kimes, D. S., Verstraete, M. M., Williams, D. L., 1995. Optical remote sensing of vegetation: Modeling, caveats, and algorithms. *Remote Sens. Environ.*, 51, 169–188.
- Okamoto, K., Yamakawa, S., Kawashima, H., 1998. Estimation of flood damage to rice production in North Korea in 1995. *Int. J. Remote Sens.*, 19, 365–371.
- O'Neill, A. L., Kupiec, J. A., Curran, P. J., 2002. Biochemical and reflectance variation throughout a Sitka spruce canopy. *Remote Sens. Environ.*, 80, 134–142.
- Ouyang, W., Xu, Y., Hao, F., Wang, X., Chen, S., Lin, C., 2013. Effect of long-term agricultural cultivation and land use conversion on soil nutrient contents in the Sanjiang Plain. *Catena*, 104, 243–250.
- Peng, D., Huang, J., Li, C., Liu, L., Huang, W., Wang, F., Yang, X., 2014. Modelling paddy rice yield using MODIS data. *Agric. For. Meteorol.*, 184, 107–116.
- Prasad, A. K., Chai, L., Singh, R. P., Kafatos, M., 2006. Crop yield estimation model for Iowa using remote sensing and surface parameters. *Int. J. Appl. Earth Obs.*, 8, 26–33.
- Rozenstein, O., Karnieli, A., 2011. Comparison of methods for land-use classification incorporating remote sensing and GIS inputs. *Appl. Geogr.*, 31, 533–544.
- Saastamoinen, J., 1972. Atmospheric correction for the troposphere and stratosphere in radio ranging satellites. In *The Use of Artificial Satellites for Geodesy*, Henriksen, S.W., Chovitz, B.H., Mancini, A., Eds., American Geophysical Union: Washington D.C., WA, USA, pp. 247–251.

- Shrestha, D. P., Zinck, J. A., 2001. Land use classification in mountainous areas: Integration of image processing, digital elevation data and field knowledge (application to Nepal). *Int. J. Appl. Earth Obs.*, 3, 78–85.
- Son, N. T., Chen, C., Chen, C., Chang, L., Duc, H. C., Nguyen, L. D., 2013. Prediction of rice crop yield using MODIS EVI-LAI data in the Mekong Delta, Vietnam. *Int. J. Remote Sens.*, 34, 7275–7292.
- Smith, P., Martino, D., Cai, Z., Gwary, D., Janzen, H., Kumar, P., McCarl, B., Ogle, S., O'Mara, F., Rice, C., Scholes, B., Sirotenko, O., Howden, M., McAllister, T., Pan, G., Romanenkov, V., Schneider, U., Towprayoon, S., Wattenbach, M., Smith, J., 2008. Greenhouse gas mitigation in agriculture. *Philos. Trans. R. Soc. B.*, 363, 789–813.
- Tan, Z. X., Lal, R., Wiebe, K. D., 2005. Global soil nutrient depletion and yield reduction. *J. Sustain. Agr.*, 26, 123–146.
- Tanaka, A., 1956. Studies on characteristics of physiological function of leaf at definite position on stem of rice plant (Part 3). Relation between nitrogen metabolism and physiological function of leaf at definite position. *J. Sci. Soil Manure Jpn.*, 26, 413–418.
- Thenkabail, P. S., 2003. Biophysical and yield information for precision farming from near-real-time and historical Landsat TM images. *Int. J. Remote Sens.*, 24, 2879–2904.
- van Niel, T. G., McVicar, T. R., 2004. Current and potential uses of optical remote sensing in rice-based irrigation systems: A review. *Crop Pasture Sci.*, 55, 155–185.
- Waldhoff, G., Curdt, C., Hoffmeister, D., Bareth, G., 2012. Analysis of multitemporal and multisensor remote sensing data for crop rotation mapping. In *Proceedings of the 2012 International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Melbourne, Australia, 25 August–1 September 2012, pp. 177–182.
- Wang, Y., Chang, K., Chen, R., Lo, J., Shen, Y., 2010. Large-area rice yield forecasting using satellite imageries. *Int. J. Appl. Earth Obs.*, 12, 27–35.
- Willmott, C. J., 1981. On the validation of models. *Phys. Geogr.*, 2, 184–194.

- Xiao, X., Boles, S., Frohling, S., Li, C., Babu, J. Y., Salas, W., Moore, B., 2006. Mapping paddy rice agriculture in South and Southeast Asia using multi-temporal MODIS images. *Remote Sens. Environ.*, 100, 95–113.
- Xiao, X.; Boles, S., Liu, J., Zhuang, D., Frohling, S., Li, C., Salas, W., Moore, B., III, 2005. Mapping paddy rice agriculture in southern China using multi-temporal MODIS images. *Remote Sens. Environ.*, 95, 480–492.
- Xie, Y., Sha, Z., Yu, M., 2008. Remote sensing imagery in vegetation mapping: A review. *J. Plant Ecol.*, 1, 9–23.
- Yan, X., Ohara, T., Akimoto, H., 2003. Development of region-specific emission factors and estimation of methane emission from rice fields in the East, Southeast and South Asian countries. *Glob. Change Biol.*, 9, 237–254.
- Yao, Y., Miao, Y., Huang, S., Gao, L., Ma, X., Zhao, G., Jiang, R., Chen, X., Zhang, F., Yu, K., Gnyp, M. L., Bareth, G., Liu, C., Zhao, L., Yang, W., Zhu, H., 2012. Active canopy sensor-based precision N management strategy for rice. *Agron. Sustain. Dev.*, 32, 925–933.
- Yu, K., Leufen, G., Hunsche, M., Noga, G., Chen, X., Bareth, G., 2013. Investigation of leaf diseases and estimation of chlorophyll concentration in seven barley varieties using fluorescence and hyperspectral indices. *Remote Sens.*, 5, 64–86.
- Yu, K., Li, F., Gnyp, M. L., Miao, Y., Bareth, G., Chen, X., 2013. Remotely detecting canopy nitrogen concentration and uptake of paddy rice in the Northeast China Plain. *ISPRS J. Photogramm. Remote Sens.*, 78, 102–115.
- Zhang, N., Wang, M., Wang, N., 2002. Precision agriculture—A worldwide overview. *Comput. Electron. Agr.*, 36, 113–132.
- Zhao, Q., Hütt, C., Lenz-Wiedemann, V.I.S., Miao, Y., Yuan, F., Zhang, F., Bareth, G., 2015. Georeferencing multi-source geospatial data using multi-temporal TerraSAR-X imagery: A case study in Qixing Farm, northeast China. *Photogrammetrie Fernerkundung Geoinformation*, 2, 173–185.

6 Detecting Spatial Variability of Paddy Rice Yield by Combining the DNDC Model with High Resolution Satellite Images

Quanying Zhao ^{1,2,*}, Sebastian Brocks^{1,2}, Victoria I.S. Lenz-Wiedemann ^{1,2}, Yuxin Miao ^{2,3}, Fusuo Zhang ³, Georg Bareth ^{1,2}

Submitted in: Agricultural Systems

Original manuscript is embedded in dissertation format.

- 1 Institute of Geography, GIS & RS Group, University of Cologne, Albertus-Magnus-Platz, 50923 Cologne
- 2 ICASD-International Center for Agro-Informatics and Sustainable Development, www.icasd.org
- 3 College of Resources and Environmental Sciences, China Agricultural University, Beijing China, 100193 Beijing

* Corresponding author, Tel.: +49-221-470-1951; E-Mail: zhaoquanying@gmail.com

Abstract

Yield estimation over large areas is critical for ensuring food security, guiding agronomical management, and designing national and international food trade strategies. Besides, analyzing the impacts of managed cropping systems on the environment is important for sustainable agriculture. In this study, the agro-ecosystem model DNDC (DeNitrification-DeComposition) and FORMOSAT-2 (FS-2) imagery were used to detect spatial variabilities of paddy rice yield in the Qixing Farm in 2009. The Qixing Farm is located at the center of the Sanjiang Plain in NE-China, which is one of the important national food bases of China. The site-specific mode of the DNDC model was adapted due to its advantages of better transferability and flexibility. It was generalized onto a regional scale by programming a set of scripts using the Python programming language. Soil data were prepared as model inputs in 100 m raster files. The spatial variabilities in modelled yields were well detected based on the detailed soil data and an accurate rice area map. Rice yield was also derived from multiple vegetation indices based on the FS-2 imagery. It was found that the highest coefficient of model determination (CD) and index of agreement (IA) for the modelled yield were 2.63 and 0.74, respectively, while for the RS-derived yield, the highest CD and IA were 1.2 and 0.55,

respectively. Results from both methods were comparable and each method has its own advantages.

Key words: yield estimation; agro-ecosystem model; regionalization; soil characteristics; FORMOSAT-2; Sanjiang Plain

6.1 Introduction

Yield estimation is significant concerning to global food security issues, food trade strategies, and agro-ecosystems studies (Godfray et al., 2010; Ericksen et al., 2009). Agro-ecosystem modelling is an effective means in yield estimation due to its capability of using various environmental factors (model input data) for simulating crop production under different scenarios. Remote sensing (RS) is another effective technology for agricultural applications in terms of yield estimation and sustainability (Atzberger, 2013; Liaghat and Balasundram, 2010). As one of the major food crops, rice is crucial not only for national economy and food security especially in the developing world, but also for the natural ecosystems (Tilman et al., 2002). Many rice-related studies have been reported. These studies have focused on agro-ecosystem analyses using process-based models and rice area delineation and status monitoring using RS technologies (e.g., Fumoto, et al., 2008; Zhang et al., 2011; Kuenzer and Knauer, 2010; Zhao et al., 2015a; Huang et al., 2015).

In agro-ecosystem models, plant development, growth, and their effects on the environment are calculated as functions of environmental parameters and agricultural management data. Many of such models have been designed and used at field scale, whereas models that integrate abundant information over large geographic extents on regional and global scales are needed as well. The significance of generalizing site-specific model into a regional extent has been proposed in recent studies (Resop et al., 2012; Thorp and Bronson, 2013). The quality and spatial coverage of the available model inputs are key considerations in implementing agro-ecosystem modellings on large geographic scales (Bareth, 2009). With high spatial heterogeneity, the quality (i.e., scale, accuracy) of soil and management data critically determine the accuracies of model results on spatial variability (Hansen and Jones, 2000; Kersebaum et al., 2007).

This study was conducted for Qixing Farm in the Sanjiang Plain (SJP) of NE-China in 2009. The DNDC model (DeNitrification-DeComposition, version 9.5; developed by Li et al., 1992, modified by Li et al., 2000) was chosen to detect spatial variabilities of paddy rice yield. The DNDC model is one of the few agro-ecosystem models that are designed with two modes -site-specific and regional (Li et al., 2004). However, the site-specific mode is

more flexible and transparent because more input parameters can be adjusted (Perlman et al., 2013). Therefore, the site-specific mode of the DNDC model was used and generalized to a regional scale through a technical improvement. Raster files of soil properties (e.g., SOC, clay fraction, etc.) in 100 m spatial resolution were prepared as model inputs. Additionally, FORMOSAT-2 (FS-2) satellite imagery with a ground pixel resolution of 8 m was used. Rice yield was estimated based on the multi-temporal vegetation indices derived from the FS-2 imagery.

The objectives are to 1) apply the DNDC model and assess its performance in estimating rice yield in the SJP; 2) generalize the site-specific model for regional applications; 3) investigate the effects of soil heterogeneity on spatial variability in rice yield; and to 4) compare and assess the modelled and RS-derived rice yields.

6.2 Materials and Methods

6.2.1 Study area

The study area Qixing Farm (47.2 °N, 132.8 °E), covering an area of approximately 120,000 ha, is a typical state-owned farm located at the center of the SJP in NE-China (**Figure 6-1**). The SJP, covering more than 100,000 km², is an alluvial plain formed by the Songhua, Heilong and Wusuli rivers. It is dominated by temperate sub-humid monsoon climate, with a mean annual precipitation of 500 – 650 mm and an annual average temperature between – 4 ° C and 4 ° C (Yan et al., 2001). Accumulated temperature (≥ 10 ° C) in the SJP ranges from 2300 °C to 2500 °C across the year. Only single-season crops are planted. Four main soil types - dark brown forest soil, meadow soil, albic soil and bog soil -cover 95% of the whole area of the SJP (Nachtergaele et al., 2008).

The SJP, with 57 % of arable land area in 2009 (Ouyang et al., 2013), is one of the crucial food bases of China. In the Qixing Farm, paddy rice areas cover three quarters of the total sown areas (SBHR & HRSTCB, 2011).

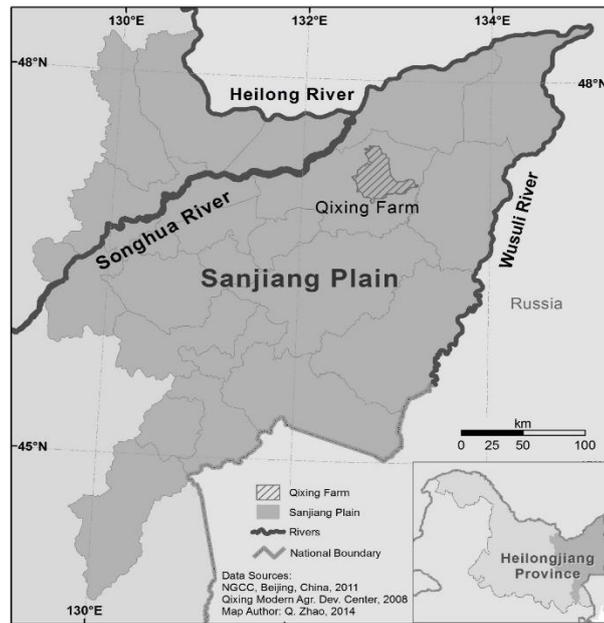


Figure 6-1: Location of the study area Qixing Farm in Heilongjiang province, NE-China.

6.2.2 Field data

Three field datasets spatially distributed in the Qixing Farm were collected in 2009 (Figure 6-2).

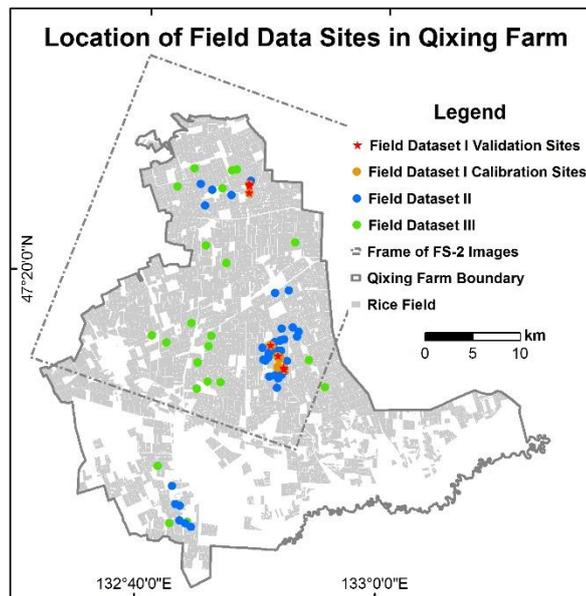


Figure 6-2: Distribution of field data sites in 2009 in the Qixing Farm.

Field dataset I was measured at 42 sites with each representing one field plot (0.5 – 3 ha). For each site, detailed crop parameters were measured, including leaf biomass, stem biomass and panicle weight. More information on data collection was given in Zhao et al. (2015a). Yield was measured by collecting grain samples at harvest time. The grain sample for each site consists of three or more sub-samples collected from 1 m². All samples were processed by grain threshing, drying and weighting. Specially, a constant water content of 14.5 % was assumed in final yield for all three field datasets. In total 33 sites from field dataset I were used to calibrate the model (i.e., maximum grain production, biomass fraction of leaf, stem and grain) and the data collected from the other nine sites of unique rice cultivar were used for site model validation. For the satellite RS approach, 28 cloud-free sites covered by the FS-2 images were used to construct the yield estimation model.

Field dataset II was acquired through a farmer survey. It includes 35 sites with each representing one farmer's field (20 – 40 ha). Yield was recorded by farmers and generally calculated as the total grain weight to the total farmer's field areas that may include field ridges, water pools, channels and even constructional areas. All 35 sites were used for regional model validation and 29 cloud-free sites covered by the FS-2 images were used for the RS approach validation.

Field dataset III consists of 22 sites with each site representing one field plot (0.5 – 3 ha). Field dataset III was collected from 'high yield' farmers according to their yield records (i.e., 3 – 10 years). Yield was measured similar to field dataset I. All 22 sites were used for the regional model validation while 18 cloud-free sites covered by the FS-2 images were used to validate the RS-derived yield.

6.2.3 The DNDC agro-ecosystem model

6.2.3.1 Model Description

The DNDC is a process-based geochemical agro-ecosystem model, which is driven by ecological factors (e.g., climate, soil, vegetation, field management) at daily or sub-daily time steps (Li et al., 1992; 1994). It was originally developed to estimate greenhouse gas emissions from agro-ecosystems. Latter, it was revised to simulate the anaerobic biogeochemistry of paddy fields (Li et al., 2004). DNDC consists of six sub-models: soil-climate, plant growth, decomposition, nitrification, denitrification, and fermentation. Crop growth is estimated using a crop growth curve that is generated as a function of elapsed fraction of the growing season (Li et al., 1994). Certain revised versions were developed in which detailed crop

algorithms associated with more biochemical soil conditions are integrated and more complex crop input parameters are needed (Zhang et al., 2002; Fumoto et al., 2008). The DNDC model has been widely applied in investigating greenhouse gas emissions and soil carbon fates (e.g. Li et al., 2000, 2004; Zhang et al., 2002), whereas only a few studies with coarse spatial resolution were reported to address rice growth (e.g., Zhang et al., 2011).

The site-specific mode of DNDC processes model input data by generating a transferable and changeable file which includes more parameters relative to its regional mode. For instance, the regional mode lacks options for parameters of water management and way and amount of nitrogen (N) fertilizer input which are critical in simulating crop yield.

6.2.3.2 Technical generalization for regional modelling

The site-specific mode of DNDC was generalized on a regional scale to take the advantages of better transferability and flexibility. Using ESRI ArcGIS 10.3 and Python, two scripts were developed to automatically assimilate raster files as model input data and generate raster files as model output data, respectively. All raster files have to be co-registered to pixel level as the site-specific model runs independently for each pixel.

Specifically, one script derives pixel values at the same coordinates from all input raster files (e.g., clay fraction raster, SOC content raster, etc.) and assigns these values to the model input parameters. The other script loops through all output files in a batch mode according to the pixel coordinates. For each output parameter (e.g. crop carbon, nitrous oxide flux, methane emission, etc.), a raster was generated by assigning its pixels with the recorded values of the corresponding coordinates.

