2013

A Land Data Assimilation System (ldas) Based Dataset For Regional Agro-Climatic Assessments

Xing Liu
Purdue University, liuxing0901@gmail.com

Graduate School ETD Form 9
(Revised 12/07)

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By Xing Liu

Entitled
A LAND DATA ASSIMILATION SYSTEM (LDAS) BASED DATASET FOR REGIONAL AGRO-CLIMATIC ASSESSMENTS

For the degree of Master of Science

Is approved by the final examining committee:

DEV NIYOGI
Chair

LARRY L. BIEHL

HAO ZHANG

YIWEI JIANG

JEFFREY A. ANDRESEN

To the best of my knowledge and as understood by the student in the *Research Integrity and Copyright Disclaimer (Graduate School Form 20)*, this thesis/dissertation adheres to the provisions of Purdue University's "Policy on Integrity in Research" and the use of copyrighted material.

Approved by Major Professor(s): DEV NIYOGI

Approved by: Joseph M. Anderson 11/24/2013
Head of the Graduate Program Date

A LAND DATA ASSIMILATION SYSTEM (LDAS) BASED DATASET FOR REGIONAL AGRO-CLIMATIC ASSESSMENTS

A Thesis

Submitted to the Faculty

of

Purdue University

by

Xing Liu

In Partial Fulfillment of the

Requirements for the Degree

of

Master of Science

December 2013

Purdue University

West Lafayette, Indiana

To the memory of my two grandfathers: Delu Ma and Changhua Liu.

ACKNOWLEDGEMENTS

I would like to express my sincere thanks to my advisor Porf. Dev Niyogi for supporting my research, for his patience and guidance. I would like to thank the rest of my thesis committee: Larry Biehl, Prof. Jeff Andresen, Prof. Hao Zhang and Prof. Yiwei Jiang for their insightful comments and helpful advice.

My sincere thanks also goes the U2U project members: Pengxuan Zheng, Melissa Widhalm, Carol Song and Lan Zhao, for sharing the data with me and helping a lot on programming.

I thank my labmates: Subashini Subramanian, Elin Karlsson, Anil Kumar, Ani Elias, Olivia Kellner and Yue Zhen, for their continuous support.

I would like to thank staff of Agronomy, for their generous help during the two years.

At last, I would like to thank my friends and families, for the support, for the trust.

TABLE OF CONTENTS

Page

Page

Page

Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

Figure Page

ABSTRACT

Liu, Xing. M.S., Purdue University, December 2013. A Land Data Assimilation System (LDAS) Based Dataset for Regional Agro-climatic Assessment. Major Professor: Dev Niyogi.

This study is part of a USDA sponsored project ----Useful to Usable (U2U): "Transforming Climate Variability and Change Information for Cereal Crop Producers". The broader objective includes improving farm resilience and profitability in the U.S. Corn Belt region by transforming existing climate/weather data into usable knowledge and tools for the agricultural community.

The specific tasks of this research are: (1) Build a high-resolution (4 km, daily) agro-climatic dataset using a Land Data Assimilation System (LDAS). (2) Estimate regional corn yield across the Corn Belt with crop models and the agro-climatic dataset. (3) Evaluate the impacts of climate variability due to El Niño–Southern Oscillation (ENSO) on corn yield in the Corn Belt.

Accordingly, a high-resolution (4 km, 1979-2012, daily) agro-climatic dataset across the U.S. Corn Belt has been built using the North America Land Data Assimilation System version 2 (NLDAS2) product. This newly developed dataset includes daily maximum/minimum temperature, precipitation, solar radiation, soil moisture, and soil temperature at four soil depths (0-10 cm, 10-40 cm, 40-100 cm, and 100-200 cm). Validations indicate strong agreement between this dataset and field measurements. The

agro-climatic dataset was then used with a Hybrid-Maize crop model to estimate regional corn yield at grid scale. The crop model was first validated at the field and county scale and found to consistently overestimate yields at the county scale. This was attributed to the optimum field conditions considered in the model and the overall uncertainties. Comparison with NASS yield survey data indicates a 0.6 multiplicative factor provides good agreement with actual yields, and is recommended for county-scale simulations.

Following the field/county scale model tests, a modeling framework was developed to simulate gridded crop yields. Results indicate that integrating spatial climatic information improved the regional performance of the Hybrid Maize model and this agro-climatic dataset shows good potential for developing agro-meteorological related applications. Finally, the impacts of the El Nino-Southern Oscillation (ENSO) on observed and simulated corn yields were examined. As a result, La Niña shows a significant negative impact on corn yield in the Corn Belt while the impact from El Niño is insignificant. It also has been found that La Niña correlates with relatively late planting dates in the Corn Belt. Based on a crop model study, the results indicate that for some counties, under optimal conditions, late planting dates can mitigate the negative impacts from the La Niña phase.

Based on the studies above, reliable performance of the Hybrid Maize crop model and superior data ability of the new agro-climatic dataset have good potential to simulate regional corn yield with climate projections. The significant impacts of ENSO on corn yield indicate that advance ENSO warning may benefit field management in the Corn Belt.

INTRODUCTION

The U.S. Corn Belt produces nearly one-third of the global corn supply and contributes 100 billion dollars annually to the economy. Weather conditions and climate variability have a