6.2.3.3 Model input data

6.3.2.3.1 Meteorological data

Spatial heterogeneity of meteorological data is important for yield variability on large scales. Considering data availability and the geographical extent of the study area, meteorological conditions were assumed to be spatially uniform for the entire Qixing Farm. Meteorological data for the year of 2009 were provided by a meteorological station in the Qixing Farm. Model input data comprise daily maximum and minimum temperature, precipitation, wind speed and humidity.

6.3.2.3.2 Soil data

Soil input data are critical in crop yield estimation, but they are the most difficult data to obtain, especially with detailed spatial resolutions (Bareth and Yu, 2004). The DNDC model

provides a soil library, in which 14 sets of default soil properties (e.g., clay fraction, field capacity, etc.) are defined regarding to soil textures. However, soils classified in a same soil texture may maintain different soil properties such as soil particle constitution. Additionally, properties such as hydraulic parameters are even affected by SOC contents (Saxton and Rawls, 2006).

During the Second National Soil Survey in 1980s, soils in China were sampled, analyzed, mapped into different scales, and afterwards digitized. For this study, a digital soil type map with a fine scale of 1: 100,000 in raster format with a spatial resolution of 100 m was provided by the Chinese Academy of Agricultural Sciences in 2012. The soil type was classified and named in the Chinese soil taxonomy. This soil type raster file (including six soil types for the Qixing Farm) was subsequently used as a spatial reference for other model input raster files. Measured point data of soil pH and soil organic matter (SOM) were provided by the Qixing Farm Modern Agriculture Research Centre. In total 1156 soil samples of the top soil layer (0 – 20 cm) evenly distributed in the Qixing Farm were collected and measured in 2007 and 2008.

Nine model input parameters of soil properties were processed in raster format. They were SOC content (kg C/kg soil), soil pH, bulk density (BD, g/cm³), soil porosity (0 – 1), soil texture (0 – 1), clay fraction (0 – 1), field capacity (FC, 0 – 1; water-filled porosity at soil field capacity), water-filled porosity at soil wilting point (WP, 0 – 1) and saturated hydraulic conductivity (HC, m/hr). Detailed procedures for generating raster files of SOC, soil pH, BD, soil porosity, FC and WP were introduced in Zhao et al. (2015b).

Soil texture and clay fraction

Soils in the U.S. soil taxonomy are required by the DNDC model. In this study, the available soil data were transferred to the U.S. soil taxonomy using following steps. First, for each soil type, the constitution of soil particles with different sizes was analyzed according to the publications for the SJP (**Table 6-1**). A database of soil particle size grading curves was constructed. Second, the soil particle size grading curves were interpolated by the particle diameter metrics of the U.S. soil taxonomy (i.e., 0.002 mm, 0.02 mm) using cubic spline interpolation method in Matlab (Cai et al., 2003). Third, according to the particle diameter metrics of the U.S. soil taxonomy, soil texture and clay fraction were calculated for each of the soil types. Finally, by merging all soil type data, the soil texture and clay fraction rasters for the Qixing Farm were generated (**Figure 6-3**).

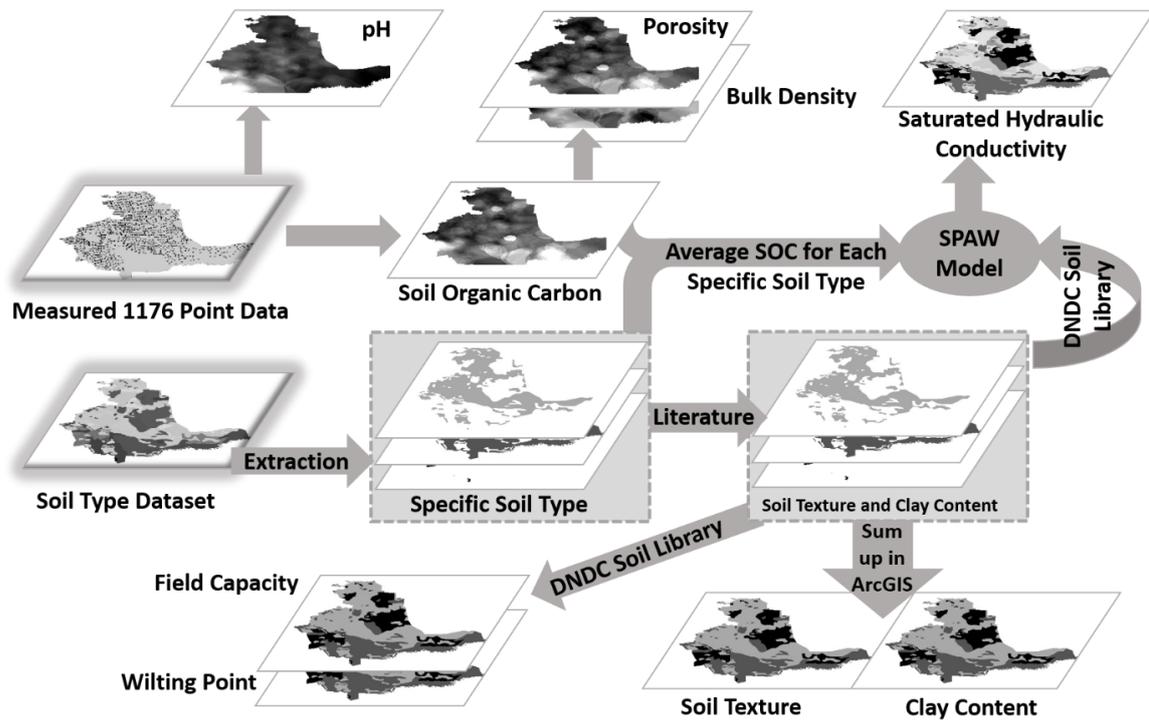


Figure 6-3: Workflow of deriving soil input raster files based on the measured point data and soil type raster file.

Saturated hydraulic conductivity

HC for each soil type was calculated in the ‘Soil Water Characteristics’ mode of the Soil-Plant-Air-Water (SPAW) model (Saxton and Rawls, 2006) (Figure 6-3). As required by the SPAW model, soil texture and mean value of SOM for each soil type was used to estimate HC (Table 6-1). The final HC raster file was generated in ArcGIS 10.3 with a pixel size of 100 m and was lined up to the soil type data

Table 6-1: Transferred soil textures and selected soil physical properties in the study area.

Soil type in Genetic Soil Classification of China Soil	Reference for soil particle constitutions	Soil texture in U.S. soil taxonomy	Soil Class No. in DNDC	Clay (%)	Silt (%)	Sand (%)	Saturated HC (m/hr)	Mean value of SOM (%)
Gleyed albic soil	Liu et al., 2012	Silt loam	4	16.5	78.8	4.7	0.0231	4.23
Meadow albic soil	Liu et al., 2012	Silt loam	4	18.9	75.5	5.6	0.0203	4.11
Typical albic soil	Zhang and Zhang, 1988	Silt loam	4	23.7	64.1	12.2	0.0165	4.12
Black soil	Wang et al., 2011; Wang et al., 2002; Ma et al., 2004,	Silty clay loam	7	30.4	55.5	14.1	0.0117	3.95
Bog soil	Zhang, 1981	Silty clay loam	7	34	55.2	10.8	0.0124	4.69
Meadow soil	Huo and Liu, 1985	Silty clay	10	47.6	40.7	11.7	0.008	4.59

6.3.2.3.3 Field management data

Field management is a critical factor in agro-ecosystem modelling applications (Kersebaum et al., 2007). Crop-specific field management depends on crop types that can be detected from land use classifications (Waldhoff and Bareth, 2009). In this study, a high accuracy rice map was used. It was derived from multiple FS-2 images and GIS topographic data (Zhao et al., 2015a).

Detailed rice management data regarding tillage, irrigation, transplanting and fertilizer application and harvest date were recorded for each site of field dataset I. The field management parameters for DNDC were adjusted using the 33 sites data from field dataset I (**Table 6-2**). Specially, N fertilizer was set as 85 kg N /ha according to a farmer survey (data not shown).

6.3.2.3.4 Crop-specific input parameters

Crop-specific input parameters were optimized using data of the 33 sites from field dataset I. For the regional application, a set of crop-specific parameters was adjusted in which characteristics of the major rice cultivars (i.e., Longjing 21, Kongyu 131) were considered (**Table 6-2**).

Table 6-2: DNDC model input parameters (selected) for regional application.

Model input parameter		value
Rice management	Tillage date / depth (cm)	Apr. 30 th / 30
	Transplanting date	May 20 th
	Times of N application per growing season	4
	Total N input (kg N/ha)	85
	Harvest date	Sep. 20 th
Biomass production (kg C/ha)	maximum grain (kg C/ha)	3100
	leaf (kg C/ha)	913
	stem (kg C/ha)	1860
	root (kg C/ha)	652
Biomass fraction (%)	grain	47.5
	leaf	14
	stem	28.5
	root*	10
Crop-specific parameter	grain	55
	leaf	60
	stem	60
	root	70
	Accumulated temperature for maturity	2300
	Water demand (g water/g dry matter) #	508
	N fixation index (crop N/N from soil) *	1
	Vascularity index (0-1) #	1
	Optimum temperature (degree C) #	22

*Referring to Zhang et al. (2011); #model default value.

The DNDC model performance is significantly affected by cultivar-specific properties, such as accumulated temperature, biomass fractions, and actual field management strategies including transplanting date and N input (Zhang et al., 2002; Fumoto et al., 2008). Therefore in the site validation, cultivar-specific parameters (i.e., accumulated temperature, biomass fraction) and actual field management data from the nine validation sites were used (**Table 6-4**).

6.2.4 Remote sensing approach

Multi-temporal vegetation indices have been used to estimate crop yields (Bolton and Friedl, 2013). In this study, multi-temporal NDVIs (normalized difference vegetation index) derived from three FS-2 images were used to estimate rice yield. The FS-2 images were captured on June 24th, July 6th and August 9th of 2009. Each image covers the main arable land area (ca. 56 000 ha) of the Qixing Farm (**Figure 6-1**).

The FS-2 satellite collects multispectral images with four bands of blue (450–520 nm), green (520–600 nm), red (630–690 nm), and near-infrared (760–900 nm). Image atmospheric correction and geometric correction were described in Zhao et al. (2015a). The NDVI for each image was calculated as:

$$NDVI = \frac{R_{NIR} - R_{red}}{R_{NIR} + R_{red}} \quad (6 - 1)$$

where R_{NIR} and R_{red} represent reflectance at the near-infrared and the red band, respectively.

Rice yield was calculated through an empirical model which was constructed in SPSS 21 (SPSS, Inc., Chicago, IL, USA) using multiple linear regression (MLR). The NDVIs of the three dates were used as descriptive factors in the MLR (**Table 6-3**).

Table 6-3: Parameters for the multiple linear regression model for the RS approach.

Para- meters	Coefficient of descriptive variables				MLR method	Linear R ²	total degree of freedom	Sig. level
	NDVI of June 24	NDVI of July 6	NDVI of August 9	Constant				
Value	-2 319.254	2 103.761	32874.675	-18 785.41	Enter	0.474	28	$p < 0.001$

6.2.5 Statistical analysis

Statistical measures including the coefficient of model efficiency (EF), the coefficient of model determination (CD), and the index of agreement (IA) were used to test the goodness-of-fit of the DNDC model and the regression model based on the satellite RS data. They were calculated as below.

$$EF = \left(\sum_{i=1}^n (O_i - \bar{O})^2 - \sum_{i=1}^n (P_i - O_i)^2 \right) / \sum_{i=1}^n (O_i - \bar{O})^2 \quad (6 - 2)$$

$$CD = \sum_{i=1}^n (O_i - \bar{O})^2 / \sum_{i=1}^n (P_i - \bar{O})^2 \quad (6 - 3)$$

$$IA = 1.0 - \sum_{i=1}^n (P_i - O_i)^2 / \sum_{i=1}^n (|P_i - \bar{O}| + |O_i - \bar{O}|)^2 \quad (6 - 4)$$

In all equations, P_i and O_i denote the model predicted (P_i) and the observed (O_i) value; \bar{O} is the mean of the observed value; n is the total number of observations. The value for EF is

up to 1 and a positive value indicates that the trend in the measured data was better described by the modelled values than the mean of the observed values. The CD value is larger than or equal to 0. A CD value of 1 or above indicates that the model describes the measured data better than the mean of the observations. The IA, ranging from 0.0 to 1.0, represents the degree of agreement between model estimations and observed values (Willmott, 1981).

6.3 Results

6.3.1 Site-specific model application and site validation

The DNDC model performed well for site validation based on the site-specific management data and the cultivar-specific crop parameters (Table 6-4).

Table 6-4: Site-specific input data and performance of DNDC for site validation.

Cultivar	T.P. date (mm/dd)	Accumulated temperature (°C)	Biomass fraction			N input (kg N / ha)	Yield (kg/ha)			Assessment index *			
			Grain	Leaf	Stem		Field observed	DNDC modelled	Yield Difference	R ²	EF	CD	IA
Longjing 21	5/10	2500 (Chen et al., 2014)	0.463	0.144	0.293	104	7824	8094	270				
Xixuan 1	5/15	2600 (HAAS, 2015)	0.433	0.168	0.299	132	8510	8503	-7				
Longdun 104	5/16	2550 (Niu and Shi, 2013)	0.42	0.167	0.313	118	6940	7942	1002				
Kendao 6	5/20	2300 (Yu, 2003)	0.522	0.14	0.238	83	8298	8249	-49				
Longjing 24	5/16	2300 (Wang, 2009)	0.478	0.155	0.267	95	7985	8032	47	0.816	0.676	2.61	0.867
Longdun 249	5/20	2550 (HPBLM, 2015a)	0.436	0.163	0.301	102	8881	9015	134				
Chaoyou 949	5/18	2300 (HPBLM, 2015b)	0.39	0.197	0.313	106	8739	8623	-116				
Jinxuan 1	5/20	2300 (Hu and Liu, 2013)	0.469	0.151	0.28	86	9252	9000	-252				
Kongyu 131	5/17	2300 (Qu et al., 2001)	0.517	0.133	0.25	82	8423	8532	109				

* $n = 9$

The site-specific DNDC model was successfully generalized to the regional scale of the entire Qixing Farm. Spatial variability was well detected over the entire Farm in the yield map (Figure 6-4). High yields appeared in the southern part while lower yields clustered in the mid-eastern part of the Qixing Farm. Even a ‘within-field’ variability was detected.

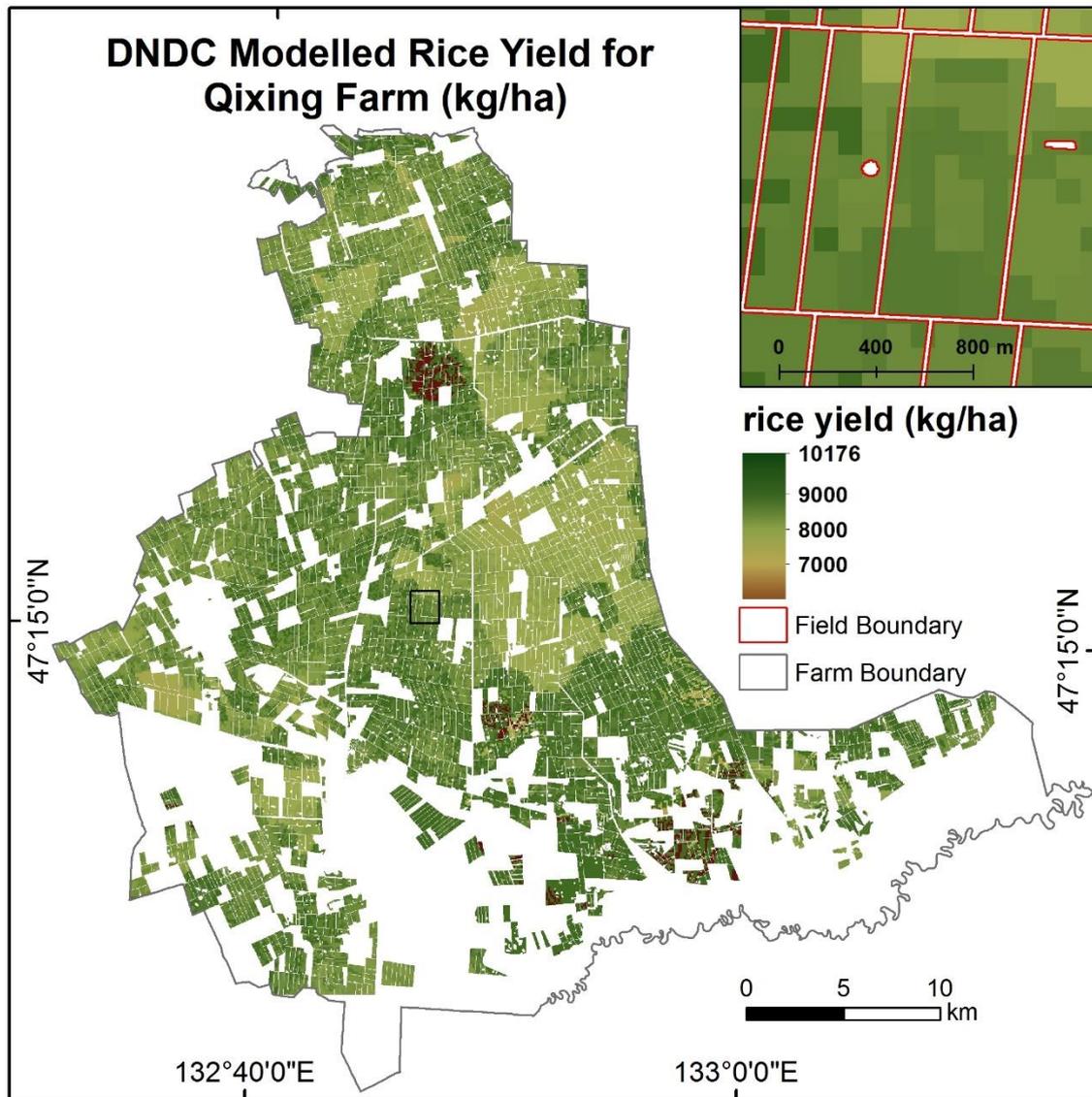


Figure 6-4: Map of rice yield modelled by the DNDC model and the within-field variability in the inset map.

In the regional validation results, satisfactory values were obtained for all assessment indices from both field dataset II and III (Table 6-5).

Table 6-5: Assessment indices in model regional validation.

Dataset	Sites	Linear R ²	EF	CD	IA
Field dataset II	35	0.498**	0.053	1.59	0.667
Field dataset III	22	0.411*	0.404	2.63	0.74

Significance test: ** $p < 0.005$; * $p < 0.05$.

Figure 6-5 shows the DNDC modelled and the observed yields at all 57 validation sites from both field dataset II and III. The modelled yields and the field observed yields were comparable at most of the validation sites. The model overestimated rice yield at low yield (< 7500 kg/ha) sites, which may be attributed to the overestimation of rice cultivation area in the yield calculation algorithm of field dataset II (see in section 2.2), whereas, the model under-estimated rice yield slightly at high yield (> 9000 kg/ha) sites.

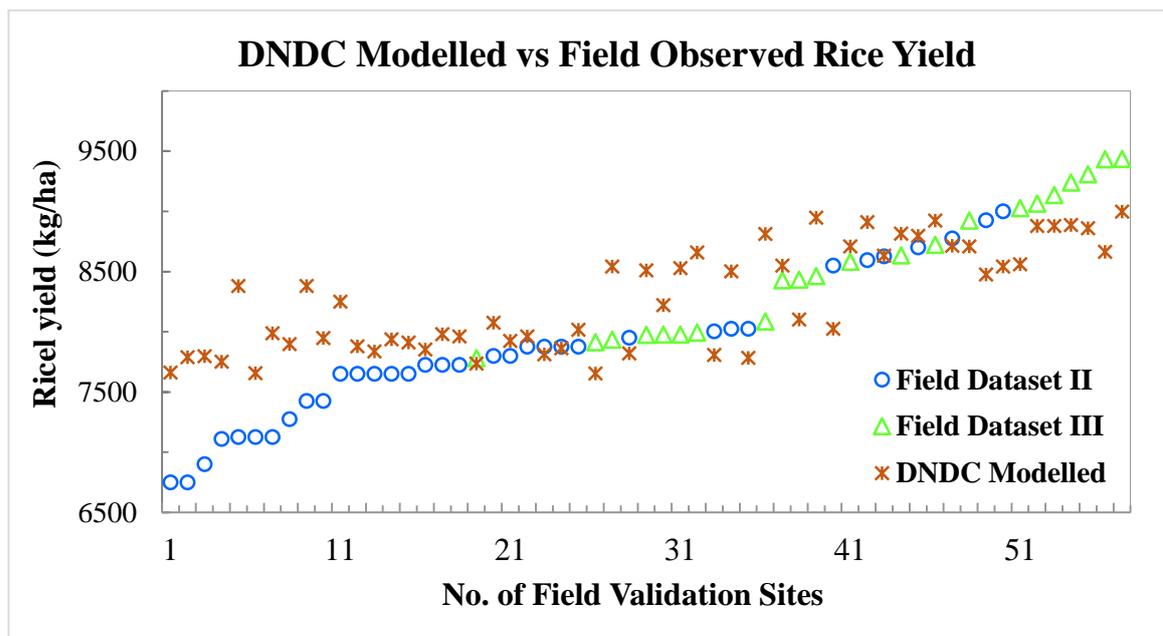


Figure 6-5: DNDC modelled and field observed rice yields at all 57 validation sites.

6.3.2 Soil-specific validation

With respect to soil types, the best validation result of modelled yield was found in meadow albic soil (Table 6-6). Specifically, validation results vary in silt loam soils, attributing to the detailed raster input files rather than the default values from the DNDC soil library.

Table 6-6: Assessment indices for soil-specific validation.

Dataset	Sites	Soil type	Linear R ²	EF	CD	IA
Field dataset II	6	meadow albic	0.515	0.079	2.04	0.622
	24	meadow	0.284	-0.707	1.06	0.497
	9	all silt loam soils	0.512	0.018	1.88	0.63
Field dataset III	8	gleyed albic	0.294	0.233	7.66	0.52
	7	meadow albic	0.873	0.413	11.68	0.633
	4	typical albic	0.08	-0.302	1.29	0.603
	19	all silt loam soils	0.347	0.302	4.60	0.628

6.3.3 Validation of RS-derived rice yield

For the validation of the RS approach, the CD value was 1.25, higher than 1, indicating the RS-derived results described the measured data better than the mean of the observations (Table 6-7). Field dataset III showed a better validation results than field dataset II, with CD and IA values of 1.42 and 0.55, respectively.

Table 6-7: Assessment indices for RS-derived yield.

Dataset	Sites	Linear R ²	EF	CD	IA
Field dataset II	29	0.051	-1.161	1.255	0.222
Field dataset III	18	0.111	-0.194	1.417	0.549

6.3.4 Comparison of modelled and RS-derived yields

Yield difference was calculated by subtracting the RS-derived value from the DNDC modelled value. A ratio map was generated to show the percentage of yield difference to the modelled yield (Figure 6-6). In an overview, the modelled and RS-derived yields are comparable. The ratio ranged from -15 - 15 % in most areas. Areas with similar ratio classes were rather clustered than randomly distributed, which is reasonable because actual yield pattern may tend to be clustered due to spatial patterns of soil conditions, as well as field-specific cultivars and management strategies. There were more positive areas than negative areas, which indicated the modelled yields were more frequently higher relative to the RS-derived yields. Specially, positive ratio areas occurred more at the field edges.

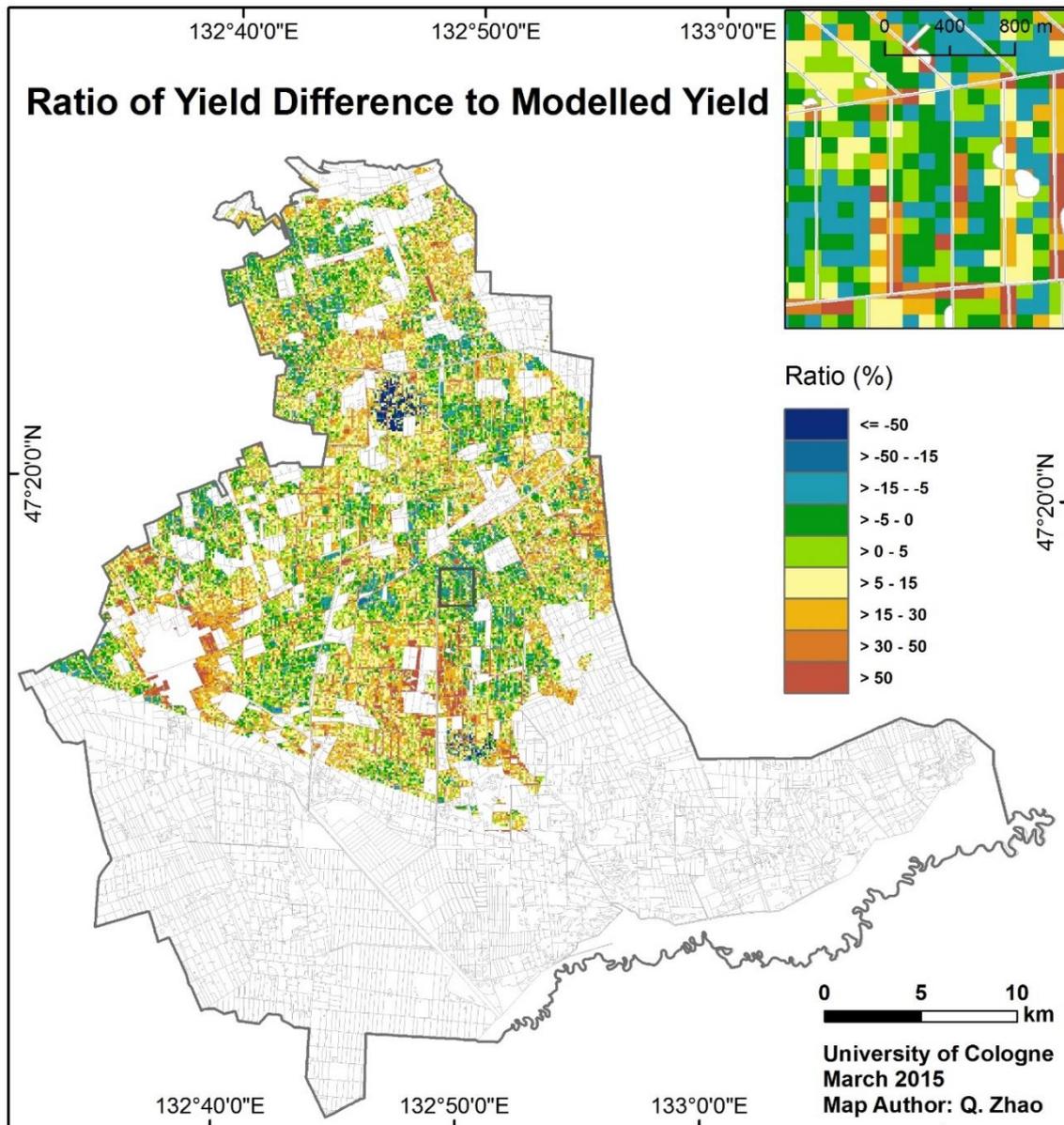


Figure 6-6: Ratio map of yield difference (subtracting the RS-derived yield from the modelled yield) to the DNDC modelled yield.

6.4 Discussion

6.4.1 Model regionalization

Site-specific models are insufficient in capturing and representing the spatial variability of the model concerns (e.g., meteorological conditions, soil properties, plant growth factors, land management practices) across landscape (Resop et al., 2012; Thorp and Bronson, 2013).

Although the DNDC model consists a regional mode, its simulation units are basically administrative areas or soil polygons which are generally determined by the geographical scales of the available data (e.g., Zhang et al., 2014). Thus, parallel analyses of model results from different study cases are limited.

This study generalized the site-specific DNDC model using additional scripts to deal with raster files. Since the raster pixels are used as the model simulation units and the pixel resolution is user-defined, it is efficient and flexible to solve practical problems. The pixel resolution can be defined regarding not only the scales of the available spatial data but also the study concerns. For instance, the pixel resolution/simulation units can be defined according to field management information. Additionally, such model regionalization facilitates synthetic analysis on crop information modelled by the DNDC and derived from remote sensing imagery.

6.4.2 Soil effects on DNDC modelled rice yield

The modelled yield highly depends on the soil properties (**Figure 6-7: a-c, e, f**). In this study, higher modelled yields occur in silt loam and silty clay loam soils with lower clay fractions and higher silt contents (**Figure 6-7: a-c, Table 6-1**). These results are consistent with R uth and Lennartz (2008), who found that the rice yield was positively related to the silt content because such soils provide better soil nutrition-holding capacity and proper nutrition status for crops. Spatial pattern of modelled yield was positively consistent with HC and SOC patterns (**Figure 6-7: a, e, f**).

The high yield sites of field dataset III are mainly located in silt loam soils (**Figure 6-7: b**), which may greatly be attributed to the factor of field management. Management measures have greater effects on rice yield in sandy soils than in clayed soils because the high nutrient content in heavy clayed soils guarantees high yields (R uth and Lennartz, 2008). The potential of field management effects on rice yield in different soils needs to be further studied.

There was no significant consistency between the spatial pattern of RS-derived yield and soil properties (**Figure 6-7: b-f**). This is reasonable because the RS-derived yield is a synthesized response from all environmental factors (e.g., climate, soil) and human activities rather than a single soil factor. Nevertheless, the RS-derived yield was lower in the mid-east part of the image covered areas (**Figure 6-7: d**) where the soil condition was composed of silty clay soil texture, high clay content, low HC, and low SOC content..

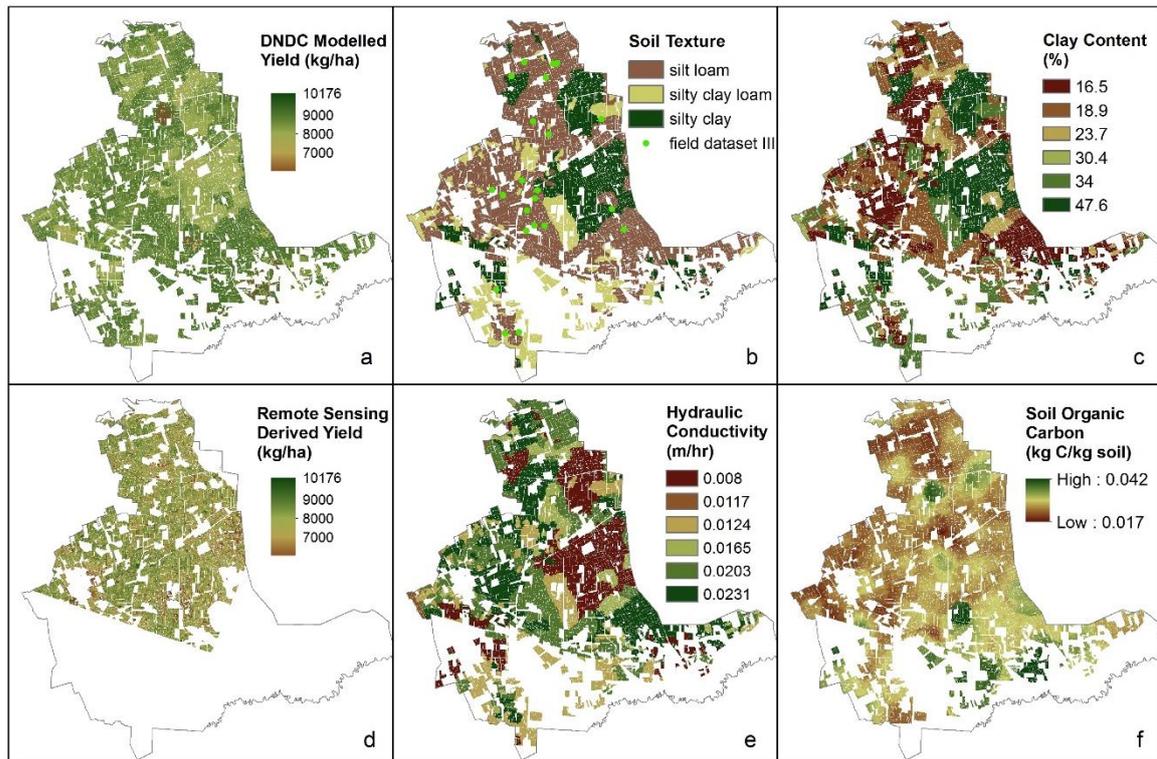


Figure 6-7: Spatial patterns of estimated rice yield maps and soil data.

6.4.3 Sources of uncertainty for regional model application

6.4.3.1 Field management

Rice management such as transplanting date and N fertilizer input varies in different paddy fields (Table 6-4). However, in the regional application, only one single field management case was implemented (Table 6-2), which was not able to reflect the true status for the whole area.

Additionally, certain measures were demonstrated to be critical for rice under the ‘cold’ climate conditions in the SJP. For example, certain farmers will keep the depth of water layer to be more than 17 cm to avoid cold injury (temperature < 17 °C) during the panicle initiation stage according to practical weather conditions (Meng et al., 2005). Such flexible measures are not considered in the DNDC model application.

6.4.3.2 Crop-specific parameters

Crop-specific parameters vary with different rice cultivars (Table 6-4). In the regional application, errors may result from the single set of crop-specific parameters (Table 6-2). For instance, the under-estimations at the ‘high yield’ sites of field dataset III (Figure 6-5) may be due to the ‘improper’ settings of crop-specific parameters because at these sites special

cultivars are generally planted rather than the domain cultivars of Longjing 21 or Kongyu 131.

6.4.4 RS-derive rice yield

The regional validation for the RS approach showed lower assessment indices (**Table 6-7**) than the modelled results (**Table 6-5**). Limitations could exist in the empirical model construction. The MLR model was constructed based on only 28 sites which were clustered in two villages. Thus the spatial heterogeneity in rice status could not be well represented.

Nonetheless, compared with the process-based DNDC model, the empirical MLR model integrated most environmental and human effects that the RS imagery detected. For instance, the RS approach successfully detected the lower yields at the field edges (**Figure 6-6**). Such lower yields may result from resource competition of agroforestry system (Rüser and Hansen, 2014) since forest belts are generally planted at the field edges in the SJP, whereas such effects are out of the scope of the DNDC model.

In addition, satellite-RS can monitor rice status (i.e., biomass, LAI, etc.) effectively (Zhao et al., 2015a). It is promising to integrate RS-derived vegetation information into process-based agro-ecosystem models for regional applications.

6.5 Conclusions

The DNDC model performed well in estimating rice yield in the SJP, especially when cultivar-specific parameters and site-specific management data were specialized. Generalizing the site-specific model onto regional scales may assist regional studies for addressing agro-environmental issues. Detailed soil data and accurate land use maps improved the model performance in detecting spatial variabilities in rice yield. High rice yields were found in silt loam soils, while the reasons for this result need to be further studied.

The regional validation showed better yield results from the agro-ecosystem model than the RS approach. Nonetheless, the DNDC modelled and RS-derived yields were comparable. The RS-derived yield represents a synthesized response of the environmental factors and human activities; therefore, it is a complementary method to the process-based modelling approach. For example, it can detect the lower yields resulted from factors that are beyond the capability of the process-based modelling. Better results could be obtained by integrating the RS-derived information into the process-based agro-ecosystem models.

Acknowledgements

This work was funded by the Natural Science Foundation of China (NSFC, project No. 31071859) and the German Research Foundation (DFG, project no. BA 2062/8-1). The authors are thankful to all people from University of Cologne, China Agricultural University and Heilongjiang Land Reclamation who supported and assisted in this research.

Conflicts of Interest

The authors declare no conflict of interest.

References

- Atzberger, C., 2013. Advances in remote sensing of agriculture: Context description, existing operational monitoring systems and major information needs. *Remote Sens.*, 5 (2), 949–981.
- Bareth, G., 2009. GIS- and RS-based spatial decision support: structure of a spatial environmental information system (SEIS). *International Journal of Digital Earth*, 2 (2), 134–154.
- Bareth, G., Yu, Z., 2004. Verfügbarkeit von digitalen Geodaten in China. *Petermanns Geographische Mitteilungen*, 148 (5), 78–85 (in German with English abstract).
- Bolton, D. K., Friedl, M. A., 2013. Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics. *Agr. Forest Meteorol.*, 173, 74–84.
- Cai, Y., Zhang, K., Li, S., 2003. Study of the conversion of different soils texture. *Acta Pedologica Sinica*. 40 (4), 511–517 (in Chinese with English abstract).
- Chen, S., Yang, L., Zhao, H., Xue, Q., Pan, G., 2014. Study on cultivation techniques with high yield of high-quality super rice Longjing 21 in cold region. *Heilongjiang Agricultural Sciences*, 2, 23–28 (in Chinese with English abstract).
- Ericksen, P. J., Ingram, J. S. I., Liverman, D. M., 2009. Food security and global environmental change: emerging challenges. *Environ. Sci. Policy*, 12 (4), 373–377.
- Fumoto, T., Kobayashi, K., Li, C. S., Yagi, K., Hasegawa, T., 2008. Revising a process based biogeochemistry model (DNDC) to simulate methane emission from rice paddy fields under various residue management and fertilizer regimes. *Glob Change Biol.*, 14, 382–402.
- Godfray, H. C. J., Beddington. J. R., Crute, I. R., Haddad, L., Lawrence, D., Muir, J. F., Pretty, J., Robinson, S., Thomas, S. M., Toulmin, C., 2010. Food security: the challenge of feeding 9 billion people. *Science*, 327 (5967), 812–818.
- HAAS (Heilongjiang Academy of Agricultural Sciences) 2015. Brief introduction to rice cultivars Xixuan 1. (<http://www.haas.cn/newsview.aspx?id=520>, accessed on August 31, 2015)
- Hansen, J. W., Jones, J. W., 2000. Scaling-up crop models for climate variability applications. *Agr. Syst.*, 65 (1), 43–72.

- HPBLM (Heilongjiang Province Breeding Limited Company), 2015a. Introduction to seeds products Longdun 249. (http://www.3456.tv/business/hljkyzy/pro_590919, accessed on August 31, 2015).
- HPBLM, 2015b. Introduction to seeds products Chaoyou 949. (http://www.3456.tv/business/jhznzw/pro_88942, accessed on August 31, 2015)
- Hu, J., Liu, H., 2013. Characteristics and key cultivation techniques of a new rice cultivar Jinxuan 1. *North Rice*, 1, 63–64 (in Chinese).
- Huang, S., Miao, Y., Zhao, G., Yuan, F., Ma, X., Tan, C., Yu, W., Gnyp, M.L., Lenz-Wiedemann, V.I.S., Rascher, U., Bareth, G., 2015. Satellite Remote Sensing-Based In-Season Diagnosis of Rice Nitrogen Status in Northeast China. *Remote Sens.*, 7 (8), 10646–10667.
- Huo, Y., Liu, X., 1985. Study on the physical characteristics of soil moisture in meadow soils in the Sanjiang Plain. *Journal of Northeast Agricultural College*, 3, 53–58 (in Chinese).
- Jones, J. W., 2014. Assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison. *PNAS*, 111 (9), 3268–3273.
- Kersebaum, K. C., Hecker, J. M., Mirschel, W., Wegehenkel, M., 2007. Modelling water and nutrient dynamics in soil–crop systems: a comparison of simulation models applied on common data sets, In: *Modelling water and nutrient dynamics in soil–crop systems*, Springer, Netherlands, pp. 1–17.
- Kuenzer, C., Knauer, K., 2010. Remote sensing of rice crop areas. *Int. J. Remote Sens.*, 34 (6), 2101–2139.
- Li, C., Aber, J., Stange, F., Butterbach-Bahl, K., Papen, H., 2000. A process - oriented model of N₂O and NO emissions from forest soils: 1. Model development. *Journal of Geophysical Research: Atmospheres*, 105 (D4), 4369–4384.
- Li, C., Cui, J., Sun, G., Trettin, C., Cui, J., 2004. Modeling impacts of management on carbon sequestration and trace gas emissions in forested wetland ecosystems. *Environ. Manage.*, 33 (1), S176–S186.

- Li, C., Frolking, S., Frolking, T. A., 1992. A model of nitrous oxide evolution from soil driven by rainfall events: 1. Model structure and sensitivity. *J. Geophys. Res-Atmos.*, 97 (D9), 9759–9776.
- Li, C., Frolking, S., Harriss, R., 1994. Modeling carbon biogeochemistry in agricultural soils. *Global biogeochem. Cy.*, 8 (3), 237–254.
- Liaghat, S., Balasundram, S. K., 2010. A review: The role of remote sensing in precision agriculture. *American journal of agricultural and biological sciences*, 5(1), 50–55.
- Liu, T., Gao, Z., Kuang, E., Liu, F., 2012. Study on Physical Characteristics of Planosol I. Some Physical Characteristics about Soil Compaction and Permeability. *Heilongjiang Agricultural Sciences*, 4, 49–52 (in Chinese with English abstract).
- Ma, Q., Yu, W., Zhao, S., Zhang, L., Shen, S., Wang, Y., 2004. Comprehensive evaluation of cultivated black soil fertility. *Chinese Journal of Applied Ecology*, 15 (10), 1916–1920 (in Chinese with English abstract).
- Meng, Z., Wang, Y., Sun, Z., Meng, Q., Liu, Y., Li, C., Zhang, J., 2005. The research of rice sterile-type chilling injury and irrigation technique. *Chinese Agricultural Science Bulletin*, 21 (6), 197–201 (in Chinese with English abstract).
- Nachtergaele, F., van Velthuisen, H., Verelst, L., 2008. Harmonized World Soil Data-base, Food and Agriculture Organization of the United Nations.
- Niu, S., Shi, X., 2005. Key techniques of breeding and cultivation for a new rice cultivar Longdun 104. *Reclaiming and Rice Cultivation*, 5, 11–12 (in Chinese).
- Ouyang, W., Xu, Y., Hao, F., 2013. Effect of long-term agricultural cultivation and land use conversion on soil nutrient contents in the Sanjiang Plain. *Catena*, 104, 243–250.
- Perlman, J., Hijmans, R. J., Horwath, W. R., 2013. Modelling agricultural nitrous oxide emissions for large regions. *Environ. Modell. Softw.*, 48, 183–192.
- Qu, G., Zhao, Y., Zhao, Y., Huo, X., 2001. Main characteristics and key cultivation techniques of the high quality rice cultivar Kongyu 131 in the cold areas. *China Rice*, 7 (2), 18 (in Chinese).
- Resop, J. P., Fleisher, D. H., Wang, Q., Timlin, D. J., Reddy, V. R., 2012. Combining explanatory crop models with geospatial data for regional analyses of crop yield using field-scale modeling units. *Comput. Electron. Agr.*, 89, 51–61.

- Riiser, N. M. and Hansen, T. R., 2014. The Favorability of Rice-Agroforestry-A Meta-Analysis on Yield and Soil Parameters. Master Thesis, Roskilde University, Denmark.
- Rüth, B., Lennartz, B., 2008. Spatial variability of soil properties and rice yield along two catenas in southeast China. *Pedosphere*, 18 (4), 409–420.
- Saxton, K. E., Rawls, W. J., 2006. Soil water characteristic estimates by texture and organic matter for hydrologic solutions. *Soil Sci. Soc. Am. J.*, 70 (5), 1569–1578.
- SBHR & HRSTCB (Statistics Bureau of Heilongjiang Reclamation and Heilongjiang Reclamation Survey Team of China Bureau), 2011. Leapfrog Development of Heilongjiang Reclamation. China Statistics Press.
- Thorp, K. R., Bronson, K. F., 2013. A model-independent open-source geospatial tool for managing point-based environmental model simulations at multiple spatial locations. *Environ. Modell. Soft.*, 50, 25–36.
- Tilman, D., Cassman, K. G., Matson, P. A., Naylor, R., Polasky, S., 2002. Agricultural sustainability and intensive production practices. *Nature*, 418 (6898), 671–677.
- Waldhoff, G., Bareth, G., 2009. GIS-and RS-based land use and land cover analysis: Case study Rur-Watershed, Germany. In *Geoinformatics 2008 and Joint Conference on GIS and Built environment: Advanced Spatial Data Models and Analyses* (pp. 714626–714626). International Society for Optics and Photonics.
- Wang, J., 2009. Characteristics and key cultivation techniques of an early mature rice cultivar Longjing 24. *China Rice*, 6, 64 (in Chinese).
- Wang, J., Wang, T., Zhang, X., Guang, L., Wang, Q., Hu, H., Zhao, Y., 2002. An approach to the changes of black soil quality (I) – Changes of the indices of black soil with the year(s) of reclamation. *Journal of Shenyang Agricultural University*, 33 (1), 43–47 (in Chinese with English abstract).
- Wang, X., Wang, Z., Song, W., Zhang, G., Zhang, W., 2011. Relationship between depth of humus horizon and particle size composition in tillage horizon of the typical black soil in Northeastern China. *Science of Soil and Water Conservation*, 8 (6), 39–44 (in Chinese with English abstract).
- Willmott, C. J., 1981. On the validation of models. *Phys. Geogr.* 2, 184–194.

- Yan, M., Deng, W., Ma, X. 2001. Climate variation in the Sanjiang Plain disturbed by large scale reclamation during the last 45 years. *Acta Geographica Sinica*, 56 (2), 159–170 (in Chinese with English abstract).
- Yu, H., 2003. Brief introduction to high quality, high yield rice cultivar. *Reclaiming and Cultivation*, 2, 004 (in Chinese).
- Zhang, Y., 1981. The genesis, nature and classification of the marshy soil of Sanjiang Plain. *Scientia Geographica Sinica*, 1 (2), 171–180 (in Chinese with English abstract).
- Zhang, Y., Li, C., Zhou, X., Moore III, B., 2002. A simulation model linking crop growth and soil biogeochemistry for sustainable agriculture. *Ecol. Model.*, 151 (1), 75–108.
- Zhang, Y., Wang, Y., Su, S., Li, C., 2011. Quantifying methane emissions from rice paddies in Northeast China by integrating remote sensing mapping with a biogeochemical model. *Biogeosciences*, 8 (5), 1225–1235.
- Zhang, Z., Zhang, Y., 1988. Compared study on soil particle constituent and microaggregate of albisol, black soil and meadow soil. *Journal of Heilongjiang August First Land Reclamation University*, 2, 9–20 (in Chinese with English abstract).
- Zhao, Q., Brocks, S., Lenz-Wiedemann, V.I.S., Miao, Y., Bareth, G., 2015b. Regional application of the site-specific biochemical process-based crop model DNDC for rice in NE-China. The 35th Annual Conference of the DGPF, March 16th – 18th, 2015, Cologne, Germany.
- Zhao, Q., Lenz-Wiedemann, V. I. S., Yuan, F., Jiang, R., Miao, Y., Zhang, F., Bareth, G., 2015a. Investigating within-field variability of rice from high resolution satellite imagery in Qixing Farm County, Northeast China. *ISPRS Int. J. Geo-Inf.*, 4 (1), 236–261.

7 Discussion

This study is a problem-oriented interdisciplinary research on agriculture and its impacts on environment. The study was conducted following the logic pyramid of ‘data-information-knowledge-understanding-decision’ system (Rowley, 2007). Several problems in the AEIS are discussed in this section.

As mentioned in **chapter 3**, setting up an AEIS is an interdisciplinary study which is based on multi-source data. The extreme heterogeneity of data with highly varying characteristics in formats, resolutions, usage constraints, information gaps, metadata inconsistencies and semantic issues has to be considered (Fegraus et al., 2012). Thus, organization of multi-source data have to be emphasized. As a key pre-processing procedure for geospatial data, the innovative method of georeferencing using TerreSAR-X imagery as ground control information source is discussed.

To derive information from the multi-source data, the knowledge-based expert classifier was applied. By combining GIS and RS data using the expert classifier, rice areas were delineated with high accuracy (**chapter 5**). Based on this accurate rice area map, FS-2 imagery was used to monitor rice status on a within-field level. Although satisfy results were obtained, disadvantages still exist due to the characteristics of satellite imagery and the natural conditions of the SJP, which are also discussed in this section.

The DNDC was selected in this study and its site-specific mode has been generalized on the regional scale to estimate rice yield (**chapter 6**). Rice yield was also derived from FS-2 images and compared with the modelling results (**chapter 6**). Furthermore, interactions between rice production and soil properties were addressed by the DNDC results and thus the knowledge of crop growth and environmental factors was better understood (**chapter 6**). The performance of the agro-ecosystem model DNDC on rice biomass simulation is discussed in this section.

7.1 Multi-source data analysis

Multi-source data are versatile and applicable in diverse areas (Li, 2010) and they are highly needed in multi-disciplinary research. However, the multi-source data are characterized by their heterogeneous data formats (e.g., GIS topographic data, satellite remote sensing, household survey data, soil surveys, vegetation, field observations, and land cover), information gaps from different data providers, inconsistency in data precision, etc. (Bareth, 2009; Fegraus et al., 2012; Khaleghi et al., 2013). Therefore, dealing with multi-source data

for the multi-disciplinary research is mandatory. There are quite a number of challenging issues in the ‘pre-processing’ of multi-source data (Li, 2010; Zhang, 2010).

In this study, the high resolution radar imagery was used to georeference the multi-source data to take advantage of its high spatial accuracy (Zhao et al., 2015a). Geospatial data of both topographic vector data and satellite raster data can be well georeferenced. However, the expense of radar imagery and the spatial coverage of high resolution radar imagery could be a constraint in research projects. In this study, the ground control points was manually selected from the TerraSAR-X imagery (**chapter 4**), which obstructs the application of this approach especially for a larger area with a big number of datasets needed. Automatic methods have been proposed for selecting ground control points from satellite images (e.g., Chen et al., 2008), or for single ground features in vector format (e.g., Song et al., 2009). However, very few automatic methods are reported for georeferencing simultaneously both topographic vector data and raster data using ancillary data. Using TerraSAR-X imagery as ground control information in multi-source data georeferencing is promising, especially in the areas where necessary ground control information is lack.

An efficient tool to ‘pre-process’ the multi-source data is the KBS, which categories and integrates multi-source data based on the human’s preliminary understanding of the information. Many studies found that the KBS is effective in multiple satellite imagery classifications (e.g., Cohen and Shoshany, 2005; Ban et al., 2010; Mwaniki and Möller, 2015). However, few research has been conducted using the KBS to integrate multi-source geospatial data in both formats of vector and raster.

The knowledge-based classification is demonstrated to be an efficient method to derive information from RS and GIS topographic data. The KBS utilized information from GIS topographic data into RS-derived classes for area detection, based on creating rules in the expert classifier. Specially, the multi-source data have better ability in distinguishing rice areas from dryland areas, as shown in the following tables (**Table 7-1, 7-2, 7-3**). This is because the different crops can be better detected by multi-temporal satellite observations due to their different phenological characteristics over through the growing season (Zhong et al, 2014). The best results from GIS and multiple satellite data in **Table 7-3** are additionally attribute to the in-situ surveyed GIS data.

Table 7-1: Confusion matrix of the land cover classification map based on a single date (August 9th) RS image.

		Reference					
		Rice	Dryland	Forest	Residential areas	Other	Total
Classification of single-date RS data	Rice	319	7	10		3	339
	Dryland	44	112	2	8	3	169
	Forest	14	16	74		3	107
	Residential areas	7	8	7	65		87
	Other	2	13	6	1	76	98
	Total	386	156	99	74	85	800

Table 7-2: Confusion matrix of the land cover classification map that was generated by integrating the three classifications from the three dates.

		Reference					
		Rice	Dryland	Forest	Residential areas	Other	Total
Classification of multiple-date RS data	Rice	354	13	19		10	396
	Dryland	15	115	1	5	1	137
	Forest	10	11	74		3	98
	Residential areas	7	8	4	68	2	89
	Other		9	1	1	69	80
	Total	386	156	99	74	85	800

Table 7-3: Confusion matrix of the land cover classification map based on both the RS and GIS data.

		Reference					
		Rice	Dryland	Forest	Residential areas	Other	Total
Classification of RS and GIS data	Rice	358	13	8		1	380
	Dryland	8	135	1	2	1	147
	Forest	13	4	85			102
	Residential areas	7	1	4	72		84
	Other		3	1		83	87
	Total	386	156	99	74	85	800

Specifically, the implementation of the KBS classification increases the producers' accuracies and decreases the omission errors (**Table 7-4**), indicating a better ability in detecting the target area of interest than single sourced (RS) data.

Table 7-4: Accuracy indices for land use classification from different data sources.

Accuracy index	RS single-date	RS Multi-dates	RS and GIS data
Commission error	0.059	0.106	0.058
User's accuracy	0.941	0.894	0.942
Omission error	0.174	0.083	0.073
Producer's accuracy	0.826	0.917	0.927
Kappa	0.733	0.782	0.881
Overall accuracy	0.808	0.85	0.916

7.2 Accuracies of RS-derived crop information

Empirical relationships between a crop parameter and the vegetation indices for a certain growth stage are the main focus in recent studies (e.g., Bannari et al., 1995; Jacquemoud et al., 2009; Yang et al., 2013; Yao et al., 2012). Thus, the empirical relationships between RS information and crop parameters are case- and time-specific which limits its transferability to other study cases. For instance in this study, rice characteristics (i.e., biomass, LAI, N concentration, N uptake) were estimated using time-specific (i.e., June 24th, July 6th, August 9th) regression models and such models may not be valid for another year. Although rice yield was estimated using one single regression model for all three growth stages (**Chapter 6**), the empirical regression model was still limited to the study case. Studies tend to find out generic relationships between RS reflectances and crop parameters over a long term (e.g., Zhong et al., 2014). However, more investigations are still needed to explore the mechanisms between RS reflectances and crop status by decomposing the originating factors of the two.

In-time response of crop status is required in precision agriculture (Mulla, 2013). With the dramatic improvement in temporal frequency of remote sensing imagery, there is growing interest in deriving crop information from time series of RS images to conduct in-time/in-season field management. Such applications could be difficult for rice in the SJP due to several reasons. First, rice grows fast during the short growing season of approximately 120 days. Specially, in some critical growing stages such as jointing stage and heading stage, rice status could change greatly in a few days (Ling, 2002). Second, the field management is time intensive during the short growing season. For instance, fertilizer could be applied once a week and irrigation could be applied twice a week according to the rice status (Yao et al., 2012). Third, weather conditions are considerable constraints for the application of satellite remote sensing in the SJP (Meng et al., 2005). Under the monsoon climate, clouds and rains gather in late June to early August in the SJP, while this period is the key time for rice growth

and conducting field management (Yao et al., 2012). Although the weather conditions are good for satellite remote sensing in the early growing stage of late May and early June, the remotely sensed rice signals are normally too weak due to the low biomass density. Last but not least, transferring the commercial satellite imagery into rice information needs processing time technically and commercially.

As aforementioned, generally the RS-based crop status monitoring rely on empirical relationships between the remotely sensed information and the field observed crop properties (e.g., Shang et al., 2015; Vicente-Guijalba et al., 2015; Moran et al., 1997). Therefore, ground truth information of crops is indispensable. Field campaigns for collecting ground truth data are time consuming and labor-intensive, especially for a large spatial extent. It is of great value to build spectral libraries for specific crops during the whole growing season. However, annual differences in crop status calendar and impacts of climate changes have to be considered for long term investigation (Rao, 2008; Zhong et al., 2014). Moreover, the strategy of conducting ground data collection is critical since the errors from ground truth data were greater than the errors from the remotely sensed data (Curran and Williamson, 1985). Furthermore, the accuracy of information derived from the remotely sensed data is highly depends on the positional accuracy of the ground truth data (Congalton and Green, 2008).

7.3 Spatial scales for an AEIS

It is demonstrated that although the multiple scales in an agro-ecosystem study is a difficult issue due to the lack of scaling up methods, endless complexity, time lags, limited predictability of models, and connections between local to macroscales (Sayer et al., 2013; Ostrom, 2007). In this research, the site-specific DNDC was enabled to apply on a regional scale driven by additional scripts. Thus the model run for every pixel to get detailed information of the study area.

Field plots are treated as the basic management units in paddy rice fields since the enclosed raised boundaries not only provide uniform crop conditions in terms of water management, tillage, and even fertilizer, but also spatially isolate one field plot from another. Nonetheless, within the field block there are still variabilities in rice growing status induced by soil conditions, human activities, or even pests and diseases. Information on a fine resolution of within-field scale assists in monitoring rice precisely and closing yield gaps through site-specific management measures, and likewise, decreasing environmental effects such as greenhouse gas emissions.

In this study, 28 ground truth experimental plots were used to construct the empirical models to extract rice status. The area of each pixel of the FS-2 satellite imagery is 64 m², and

the areas of experimental field plots range from 1000 m² to 3000 m². In each experimental plot, up to four samples were collected. Limitation exists when using such four samples to represent all the pixels (ca. 15 to 47 pixels) in the experimental plot. The differences of pixels in one experimental plots may be considerable but averaged by the limited number of samples. As a result, the pixel-based ‘within-field’ detection was relatively mixed with neighbouring pixels. In addition, the locations of the 28 ground truth plots were clustered, which may lead to uncertainties in the rice monitoring procedure. Specially, in **chapter 5**, although specific polynomial growing curves for each site were constructed to interpolate the ground truth status at the imagery collection dates, it is challengeable to represent the heterogeneity of the rice growing process over the entire Qixing Farm by using only 28 ground truth plots.

7.4 Model uncertainty

Process-based models are widely used to explore the interactions between the factors of the target objects, predict results under certain scenarios, and anticipate possible outcomes under future conditions (Friend et al., 1997; An, 2012). These models are constructed based on a series of functions that are represented by mechanism or empirical equations. It is challengeable for a process-based agro-ecosystem model to consider every factor relating to the intricate and manifold agro-ecosystems (Belcher et al., 2004). On the other hand, redundant parameters in a model will indisputably block the model transferability (Perlman et al., 2013).

With the focus on food production in the SJP, rice yield was first estimated using the process-based agro-ecosystem model DNDC. Effects of soil properties on rice yield were analyzed on the basis of soil characteristics determining the status of available water and nutrients that drive the crop growth. Specially, in the regional modelling application, detailed soil data with fine resolution decrease the model uncertainties (Zhang et al., 2016). Validation results demonstrated good performance of the DNDC model in regional rice yield estimation (**chapter 7**).

However, the DNDC model showed weakness in capturing the time series biomass status (**Figure 7-1, 7-2**). The DNDC overestimated the biomass for both the ‘high yield’ and ‘medium yield’ fields (see definitions in **chapter 7**) across the growing season in 2009. This may due to the crop phenological development scheme in the DNDC model, which affects the simulation results on plant growth greatly (Fumoto et al., 2008; Zhang et al., 2002). A similar shape of the biomass accumulation curves is detected between modelled values and field observed values. However, the modelled biomass accumulation curves start earlier than the

field observed values (Figure 7-1, 7-2). This may result from the generic crop growth curve inherent in the DNDC, which may not be suitable for rice in the SJP.

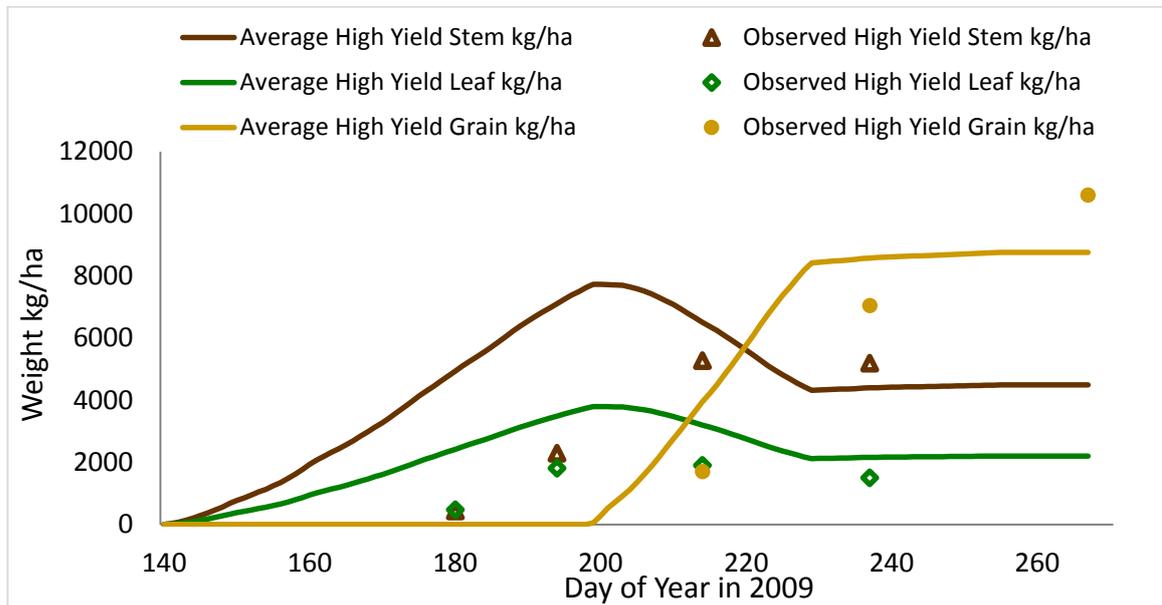


Figure 7-1: Time series data of DNDC modelled biomass vs. field observed values in the ‘high yield’ fields.

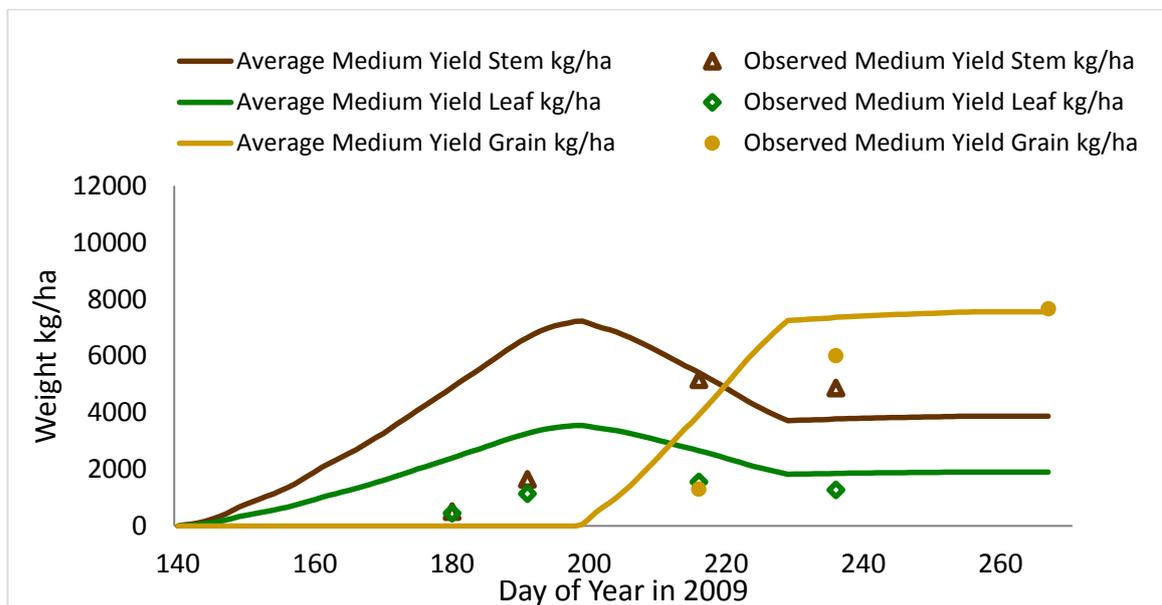


Figure 7-2: Time series data of DNDC modelled biomass vs. field observed values in the ‘medium yield’ fields.

Due to the limited thermal resource in the SJP, rice has to be sown in the greenhouse in early April and the seedlings are grown in the greenhouse until mid-May to improve the local

thermal efficiency. Thus, as a starting step for rice in the paddy field, seedling transplanting is a critical activity which greatly determines the phenological stage that drives the partitioning assimilates for rice. However, the DNDC version utilized in this study did not include the transplanting options for rice. Therefore, it is challengeable for the current model version to capture the time series of biomass accumulation due to considerable uncertainties in defining the phenological stages. The DNDC-Rice (Fumoto et al. 2008), a revised version of DNDC, is promising in future rice applications.

7.5 Environmental effects

In an AEIS, impacts of agriculture production on environment have to be interpreted. Greenhouse gases emissions from the rice-ecosystems are treated as one important environmental effects on global climate change which have been attracted many focuses (e.g., Huang et al., 2010; Linquist et al., 2012). As nitrous oxide (N_2O) emissions occur via nitrification and denitrification, the anaerobic-aerobic cycling in paddy soils promotes N_2O emission (Granli and Bøckman, 1994; Smith and Patrick, 1983). It is important to better understand the N_2O flux coupled with water regimes and fertilizer applications in the paddy field to decrease the greenhouse gas effects by conducting field management strategies (Cai et al., 1997).

The DNDC model is one of the few agro-ecosystem models that simulate N_2O emissions. In DNDC, N_2O production/consumption is directly regulated by three factors of soil redox potential (i.e. Oxidation-Reduction Potential [Eh]), dissolved organic carbon (DOC) concentration and available N (i.e. ammonium or nitrate) concentration (Giltrap et al., 2010). One of the key processes controlling N_2O production/consumption in paddy soils is soil Eh dynamics (Yu and Patrick, 2004). Paddy rice is characterized by the frequent changes between saturated and unsaturated conditions driven by the water management. During these changes in soil water content, the soil redox potential changes is subjected to substantial changes.

The N_2O emission pattern (**Figure 7-3**, left) of the Qixing Farm is consistent with the soil input data patterns (Figure in **chapter 7**). The N_2O emissions of this study is lower than the results of Brocks et al. (2014), who explored an empirical approach to estimate the N_2O emissions all over Germany. The reason could be due to the climate conditions and the frequency of anaerobic-aerobic exchanges in the soil. In the SJP, NE-China, the snow-covered winter with temperature below 0 °C lasts from late October to early April, which is much longer than that in Germany. The lower temperature prevents the N_2O flux. Although

the agriculture land use in this study is paddy rice, the frequency of human-induced anaerobic-aerobic exchanges is lower than that in Germany due to higher precipitation (Brocks et al., 2014).

As a typical greenhouse gas emitted from paddy rice field, carbon dioxide (CO₂) characterizing the SOC dynamics as well (Eswaran et al., 1993). In this study, soil CO₂ emissions were also modelled using the DNDC model (**Figure 7-3**, right).

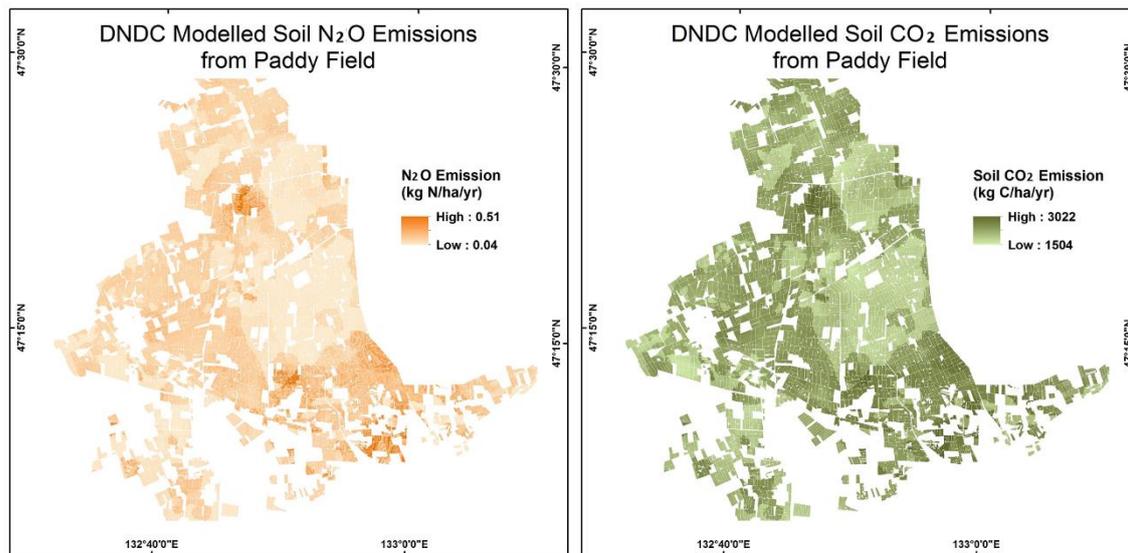


Figure 7-3: Maps of DNDC modelled greenhouse gas emissions from paddy rice field in Qixing Farm.

7.6 Summary

The AEIS 1) assists in multi-source data organization, i.e. using the KBS to integrate multi-source data for deriving required information; 2) provides information about the agriculture in the SJP, i.e. rice status derived from the FS-2 imagery; 3) represents the interactions between environmental factors and agriculture, i.e. the effects of soil characteristics on rice yield; and 4) estimates the environmental effects of agriculture, i.e. greenhouse gas emissions.

DNDC is a powerful process-based geochemical model that can be used not only to simulate crop growth, but also to estimate greenhouse gas emissions and soil C changes (Giltrap et al., 2012). The results of this study showed that DNDC is one of the suitable models that can be implemented in an AEIS for the research regarding sustainable agriculture.

However, issues of water resource were not considered in this study. As water is one of the basic resource required for paddy rice production, there is a crucial need to investigate water resources accompanied with rice and other environmental factors (e.g., climate change).

8 Future Challenges and Outlook

8.1 Requirement of a strong network for gathering sufficient high-quality, in-time data

As an interdisciplinary research, implementing an AEIS needs multi-source data including crop growth, soil characteristics and soil nutrients dynamics, water availability, landscape changes, field management evolution. With the dramatic development of the earth observation and monitoring systems, there is a demand for deploy data to provide synthesized information for the intended stakeholders. To get sufficient data in time and in high quality, strong cooperative networks are need to acquire, manage, archive, interpret and disseminate the multi-source data with systematic data metrics, formats, collection protocols, spatial resolutions and measurement frequencies. There is a need to explore approaches for pre-processing the multi-source data with respect to its heterogeneous and complex characteristics.

8.2 Improve model capability

The DNDC model is one of the optimum agro-ecosystem models for an AEIS because it enables the users to interpret the interactions between agriculture and environmental factors (e.g., greenhouse emissions) on the basis of simulating biochemical processes in the soil-vegetation-atmosphere system. Whereas the crop module of the DNDC has to be further improved to capture the crop status precisely.

Spatial distribution of model responses on a ‘within-field’ level would be beneficial for many applications. In future studies, there is a need to explore feasible technologies to generalize site-specific models onto regional applications. To adapt process-based models into a GIS environment could be promising in the modelling application on a regional extent. However, processing time could be a considerable limitation in the regional application of site-specific models. Uncertainties sourced from the spatial data (i.e., climate data, soil data) could aggregate model simulation uncertainties. Proper uncertainty analyses are needed to better understand the model predictions, especially for a regional applications of site-specific models.

8.3 Assimilate RS-derived crop information into process-based agro-ecosystem modelling on large extent

Process-based agro-ecosystem modelling connects the crop variables such as LAI, biomass and the developmental phase to the environmental factors of climate, soil and field management on intensive time steps. Satellite RS captures unprecedented data in forms of electromagnetic signals over a large geographical extent. Such signals are interpreted into earth

surface information in terms of vegetation, climate, and soils. The parallel development of RS and process-based agro-ecosystem models drives scientists' interest to develop synergic applications. To assimilate RS-derived crop information (i.e., LAI, biomass) into process-based agro-ecosystem modelling would largely improve the model performances (Dorigo et al., 2007). Moreover, model processing time and uncertainties have to be quantitatively controlled.

8.4 Describing regional long-term stories using the AEIS

The landscape, soil conditions, water resources, and even micro-climate in the SJP have been changed greatly in the past 60 years due to the intensive reclamation events of crop lands. These changes from a natural system to a human dominated agricultural system comprise many valuable processes for scientists to describe. Long-term dynamics in its agro-ecosystems are worth to study for a better understanding of the complex interactions between the agricultural production and the environmental factors, and ultimately, for sustainable agriculture. Future work needs to position the AEIS in a long-term scope, to get better knowledge of nature and to guide human beings to perform in a more 'proper' way to nature.

References*Chapters 1,2,3,7,8

- An, L., 2012. Modeling human decisions in coupled human and natural systems: review of agent-based models. *Ecological Modelling*, 229, 25–36.
- An, S., Li, H., Guan, B., Zhou, C., Wang, Z., Deng, Z., Zhi, Y., Liu, Y., Xu, C., Fang, S., Jiang, J., Li, H., 2007. China's natural wetlands: past problems, current status, and future challenges. *AMBIO: A Journal of the Human Environment*, 36 (4), 335–342.
- Baldocchi, D. D., Wilson, K. B., Gu, L., 2002. How the environment, canopy structure and canopy physiological functioning influence carbon, water and energy fluxes of a temperate broad-leaved deciduous forest—an assessment with the biophysical model CANOAK. *Tree Physiology*, 22 (15-16), 1065–1077.
- Ban, Y., Hu, H., Rangel, I. M., 2010. Fusion of Quickbird MS and RADARSAT SAR data for urban land-cover mapping: Object-based and knowledge-based approach. *International Journal of Remote Sensing*, 31 (6), 1391–1410.
- Bannari, A., Morin, D., Bonn, F., Huete, A.R., 1995. A review of vegetation indices. *Remote sensing reviews*, 13 (1-2), 95–120.
- Bareth, G. Yu, Z., 2004. Verfügbarkeit von digitalen Geodaten in China. – *Petermanns Geographische Mitteilungen*, 148 (5), 78–85 (written in German with English abstract).
- Bareth, G., 2009, GIS- and RS-based spatial decision support: structure of a spatial environmental information system (SEIS). *International Journal of Digital Earth*, 2 (2), 134–154.
- Belcher, K. W., Boehm, M. M., Fulton, M. E., 2004. Agroecosystem sustainability: a system simulation model approach. *Agricultural Systems*, 79 (2), 225–241.
- Brocks, S., Jungkunst, H. F., Bareth, G., 2014. A regionally disaggregated inventory of nitrous oxide emissions from agricultural soils in Germany—a GIS-based empirical approach. *Erdkunde*, 68 (2), 125–144.
- Campbell, J. B., Wynne R. H. (Eds.), 2011. *Introduction to remote sensing*. CRC Press (fifth edition).

- Cai, Z., Xing, G., Yan, X., Xu, H., Tsuruta, H., Yagi, K., Minami, K., 1997. Methane and nitrous oxide emissions from rice paddy fields as affected by nitrogen fertilisers and water management. *Plant and Soil*, 196 (1), 7–14.
- Cao, H., Fang, D., Zhang, J., Sun, T., Shang, W., Jin, X., 2005. The principle of increase production and studies on the key technique of 3-S cultivating technique of rice in cold zone. *Review of China Agriculture Science and Technology*, 5, 8–12 (Written in Chinese with English abstract).
- Cao, Z., Ding, J., Hu, Z., Knicker, H., Kögel-Knabner, I., Yang, L., Yin, R., Lin, X., Dong, Y., 2006. Ancient paddy soils from the Neolithic age in China's Yangtze River Delta. *Naturwissenschaften*, 93, 232–236.
- Câmara, G., Souza, R. C. M., Freitas, U. M., Garrido, J., 1996. SPRING: Integrating remote sensing and GIS by object-oriented data modelling. *Computers & graphics*, 20 (3), 395–403.
- Curtis, G., Cobham, D. (Eds.), 2008. *Business information systems: Analysis, design and practice*. Pearson Education.
- Chang, C. H., Liu, C. C., Tseng, P. Y., 2013. Emissions inventory for rice straw open burning in Taiwan based on burned area classification and mapping using FORMOSAT-2 satellite imagery. *Aerosol and Air Quality Research*, 13, 474–487.
- Chang, K., Shen, Y., Lo, J., 2005. Predicting rice yield using canopy reflectance measured at booting stage. *Agronomy Journal*, 97, 872–878.
- Chen, C., Knoblock C. A., Shahabi, C., 2008. Automatically and accurately conflating raster maps with orthoimagery. *Geoinformatica*, 12, 377–410.
- Chen, L., Wang, L., Ma, Y., Liu, P., 2014. Overview of Ecohydrological Models and Systems at the Watershed Scale. *IEEE Systems Council*, 1091–1099.
- Cohen, Y., Shoshany, M., 2005. Analysis of convergent evidence in an evidential reasoning knowledge-based classification. *Remote Sensing of Environment*, 96 (3), 518–528.
- Congalton, R. G., Green, K. (Eds.), 2008. *Assessing the accuracy of remotely sensed data: principles and practices*. 183 p., CRC press, Boca Raton, FL, USA.
- Curran, P. J., Williamson, H. D., 1985. The accuracy of ground data used in remote-sensing investigations. *International Journal of Remote Sensing*, 6 (10), 1637–1651.

- Curdt, C., 2014. Design and Implementation of a Research Data Management System: The CRC/TR32 Project Database (TR32DB) (Doctoral dissertation, Universität zu Köln).
- Dalgaard, T., Nicholas J. H., . John R. P., 2003. Agroecology, scaling and interdisciplinarity. *Agriculture, Ecosystems & Environment* 100 (1), 39–51.
- Dangermond, J., 2009. GIS: Designing our future. *ArcNews*, 31 (2), 1.
- de Wit, C. T., 1965. Photosynthesis of leaf canopies. Agricultural Research Report, 663. Wageningen, the Netherlands.
- Devogele, T., Parent, C., Spaccapietra, S., 1998. On spatial database integration. *International Journal of Geographical Information Science*, 12 (4), 335–352.
- Dorigo, W. A., Zurita-Milla, R., de Wit, A. J., Brazile, J., Singh, R., Schaepman, M. E., 2007. A review on reflective remote sensing and data assimilation techniques for enhanced agroecosystem modeling. *International Journal of Applied Earth Observation and Geoinformation*, 9 (2), 165–193.
- Duncan, W. G., Loomis, R. S, Williams, W. A, Hanau, R., 1967. A model for simulating photosynthesis in plant communities. *Hilgardia*, 4, 181–205.
- Duveiller, G., López-Lozano, R., Seguíni, L., Bojanowski, J. S., Baruth, B., 2012b. Optical remote sensing requirements for operational crop monitoring and yield forecasting in Europe. In *Proceedings of 2012 nel-3 OLCI/SLSTR and MERIS/(A) ATSR Workshop*, Frascati, Italy, 15–19, October 2012.
- Eitzinger, J., Trnka, M., Hösch, J., Žalud, Z., Dubrovský, M., 2004. Comparison of CERES, WOFOST and SWAP models in simulating soil water content during growing season under different soil conditions. *Ecological Modelling*, 171 (3), 223–246.
- Estes, J., Star, J. (Eds.), 1990. *Geographic information systems*. University of California. Santa Barbara-EEUU. 295.
- Eswaran, H., van Den Berg, E., Reich, P., 1993. Organic carbon in soils of the world. *Soil Science Society of America Journal*, 57 (1), 192–194.
- Ewert, F., van Ittersum, M. K., Heckeley, T., Therond, O., Bezlepkina, I., Andersen, E., 2011. Scale changes and model linking methods for integrated assessment of agri-environmental systems. *Agriculture, Ecosystems & Environment*, 142 (1), 6–17.

- Fegraus, E. H., Zaslavsky, I., Whitenack, T., Dempewolf, J., Ahumada, J., Lin, K., Andelman, S. J., 2012. Interdisciplinary decision support dashboard: A new framework for a Tanzanian agricultural and ecosystem service monitoring system pilot. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5 (6), 1700–1708.
- FAO (Food and Agriculture Organization of the United Nations), 2014. FAO Statistical Database. Available online: <http://www.fao.org/docrep/015/i2490e/i2490e03d.pdf> (accessed on Jan. 21, 2016).
- Fumoto, T., Kobayashi, K., Li, C., Yagi, K., Hasegawa, T., 2008. Revising a process based biogeochemistry model (DNDC) to simulate methane emission from rice paddy fields under various residue management and fertilizer regimes. *Global Change Biology*, 14, 382–402.
- Foley, J. A., DeFries, R., Asner, G. P., Barford, C., Bonan, G., Carpenter, S. R., Chapin, F. S., Coe, M. T., Daily, G. C., Gibbs, H. K., Helkowski, J. H., Holloway, T., Howard, E. A., Kucharik, C. J., Monfreda, C., Patz, J. A., Prentice, C., Ramankutty, N., Snyder, P. K., 2005. Global consequences of land use. *Science*, 309 (5734), 570–574.
- Fountas, S., Wulfsohn, D., Blackmore, B. S., Jacobsen, H. L., Pedersen, S. M., 2006. A model of decision-making and information flows for information-intensive agriculture. *Agricultural Systems*, 87 (2), 192–210.
- Friend, A. D., Stevens, A. K., Knox, R. G., Cannell, M. G. R., 1997. A process-based, terrestrial biosphere model of ecosystem dynamics (Hybrid v3.0). *Ecological Modelling*, 95 (2), 249–287.
- Frolking, S., Qiu, J., Boles, S., Xiao, X., Liu, J., Zhuang, Y., Li, C., Qin, X., 2002. Combining remote sensing and ground census data to develop new maps of the distribution of rice agriculture in China. *Global Biogeochemical Cycles*, 16.
- Georghiou, G. P. (Eds.), 2012. *Pest resistance to pesticides*. Springer Science & Business Media.
- GEOSS (The Global Earth Observation System of Systems). Available online: http://www.geoportal.org/web/guest/geo_home_stp (accessed on Sep. 8, 2015).
- Gessner, U., Machwitz, M., Esch, T., Tillack, A., Naeimi, V., Kuenzer, C., Dech, S., 2015. Multi-sensor mapping of West African land cover using MODIS, ASAR and TANDEM-X/TerraSAR-X data. *Remote Sensing of Environment*, 164, 282–297.

- Giltrap, D. L., Li, C., Saggarr, S., 2010. DNDC: A process-based model of greenhouse gas fluxes from agricultural soils. *Agriculture, Ecosystems & Environment*, 136 (3), 292–300.
- Girardin, P., Bockstaller, C., Werf, H. V. D., 1999. Indicators: tools to evaluate the environmental impacts of farming systems. *Journal of Sustainable Agriculture*, 13 (4), 5–21.
- Gitelson, A. A., Peng, Y., Arkebauer, T. J., Schepers, J., 2014. Relationships between gross primary production, green LAI, and canopy chlorophyll content in maize: Implications for remote sensing of primary production. *Remote Sensing of Environment*, 144, 65–72.
- Gong, L., Li, S., Jiang, L., Zhu, H., 2015. Spatial temporal characteristics of delayed cool injury for rice from 1961 to 2010 in Heilongjiang Province. *Journal of Meteorology and Environment*, 31 (1), 76–83 (Written in Chinese with English abstract).
- Goodchild, M. F., 2009. Geographic information system. In *Encyclopedia of Database Systems*, pp. 1231–1236. Springer US.
- Goodenough, D. G., Goldberg, M., Plunkett, G., Zelek, J., 1987. An expert system for remote sensing. *Geoscience and Remote Sensing, IEEE Transactions on*, (3), 349–359.
- Güting, R. H., 1994. An introduction to spatial database systems. *The VLDB Journal—The International Journal on Very Large Data Bases*, 3 (4), 357–399.
- Granli, T., Bøckman, O. C., 1994. Nitrous oxide from agriculture. *Norwegian Journal of Agricultural Sciences*, Norway, No. 12.
- Hansen, J. W., Jones, J. W., 2000. Scaling-up crop models for climate variability applications. *Agricultural Systems*, 65 (1), 43–72.
- HOHLRB (History Office of Heilongjiang Land Reclamation Bureau), 1991. *Annals of Heilongjiang Province—Annals of National Farms*, 1991. (www.zglz.gov.cn, Accessed in July, 2015).
- Holmgren, P., Thuresson, T., 1998. Satellite remote sensing for forestry planning—a review. *Scandinavian Journal of Forest Research*, 13 (1-4), 90–110.
- Howden, S. M., Soussana, J. F., Tubiello, F. N., Chhetri, N., Dunlop, M., Meinke, H., 2007. Adapting agriculture to climate change. *Proceedings of the National Academy of Sciences*, 104 (50), 19691–19696.

- HSB&HSTNB (Heilongjiang Statistical Bureau and Heilongjiang Survey Team of the National Bureau), 2011. Heilongjiang Statistical Yearbook, China statistics press, Beijing (in Chinese).
- Huang, Y., Sun, W., Zhang, W., Yu, Y., Su, Y., Song, C., 2010. Marshland conversion to cropland in northeast China from 1950 to 2000 reduced the greenhouse effect. *Global Change Biology*, 16 (2), 680–695.
- ILTER (International Long Term Ecological Research). Available online: www.ilternet.edu (accessed pm Sep. 8, 2015).
- Jacquemoud, S., Verhoef, W., Baret, F., Bacour, C., Zarco-Tejada, P. J., Asner, G. P., François, C., Ustin, S. L., 2009. PROSPECT+ SAIL models: A review of use for vegetation characterization. *Remote Sensing of Environment*, 113, S56–S66.
- Jiang, L., Deng, X., Seto, K. C., 2013. The impact of urban expansion on agricultural land use intensity in China. *Land Use Policy*, 35, 33–39.
- Jiang, Q., Fu, Q., Wang, Z., 2011. Evaluation and regional differences of water resources carrying capacity in Sanjiang plain. *Transactions of the Chinese Society of Agricultural Engineering*, 27 (9), 184–190 (written in Chinese with English abstract).
- Jin, C., Xiao, X., Dong, J., Qin, Y., Wang, Z., 2015. Mapping paddy rice distribution using multi-temporal Landsat imagery in the Sanjiang Plain, northeast China. *Frontiers of Earth Science*, 1–14.
- Ju, X., Xing, G., Chen, X., Zhang, S., Zhang, L., Liu, X., Cui, Z., Yin, B., Christie, P., Zhu, Z., Zhang, F., 2009. Reducing environmental risk by improving N management in intensive Chinese agricultural systems. *Proceedings of the National Academy of Sciences*, 106 (9), 3041–3046.
- Kako, T., Zhang, J., 2000. Problems concerning Grain Production and Distribution in China: The Case of Heilongjiang Province. *The Developing Economies*, 38 (1), 51–79.
- Khaleghi, B., Khamis, A., Karray, F. O., Razavi, S. N., 2013. Multisensor data fusion: A review of the state-of-the-art. *Information Fusion*, 14 (1), 28–44.
- Kim, H. O., Yeom, J. M. 2014. Effect of red-edge and texture features for object-based paddy rice crop classification using RapidEye multi-spectral satellite image data. *Int. J. Remote Sensing*, 35, 7046–7068.

- Kim, H. O., Yeom, J. M., 2012. Multi-temporal spectral analysis of rice fields in South Korea using MODIS and RapidEye satellite imagery. *Journal of Astronomy and Space Sciences*, 29, 407–411.
- Kontoes, C., Wilkinson, G. G., Burrill, A., Goffredo, S., Megier, J., 1993. An experimental system for the integration of GIS data in knowledge-based image analysis for remote sensing of agriculture. *International Journal of Geographical Information Systems*, 7 (3), 247–262.
- Kuenzer, C., Knauer, K., 2013. Remote sensing of rice crop areas. *International Journal of Remote Sensing*, 34, 2101–2139.
- Lambin, E. F., Meyfroidt, P., 2011. Global land use change, economic globalization, and the looming land scarcity. *Proceedings of the National Academy of Sciences*, 108 (9), 3465–3472.
- Li, C., 2000. Modeling trace gas emissions from agricultural ecosystems. *Nutrient Cycling in Agroecosystems*, 58 (1-3), 259–276.
- Li, C., 2007. Quantifying greenhouse gas emissions from soils: Scientific basis and modeling approach. *Soil Science and Plant Nutrition*, 53 (4), 344–352.
- Li, C., Frohling S., Frohling, T.A., 1992. A model of nitrous oxide evolution from soil driven by rainfall events: 1. Model structure and sensitivity. *Journal of Geophysical Research: Atmospheres*, 97 (D9), 9759–9776.
- Li, C., Mosier, A., Wassmann, R., Cai, Z., Zheng, X., Huang, Y., Tsuruta, H., Boonjawat, J., Lantin, R., 2004. Modeling greenhouse gas emissions from rice-based production systems: Sensitivity and upscaling. *Global Biogeochemical Cycles*, 18, GB1043.
- Li, C., Qiu, J., Frohling S., Xiao, X., Salas, W., Moore III, B., Boles, S., Huang Y., Sass, R., 2002. Reduced methane emissions from large-scale changes in water management of China's rice paddies during. *Geophysical Research Letters*, 29 (20), 33-1–33-4.
- Li, D., 2010. Remotely sensed images and GIS data fusion for automatic change detection. *International Journal of Image and Data Fusion*, 1 (1), 99–108.
- Li, H., Qiu, J., Wang, L., Tang, H., Li, C., van Ranst, E., 2010. Modelling impacts of alternative farming management practices on greenhouse gas emissions from a winter wheat–maize rotation system in China. *Agriculture, Ecosystems & Environment*, 135 (1), 24–33.

- Ling, Q. (Eds.), 2000. Traits of crop group biomass. Shanghai Science and Technology Press (in Chinese).
- Linquist, B., Groenigen, K. J., Adviento-Borbe, M. A., Pittelkow, C., Kessel, C., 2012. An agronomic assessment of greenhouse gas emissions from major cereal crops. *Global Change Biology*, 18 (1), 194–209.
- Liang, C., Zhuang, H., Gao, H., Qiu, H., 2013. The application and research progress of GIS technology in the cultivation of rice with high quality and high yield. *China Rice*, 19 (2), 14–17 (Written in Chinese).
- Liao, S. H., 2005. Expert system methodologies and applications—a decade review from 1995 to 2004. *Expert Systems with Applications*, 28 (1), 93–103.
- Lillesand, T., Kiefer, R. W., Chipman, J. (Eds.), 2014. Remote sensing and image interpretation. John Wiley & Sons.
- Liu, B., 2010. Regimes of Northeast rice cultivation and pest and disease control. *Applied Technology*, 12, 168, 152 (written in Chinese).
- Liu, X., Dong, G., Zhang, Y., Lu, X., Jiang, M., 2013. Contribution to the Global Warming Mitigation of Marshlands Conversion to Croplands in the Sanjiang Plain, Northeast China. *CLEAN—Soil, Air, Water*, 41 (4), 319–324.
- Liu, X., Ma, X., 2000. Influence of large-scale reclamation on natural environment and regional environmental protection in the Sanjiang Plain. *Scientia Geographica Sinica*, 20 (1), 14–19.
- Lobell, D. B., Burke, M. B., 2010. On the use of statistical models to predict crop yield responses to climate change. *Agricultural and Forest Meteorology*, 150 (11), 1443–1452.
- Long, S. P., Ainsworth, E. A., Leakey, A. D., Nösberger, J., Ort, D. R., 2006. Food for thought: lower-than-expected crop yield stimulation with rising CO₂ concentrations. *Science*, 312 (5782), 1918–1921.
- Lu, H., Tian, F., Hu, H., Tang Y., 2002. Irrigation Decision Support System Based on Genetic Algorithm and Geographic Information System. *Water Resources and Hydro-power Engineering*, 7, 009 (in Chinese with English abstract).
- Luan, Z., Zhang, G., Deng, W., Hu, J., Zhou, D., 2007. Study on the changes of air temperature and precipitation in the last 50 years in the Sanjiang Plain. *Journal of Arid Land*

- Resources and Environment, 21 (11), 39–43 (written in Chinese with English abstract).
- Lv, D., 2014. Application of agricultural water-saving irrigation techniques. *Agriculture and Technology*, 34 (11), 40–41 (written in Chinese).
- Ma, Z., 2015. An ancillary decision support system for rice field management based on Android platform. Master Thesis, Heilongjiang Bayi Agricultural University, Daqing, China (in Chinese).
- Macary, F., Almeida-Dias, J., Uny, D., Probst, A., 2013. Assessment of the effects of best environmental practices on reducing pesticide contamination in surface water, using multi-criteria modelling combined with a GIS. *International Journal of Multicriteria Decision Making* 73, 3(2-3), 178–211.
- MacDonald, R. B., Hall, F. G., 1980. Global crop forecasting. *Science*, 208, 670–679.
- McCloy, K. R., 2005. Resource management information systems: Remote sensing, GIS and modelling. CRC Press.
- McCoy, M. D., Asner, G. P., Graves, M. W., 2011. Airborne lidar survey of irrigated agricultural landscapes: an application of the slope contrast method. *Journal of Archaeological Science*, 38 (9), 2141–2154.
- McKeown Jr, D. M., 1987. The role of artificial intelligence in the integration of remotely sensed data with geographic information systems (No. CMU-CS-86-174). Carnegie-Mellon Univ Pittsburgh pa Dept of Computer Science.
- Meng, Z., Wang, Y., Sun, Z., Meng, Q., Liu, Y., Li, C., Zhang, J., 2005. The research of rice sterile-type chilling injury and irrigation technique. *Chinese Agricultural Science Bulletin*, 21 (6), 197–201 (in Chinese with English abstract).
- Mitsch, W. J., Wilson, R. F., 1996. Improving the success of wetland creation and restoration with know-how, time, and self-design. *Ecological Applications*, 77–83.
- Moran, M. S., Inoue, Y., Barnes, E. M., 1997. Opportunities and limitations for image based remote sensing in precision crop management. *Remote Sensing of Environment*, 61, 319–346.
- Mulla, D. J., 2013. Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. *Biosystems Engineering*, 114, 358–371.

- Munyati, C., 2000. Wetland change detection on the Kafue Flats, Zambia, by classification of a multitemporal remote sensing image dataset. *International Journal of Remote Sensing*, 21 (9), 1787–1806.
- Mückschel, C., Weist, C., Köhler, W., Populationsgenetik, B., 2008. Central Data Management in Environmental Research Projects-Selected Problems and Solutions. In *GIL Jahrestagung* (pp. 101-104).
- Mwaniki, W. M., Möller, S. M., 2015. Knowledge based multi-source, time series classification: A case study of central region of Kenya. *Applied Geography*, 60, 58–68.
- Nachtergaele, F., van Velthuisen, H., Verelst, L., 2008. Harmonized World Soil Database, Food and Agriculture Organization of the United Nations.
- Naeem, S., Duffy, J. E., Zavaleta, E., 2012. The functions of biological diversity in an age of extinction. *Science*, 336 (6087), 1401–1406.
- NEON (National Ecological Observatory Network). Available online: www.neoninc.org (accessed: Sep. 8, 2015).
- Ni, H., Li, J., 1999. Plant diversity of Honghe Nature Reserve in Sanjiang Plain. *Territory & Natural Resources Study*, (3), 12–18 (written in Chinese).
- OOI (Ocean Observatories Initiative). Available online: www.oceanobservatories.org (accessed pm Sep. 8, 2015).
- Ostrom, E., 2007. A diagnostic approach for going beyond panaceas. *Proceedings of the National Academy of Sciences*, 104 (39), 15181–15187.
- Ozesmi, S. L., Bauer, M. E., 2002. Satellite remote sensing of wetlands. *Wetlands Ecology and Management*, 10 (5), 381–402.
- Peng, S., Tang, Q., Zou, Y., 2009. Current status and challenges of rice production in China. *Plant Production Science*, 12 (1), 3–8.
- Perlman, J., Hijmans, R. J., Horwath, W. R., 2013. Modelling agricultural nitrous oxide emissions for large regions. *Environmental Modelling and Software*, 48, 183–192.
- Pingali, P. L., Roger, P. A. (Eds.), 2012. *Impact of pesticides on farmer health and the rice environment* (Vol. 7). Springer Science & Business Media.
- Popp, J., Pető, K., Nagy, J., 2013. Pesticide productivity and food security. A review. *Agronomy for Sustainable Development*, 33 (1), 243–255.

- Poudel, D. D., Lee, T., Srinivasan, R., Abbaspour, K., Jeong, C. Y., 2013. Assessment of seasonal and spatial variation of surface water quality, identification of factors associated with water quality variability, and the modeling of critical nonpoint source pollution areas in an agricultural watershed. *Journal of Soil and Water Conservation*, 68 (3), 155–171.
- Power, D. J. (Eds.), 2008. Decision support systems: a historical overview. In *Handbook on Decision Support Systems 1* (pp. 121–140). Springer Berlin Heidelberg.
- Rao, N. R., 2008. Development of a crop-specific spectral library and discrimination of various agricultural crop varieties using hyperspectral imagery. *International Journal of Remote Sensing*, 29 (1), 131–144.
- Reidsma, P., Ewert, F., Lansink, A. O., Leemans, R., 2010. Adaptation to climate change and climate variability in European agriculture: the importance of farm level responses. *European Journal of Agronomy*, 32 (1), 91–102.
- Reinartz, P., Müller, R., Schwind, P., Suri, S., Bamler, R., 2011. Orthorectification of VHR optical satellite data exploiting the geometric accuracy of TerraSAR-X data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66 (1), 124–132.
- Ren, J., 2014. Strategies for processing relationships between Rice Farming and Groundwater Decrease in Jiansanjiang Region. Master thesis, China agricultural university, Beijing, China (written in Chinese with English abstract).
- Resop, J. P., Fleisher, D. H., Wang, Q., Timlin, D. J., Reddy, V. R., 2012. Combining explanatory crop models with geospatial data for regional analyses of crop yield using field-scale modeling units. *Computers and Electronics in Agriculture*, 89, 51–61.
- Ronald, P., 2011. Plant genetics, sustainable agriculture and global food security. *Genetics*, 188 (1), 11–20.
- Rosenzweig, C., Elliott, J., Deryng, D., Ruane, A. C., Müllere, C., Arnethf, A., Boote, K. J., Folberth, C., Glotter, M., Khabarov, N., Neumann, K., Pionteke, F., Pughf, T. A. M., Schmidm, E., Stehfestk, E., Yang, H., Jones, J. W., 2014. Assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison. *Proceedings of the National Academy of Sciences*, 111 (9), 3268–3273.
- Rowley, J. E., 2007. The wisdom hierarchy: representations of the DIKW hierarchy. *Journal of information science*.

- SAEON (South African Environmental Observation Network). Available online: www.saeon.ac.za (accessed pm Sep. 8, 2015).
- Sayer, J., Sunderland, T., Ghazoul, J., Pfund, J. L., Sheil, D., Meijaard, E., Venter, M., Boedhihartono, A. K., Day, M., Garcia, C., van Oosten, C., 2013. Ten principles for a landscape approach to reconciling agriculture, conservation, and other competing land uses. *Proceedings of the National Academy of Sciences*, 110 (21), 8349–8356.
- Schaller, N., 1993. The concept of agricultural sustainability. *Agriculture, Ecosystems & Environment*, 46 (1), 89–97.
- Schmidt, M. W., Torn, M. S., Abiven, S., Dittmar, T., Guggenberger, G., Janssens, I. A., Kleber, M., Kögel-Knabner, I., Lehmann, J., Manning, D. A. C., Nannipieri, P., Rasse, D. P., Weiner, S., Trumbore, S. E., 2011. Persistence of soil organic matter as an ecosystem property. *Nature*, 478 (7367), 49–56.
- Schreinemachers, P., Berger, T., 2011. An agent-based simulation model of human–environment interactions in agricultural systems. *Environmental Modelling & Software*, 26 (7), 845–859.
- Seufert, V., Ramankutty, N., Foley, J. A., 2012. Comparing the yields of organic and conventional agriculture. *Nature*, 485 (7397), 229–232.
- Shang, J., Liu, J., Ma, B., Zhao, T., Jiao, X., Geng, X., Huffman, T., Kovacs, J. M., Walters, D., 2015. Mapping spatial variability of crop growth conditions using RapidEye data in Northern Ontario, Canada. *Remote Sensing of Environment*, 168, 113–125.
- Shao, H., Liu, M., Shao, Q., Sun, X., Wu, J., Xiang, Z., Yang, W., 2014. Research on eco-environmental vulnerability evaluation of the Anning River Basin in the upper reaches of the Yangtze River. *Environmental Earth Sciences*, 72 (5), 1555–1568.
- Siciliano, G., 2012. Urbanization strategies, rural development and land use changes in China: A multiple-level integrated assessment. *Land Use Policy*, 29 (1), 165–178.
- SIGEO (Smithsonian Institution Global Earth Observatories). Available online: www.sigeo.si.edu (accessed pm Sep. 8, 2015).
- Sigrimis, N., Hashimoto, Y., Munack, A., De Baerdemaeker, J., 1999. Prospects in agricultural engineering in the information age: technological developments for the producer and the consumer. *Agricultural Engineering International: CIGR Journal I*, 1–20.

- Smil, V., 1999. China's agricultural land. *The China Quarterly*, 158, 414–429.
- Smith, C. J., Patrick, Jr W. H., 1983. Nitrous oxide emission as affected by alternate anaerobic and aerobic conditions from soil suspensions enriched with $(\text{NH}_4)_2\text{SO}_4$. *Soil Biology & Biochemistry*, 15, 693–696.
- Song, W., Keller, J. M., Haithcoat, T. L., Davis, C. H., 2009. Automated Geospatial Conflation of Vector Road Maps to High Resolution Imagery. *IEEE Transactions on Image Processing*, 18 (2), 388–400.
- Song, X., Zhang, B., Zhang, Y., Han, D., 2010. The harmonious ideal land-natural environment and water resources development in Sanjiang plain, Northeast China. *ICCS Journal of Modern Chinese Studies*, 2 (1), 263–275 (written in Chinese).
- Stafford, J. V., 2000. Implementing precision agriculture in the 21st century. *Journal of Agricultural Engineering Research*, 76 (3), 267–275.
- Stefanov, W. L., Ramsey, M. S., Christensen, P. R., 2001. Monitoring urban land cover change: An expert system approach to land cover classification of semiarid to arid urban centers. *Remote Sensing of Environment*, 77 (2), 173–185.
- TERN (Terrestrial Ecosystem Research Network). Available online: www.tern.org.au (accessed pm Sep. 8, 2015).
- Tilman, D., Balzer, C., Hill, J., Befort, B. L., 2011. Global food demand and the sustainable intensification of agriculture. *Proceedings of the National Academy of Sciences*, 108 (50), 20260–20264.
- van Ittersum, M. K., Ewert, F., Heckelei, T., Wery, J., Olsson, J. A., Andersen, E., Bezlepkina, I., Brouwer, F., Donatelli, M., Flichman, G., Olsson, L., Rizzoli, A. E., van der Wal, T., Wien, J. E., Wolf, J., 2008. Integrated assessment of agricultural systems—A component-based framework for the European Union (SEAMLESS). *Agricultural Systems*, 96 (1), 150–165.
- van Cauwenbergh, N., Biala, K., Bielders, C., Brouckaert, V., Franchois, L., Cidat, V. G., Hermy, M., Mathijs, E., Muys, B., Reijnders, J., Sauvenier, X., Valckx, J., Vanclooster, M., van der Veken, B., Wauters, E., Peeters, A., 2007. SAFE—A hierarchical framework for assessing the sustainability of agricultural systems. *Agriculture, Ecosystems & Environment*, 120 (2), 229–242.

- van Niel, T. G., McVicar, T. R., 2004. Current and potential uses of optical remote sensing in rice-based irrigation systems: A review. *Crop & Pasture Science*, 55, 155–185.
- Verburg, P. H., Neumann, K., Nol, L., 2011. Challenges in using land use and land cover data for global change studies. *Global Change Biology*, 17 (2), 974–989.
- Vernier, F., Miralles, A., Pinet, F., Carluer, N., Gouy, V., Molla, G., Petit, K., 2013. EIS Pesticides: An environmental information system to characterize agricultural activities and calculate agro-environmental indicators at embedded watershed scales. *Agricultural Systems*, 122, 11–21.
- Vicente-Guijalba, F., Martinez-Marin, T., Lopez-Sanchez, J. M., 2015. Dynamical Approach for Real-Time Monitoring of Agricultural Crops. *IEEE Transactions on Geoscience and Remote Sensing*, 53 (6), 3278–3293.
- Vrieling, A., 2006. Satellite remote sensing for water erosion assessment: A review. *Catena*, 65 (1), 2–18.
- Vörösmarty, C. J., Green, P., Salisbury, J., Lammers, R. B., 2000. Global water resources: vulnerability from climate change and population growth. *Science*, 289 (5477), 284–288.
- Wahid, S. M., Babel, M. S., Gupta, A. D., Routray, J. K., Clemente, R. S., 2008. Degradation–environment–society spiral: A spatial auto-logistic model in Thailand. In *Natural Resources Forum*, 32, (4), 290–304.
- Waldhoff, G., 2014. Multidaten-Ansatz zur fernerkundungs-und GIS-basierten Erzeugung multitemporaler, disaggregierter Landnutzungsdaten (Doctoral dissertation, Universität zu Köln).
- Waldhoff, G., Bareth, G., 2009. GIS-and RS-based land use and land cover analysis: Case study Rur-Watershed, Germany. In *Geoinformatics 2008 and Joint Conference on GIS and Built environment: Advanced Spatial Data Models and Analyses*. International Society for Optics and Photonics, 7146, 714626-1–714626-8.
- Wang, G., Liu, J., Wang, J., Yu, J., 2006. Soil phosphorus forms and their variations in depressional and riparian freshwater wetlands (Sanjiang Plain, Northeast China). *Geoderma*, 132 (1), 59–74.

- Wang, J., Lu, X., Jiang, M., Li, X., Tian, J., 2009. Fuzzy synthetic evaluation of wetland soil quality degradation: A case study on the Sanjiang Plain, Northeast China. *Pe-dosphere*, 19 (6), 756–764.
- Wang, L., Shi, Y., Ma, J., 2011. Rice standard production regime in albic soil area. *North Rice* 41 (5), 49–50, 56 (in Chinese).
- Wang, X., Lei, J., Yu, C., Xiao, Y., Wang, Z., Zhou, Z., Li, M., 2013. Comparative study on rice eating quality with parameters-related differences between *indica* and *japonica* rice subspecies. *Molecular Plant Breeding*, 4.
- Wang, Z., Song, K., Ma, W., Ren, C., Zhang, B., Liu, D., Chen, J., Song, C., 2011. Loss and fragmentation of marshes in the Sanjiang Plain, Northeast China, 1954–2005. *Wet-lands*, 31 (5), 945–954.
- Wang, Z., Zhang, B., Zhang, S., Li, X., Liu, D., Song, K., Li, J., Li, F., Duan, H., 2006. Changes of land use and of ecosystem service values in Sanjiang Plain, Northeast China. *Environmental Monitoring and Assessment*, 112 (1-3), 69–91.
- Wang, Y., Wang, S., Yang, S., Zhang, L., Zeng, H., Zheng, D., 2014. Using a Remote Sensing Driven Model to Analyze Effect of Land Use on Soil Moisture in the Weihe River Basin, China. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7 (9), 3892–3902.
- Wang, Y., Chang, K., Chen, R., Lo, J., Shen, Y., 2010. Large-area rice yield forecasting using satellite imageries. *International Journal of Applied Earth Observation*, 12, 27–35.
- Wang, Z., Mao, D., Li, L., Jia, M., Dong, Z., Miao, Z., Ren, C., Song, C., 2015. Quantifying changes in multiple ecosystem services during 1992–2012 in the Sanjiang Plain of China. *Science of the Total Environment*, 514, 119–130.
- Wentz, E. A., Nelson, D., Rahman, A., Stefanov, W. L., Roy, S. S., 2008. Expert system classification of urban land use/cover for Delhi, India. *International Journal of Re-mote Sensing*, 29 (15), 4405–4427.
- Wu, C., Jin, C., Xiao, T., Zhang, M., Yuan, W., 2007. Analysis on Rape Mechanization Pre-sent Situation and Technical Affection Factors in Whole Productive Course in China. *Journal of Agricultural Mechanization Research*, 12, 061 (in Chinese with English abstract).

- Xiao, X., Boles, S., Liu, J., Zhuang, D., Frohling, S., Li, C., Salas, W., Moore III, B., 2005. Mapping paddy rice agriculture in southern China using multi-temporal MODIS images. *Remote Sensing of Environment*, 95, 480–492.
- Xu, Y., Yan, B., Tang, J., 2015. The Effect of Climate Change on Variations in Dew Amount in a Paddy Ecosystem of the Sanjiang Plain, China. *Advances in Meteorology*, 1–9.
- Yan, M., Deng, W., Chen, P., 2002. Climate change in the Sanjiang Plain disturbed by large-scale reclamation. *Journal of Geographical Sciences*, 12 (4), 405–412.
- Yang, C., Everitt, J. H., Du, Q., Luo, B., Chanussot, J., 2013. Using high-resolution airborne and satellite imagery to assess crop growth and yield variability for precision agriculture. *Proceedings of the IEEE*, 101 (3), 582–592.
- Yao, Y., Miao, Y., Huang, S., Gao, L., Ma, X., Zhao, G., Jiang, R., Chen, X., Zhang, F., Yu, K., Gnyp, M. L., Bareth, G., Liu, C., Zhao, L., Yang, W., Zhu, H., 2012. Active canopy sensor-based precision N management strategy for rice. *Agronomy for Sustainable Development*, 32 (4), 925–933.
- Yu, K., Patrick, W. H. Jr., 2004. Redox window with minimum global warming potential contribution from rice soils. *Soil Science Society of America Journal*, 68, 2086–2091.
- Yuan, L., 1997. Hybrid rice breeding for super high yield. *Hybrid Rice*, 12 (6), 1–6.
- Zhang, J., Song, C., Wenyan, Y., 2007. Tillage effects on soil carbon fractions in the Sanjiang Plain, Northeast China. *Soil and tillage Research*, 93 (1), 102–108.
- Zhang, S., Wang, W., Li, Y., Bu, K., Yan, Y., 2008. Dynamics of hillslope soil erosion in the Sanjaing Plain in the past 50 years. *Resources Science*, 30 (6), 843–849 (written in Chinese with English abstract).
- Zhang, F., Chen, X., Vitousek, P., 2013. Chinese agriculture: An experiment for the world. *Nature*, 497 (7447), 33–35.
- Zhang, F., Shen, J., Li, L., Liu, X., 2004. An overview of rhizosphere processes related with plant nutrition in major cropping systems in China. *Plant and Soil*, 260 (1-2), 89–99
- Zhang, G., Xiao, X., Dong, J., Kou, W., Jin, C., Qin, Y., Zhou, Y., Wang, J., Menarguez, M.A., Biradar, C., 2015. Mapping paddy rice planting areas through time series analysis of MODIS land surface temperature and vegetation index data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 106, 157–171.

- Zhang, J., 2010. Multi-source remote sensing data fusion: status and trends. *International Journal of Image and Data Fusion*, 1 (1), 5–24.
- Zhang, L., Yu, D., Shi, X., Xu, S., Xing, S., Zhao, Y., 2014. Effects of soil data and simulation unit resolution on quantifying changes of soil organic carbon at regional scale with a biogeochemical process model. *PLoS ONE*, 9 (2), e88622.
- Zhang, L., Zhuang, Q., Zhao, Q., He, Y., Yu, D., Shi, X., Xing, S., 2016. Uncertainty of organic carbon dynamics in Tai-Lake paddy soils of China depends on the scale of soil maps. *Agriculture, Ecosystems & Environment*, 222, 13–22.
- Zhang, Y., Li, C., Zhou, X., Moore III, B., 2002. A simulation model linking crop growth and soil biogeochemistry for sustainable agriculture. *Ecological Modelling*, 151 (1), 75–108.
- Zhang, Y., Su, S., Zhang, F., Shi, R., Gao, W., 2012. Characterizing spatiotemporal dynamics of methane emissions from rice paddies in northeast China from 1990 to 2010. *Plo-SONE*, 7 (1), e29156.
- Zhao, Q., Brocks, S., Lenz-Wiedemann V.I.S., Miao, Y., Jiang, R., Chen, X., Zhang, F., and Bareth, G., 2012. Implementation of an agricultural environmental information system (AEIS) for the Sanjiang Plain, NE-China. *Proceedings of the XXII Congress of the ISPRS, Melbourne, Australia*.
- Zhao, Q., Hütt, C., Lenz-Wiedemann, V.I.S., Miao, Y., Yuan, F., Zhang, F., Bareth, G., 2015a. Georeferencing multi-source geospatial data using multi-temporal TerraSAR-X imagery: A case study in Qixing Farm, northeast China. *Photogrammetrie Fernerkundung Geoinformation*, 2, 173–185
- Zhao, Q., Lenz-Wiedemann, V.I.S., Yuan, F., Jiang, R., Miao, Y., Zhang, F., Bareth, G., 2015b. Investigating within-field variability of rice from high resolution satellite imagery in Qixing Farm County, Northeast China. *ISPRS International Journal of Geo-Information*, 4 (1), 236–261.
- Zhong, Y., Wang, Q., 2008. Investigating the current water resource status and the exploration strategies in the Sanjiang Plain. *Heilongjiang Science and Technology of Water Conservancy*, 6 (36), 133–136 (written in Chinese).
- Zhong, L., Gong, P., Biging, G. S., 2014. Efficient corn and soybean mapping with temporal extendability: A multi-year experiment using Landsat imagery. *Remote Sensing of Environment*, 140, 1–13.

- Zhou, D., Gong, H., Wang, Y., Khan, S., Zhao, K., 2009. Driving forces for the marsh wetland degradation in the Honghe National Nature Reserve in Sanjiang Plain, Northeast China. *Environmental Modeling & Assessment*, 14 (1), 101–111.
- Zhou, Z., Liu, T., 2005. The current status, threats and protection way of Sanjiang Plain wetland, Northeast China. *Journal of Forestry Research* 16, 148–152. doi:10.1007/BF02857910
- Zou, L., Guo, S., Niu, N., 2010. Interdecadal change of agroclimatic environment in Sanjiang Plain during 1960 – 2004. *Advances in Earth Science*, 25 (8), 932–938.
- Zhu, Z., Chen, D., 2002. Nitrogen fertilizer use in China—Contributions to food production, impacts on the environment and best management strategies. *Nutrient Cycling in Agroecosystems*, 63 (2-3), 117–127.
- Zwart, S. J., Bastiaanssen, W. G., 2004. Review of measured crop water productivity values for irrigated wheat, rice, cotton and maize. *Agricultural Water Management*, 69 (2), 115–133.

List of Figures and Tables

- **Figures**

Figure 2-1: The study area Sanjiang Plain in Heilongjiang province, NE-China (Zhao et al., 2015b).	18
Figure 2-2: The reclamation area in the SJP from 1949 to 1999 (Wang et al., 2006).	21
Figure 2-3: Surface water system and elevations in the Sanjiang Plain.....	22
Figure 2-4: Change of groundwater in the Jiansanjiang area, coupled with the increased rice cultivation area (Ren 2014).	23
Figure 3-1: Components of a knowledge-based system.	32
Figure 4-1: Location of the study area Qixing Farm in Northeast China.	39
Figure 4-2: Georeferencing workflow of the multi-source geospatial data, PE = positional error.	42
Figure 4-3: Field boundary data, before (yellow) and after (cyan) the georeferencing; red arrows in the left figure show the vector force of the rubber sheeting procedure. Background data in the right figure: TSX reference image.	45
Figure 4-4: An example of georeferenced multi-source RS images in comparison to the TSX image. From left to right, 1st row: FS-2, TSX, LS 5, HJ, RE, 2nd row: LS 5, HJ, RE, FS-2, TSX, 3rd row: RE, FS-2, TSX, LS 5, HJ, 4th row: TSX, LS 5, HJ, RE, FS-2, 5th row: HJ, RE, FS-2, TSX, LS 5.....	46
Figure 4-5: Georeferenced multi-source data for the study area of Qixing Farm.	49
Figure 5-1: Location of the study area in Northeast China (the upper left corner shows a subset of the FS-2 image acquired on 9 August 2009).	60
Figure 5-2: Relationships between RS derived values and ground truth values for (a) biomass; (b) LAI; (c) N concentration; and (d) N uptake at the tillering (Δ), booting (\diamond), and heading (\circ) stages.	72
Figure 5-3: Within-field spatial variability of biomass derived from the FS-2 image of 6 July 2009.	73
Figure 5-4: Within-field spatial variability of LAI derived from the FS-2 image of 6 July, 2009.	74
Figure 5-5: Within-field spatial variability of N concentration derived from the FS-2 image of 6 July 2009.	75
Figure 5-6: Within-field spatial variability of N uptake derived from the FS-2 image of 6 July 2009.	76
Figure 5-7: Rice growth on 24 June (left), 6 July (middle) and 9 August (right).	79

Figure 6-1: Location of the study area Qixing Farm in Heilongjiang province, NE-China..	91
Figure 6-2: Distribution of field data sites in 2009 in the Qixing Farm.	91
Figure 6-3: Workflow of deriving soil input raster files based on the measured point data and soil type raster file.	95
Figure 6-4: Map of rice yield modelled by the DNDC model and the within-field variability in the inset map.	100
Figure 6-5: DNDC modelled and field observed rice yields at all 57 validation sites.	101
Figure 6-6: Ratio map of yield difference (subtracting the RS-derived yield from the modelled yield) to the DNDC modelled yield.	103
Figure 6-7: Spatial patterns of estimated rice yield maps and soil data.	105
Figure 7-1: Time series data of DNDC modelled biomass vs. field observed values in the 'high yield' fields.	119
Figure 7-2: Time series data of DNDC modelled biomass vs. field observed values in the 'medium yield' fields.	119
Figure 7-3: Maps of DNDC modelled greenhouse gas emissions from paddy rice field in Qixing Farm.	121

- Tables

Table 2-1: Land use change (%) for selected classes from the 1980s until 2000 in the Sanjiang Plain for the representative counties (Zhao et al., 2012).	21
Table 4-1: Characteristics of the RS images.	41
Table 4-2: Accuracy of the selected GCPs (PE = positional error, Std. = standard deviation).	45
Table 4-3: Accuracy of the independent check points.	47
Table 5-1: Agronomic parameters of the sample sites in 2009.	63
Table 5-2: Main atmospheric correction parameters for the FS-2 images.	64
Table 5-3: Main georeferencing parameters for the FS-2 images.	65
Table 5-4: Accuracies of rice maps produced from different data sources.	68
Table 5-5: Parameters of empirical regression models.	69
Table 5-6: Validation results for the regression models.	71
Table 6-1: Transferred soil textures and selected soil physical properties in the study area.	96
Table 6-2: DNDC model input parameters (selected) for regional application.	97

Table 6-3: Parameters for the multiple linear regression model for the RS approach.	98
Table 6-4: Site-specific input data and performance of DNDC for site validation.	99
Table 6-5: Assessment indices in model regional validation.	101
Table 6-6: Assessment indices for soil-specific validation.	102
Table 6-7: Assessment indices for RS-derived yield.	102
Table 7-1: Confusion matrix of the land cover classification map based on a single date (August 9 th) RS image.	115
Table 7-2: Confusion matrix of the land cover classification map that was generated by integrating the three classifications from the three dates.	115
Table 7-3: Confusion matrix of the land cover classification map based on both the RS and GIS data.	115
Table 7-4: Accuracy indices for land use classification from different data sources.	116

Appendix A: Eigenanteil zu Kapitel 4

Titel	Georeferencing multi-source geospatial data using multi-temporal TerraSAR-X imagery: a case study in Qixing Farm, Northeast China
Autoren	Zhao, Q. , Hütt, C., Lenz-Wiedemann, V.I.S., Miao, Y., Yuan, F., Zhang, F. and Bareth, G.
Status	veröffentlicht
Verlag	PFG (Photogrammetrie Fernerkundung Geoinformation)
Publikationsjahr	2015
Band und Seitenangabe	2: 0173-0185
DOI	10.1127/pfg/2015/0262
Eigenanteil	<p>Organisation der Feldkampagnen vor Ort in China (März-Oktober 2009; März-September 2010; Juni-August 2011; Juni-Juli 2012)</p> <p>Datenerhebung im Gelände (in China) – Durchführung einer Feldkamapagne zur Erhebung der der Landnutzung und Landbedeckung (2009-2012)</p> <p>Datenbeschaffung - Kooperation mit Wissenschaftseinrichtungen zur Beschaffung von Daten aus verschiedenen Quellen darunter topographische Vektordaten, Download und Auswertung von Satellitendaten aus mehreren Quellen: Prozessierung optischer Satellitendaten (Landsat 5, FORMOSAT-2 und RapidEye) einschließlich der Provinz Huanjing, Auswahl von Kontrollpunkten aus TerraSAR-X Daten und Passpunkten aus optischen Daten, Ergebniskontrolle, Georeferenzierung der topographischen Vektordatensätze</p> <p>Entwicklung von Analysemethoden</p> <p>Schreiben des Manuskriptes</p> <p>Erstellung von Abbildungen und Tabellen</p> <p>Review/Korrekturlesen des Manuskriptes</p>

Appendix B: Eigenanteil zu Kapitel 5

Titel	Investigating Within-field Variability of Rice from High Resolution Satellite Imagery in Qixing Farm County, Northeast China
Autoren	Zhao, Q. , Lenz-Wiedemann, V.I.S., Yuan, F., Jiang, R., Miao, Y., Zhang, F. and Bareth, G.
Status	veröffentlicht
Verlag	IJGI (ISPRS International Journal of Geo-Information)
Publikationsjahr	2015
Band und Seitenangabe	4: 2361-261
DOI	10.3390/ijgi4010236
Eigenanteil	<p>Organisation der Feldkampagnen vor Ort in China (März-Oktober 2009)</p> <p>Datenerhebung im Gelände – Erhebung agronomischer Daten darunter Biomasseentnahme, Entnahme und Dreschen von Ertragsproben, leaf area index (LAI) Bestimmung, schneiden, wiegen und trocknen von Pflanzenproben, Entnahme von Bodenproben, Befragungen von Bauern zur Feldbewirtschaftung,</p> <p>Kooperation mit lokalen Wissenschaftseinrichtungen zur Beschaffung von Daten aus verschiedenen Quellen darunter topographische Vektordaten</p> <p>Laborarbeiten – Stickstoffanalyse, Bodenparameter</p> <p>Datenanalyse – Eingabe agronomischer Daten, Entwicklung eines Regressionsmodells, Prozessierung von Satellitendaten, Landnutzungsklassifikation</p> <p>Entwicklung von Analysemethoden</p> <p>Schreiben des Manuskriptes</p> <p>Erstellung von Abbildungen und Tabellen</p> <p>Review/Korrekturlesen des Manuskriptes</p>

Appendix C: Eigenanteil zu Kapitel 6

Titel	Detecting spatial variability in paddy rice yield using the DNDC model and satellite images in the Sanjiang Plain, NE-China
Autoren	Zhao, Q. , Brocks, S., Lenz-Wiedemann, V.I.S., Miao, Y., Zhang, F. and Bareth, G.
Status	liegt in Manuskriptform vor
Verlag	
Publikationsjahr	
Band und Seitenangabe	
DOI	
Eigenanteil	<p>Organisation der Feldkampagnen vor Ort in China (März-Oktober 2009; März-September 2010; Juni-August 2011; Juni-Juli 2012)</p> <p>Datenerhebung im Gelände – Erhebung agronomischer Daten darunter Biomasseentnahme, Entnahme und Dreschen von Ertragsproben, leaf area index (LAI) Bestimmung schneiden, wiegen und trocknen von Pflanzenproben, Entnahme von Bodenproben, Befragungen von Bauern zur Feldbewirtschaftung,</p> <p>Kooperation mit lokalen Wissenschaftseinrichtungen zur Beschaffung von Daten aus verschiedenen Quellen darunter topographische Vektordaten</p> <p>Laborarbeiten – Analyse von Wachstumsparameter, Ertragsanalysen</p> <p>Datenanalyse – Auswertung von Klimadaten, Interpolation von Bodenkarten, Auswertung der Daten zur Feldbewirtschaftung</p> <p>Dateneingabe und Verarbeitung der agronomischen Parameter DNDC crop model Kalibrierung, Ergebnisanalyse, Sensitivitätstest</p> <p>Analyse von Satellitendaten und Ertragsvorhersage</p> <p>Entwicklung von Analysemethoden</p> <p>Schreiben des Manuskriptes</p> <p>Erstellung von Abbildungen und Tabellen</p> <p>Review/Korrekturlesen des Manuskriptes</p>

Appendix D: Erklärung

Ich versichere, dass ich die von mir vorgelegte Dissertation selbständig angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit – einschließlich Tabellen, Karten, und Abbildungen –, die anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie - abgesehen von unten angegebenen Teilpublikationen - noch nicht veröffentlicht worden ist sowie, dass ich eine solche Veröffentlichung vor Abschluss des Promotionsverfahrens nicht vornehmen werde.

Die Bestimmungen der Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Prof. Dr. Georg Bareth betreut worden.

Köln, den

Folgende Teilpublikationen liegen vor:

Appendix E: Curriculum Vitae

Name: Quanying Zhao;

Born: April, 1984;

Place of born: Jilin, China;

PhD study: Nov. 2010 – Jan. 2016, major in Physical Geography, Institute of Geography, University of Cologne, Germany;

Master study: Sep. 2008 – Jun. 2010, major in Environmental Sciences, College of Environmental Sciences, China Agricultural University, Beijing, China;

Bachelor study: Sep. 2004 – Jun. 2008, major in Environmental Engineering, College of Environmental Sciences, China Agricultural University, Beijing, China.

Appendix F: Publication list

* = part of dissertation

- ***Q. Zhao**, S. Brocks, V. Lenz-Wiedemann, Y. Miao, F. Zhang, G. Bareth. Detecting Spatial Variability of Paddy Rice Yield by Combining the DNDC Model with High Resolution Satellite Images. *Agricultural Systems* (Under review).
- ***Q. Zhao**, V.I.S. Lenz-Wiedemann, G. Bareth, F. Yuan, R. Jiang, Y. Miao, F. Zhang. Investigating within-field variability of rice from high resolution satellite imagery on county level. *ISPRS International Journal of Geo-Information*, 2015, 4(1), 236-261, doi:10.3390/ijgi4010236.
- ***Q. Zhao**, C. Hütt, V.I.S. Lenz-Wiedemann, Y. Miao, F. Yuan, F. Zhang, G. Bareth. Georeferencing multi-source geospatial data using multi-temporal TerraSAR-X imagery: a case study in Qixing Farm, northeast China. *Photogrammetrie, Fernerkundung, Geoinformation*, 2015, 2, 0173-0185, doi: 10.1127/pfg/2015/0262.
- Q. Zhao**, S. Brocks, V.I.S. Lenz-Wiedemann, Y. Miao, G. Bareth. Regional application of the site-specific biochemical process-based crop model DNDC for rice in NE-China. *Proceedings of the 35 Wissenschaftlich-Technische Jahrestagung der DGPF*, 2015, Cologne, Germany.
- Q. Zhao**, S. Brocks, V. Lenz-Wiedemann, Y. Miao, R. Jiang, X. Chen, F. Zhang, G. Bareth. Implementation of an agricultural environmental information system (AEIS) for the Sanjiang Plain, NE-China. *Proceedings of the XXII Congress of the ISPRS*, 2012, Melbourne, Australia.
- Q. Zhao**, X. Feng, L. Tang, Z. Ma. Performance of a natural aeration enhanced artificial soil rapid infiltration system for removing nitrogen and phosphorus in sewage. *Occupational Circle*, 2007 (23), 130-132 (in Chinese).
- L. Zhang, Q. Zhuang, **Q. Zhao**, Y. He, D. Yu, X. Shi, S. Xing. Uncertainty of organic carbon dynamics in Tai-Lake paddy soils of China depends on the scale of soil maps. *Agriculture, Ecosystems & Environment*, 2016 (222), 13-22.
- L. Zhang, Q. Zhuang, X. Li, **Q. Zhao**, D. Yu, Y. Liu, X. Shi, , S. Xing, G. Wang. Carbon sequestration in the uplands of Eastern China: an analysis with high-resolution model simulations. *Soil & Tillage Research*. 2016, 158, 165-176.

- L. Zhang, Q. Zhuang, Y. He, Y. Liu, D. Yu, **Q. Zhao**[#], X. Shi, S. Xing, G. Wang. Toward optimal Soil Organic Carbon sequestration with effects of Agricultural Management Practices and Climate Change in Tai-Lake paddy soils of China. *Geoderma* (in review) ([#]co-corresponding author).
- M. Su, S. Huang, G. Zhao, **Q. Zhao**, Y. Yao, Y. Miao. Current rice management practices of farmers in Heilongjiang land reclamation area and improvement strategies. *North Rice*, 2012, 42 (2), 28-33 (in Chinese with English abstract).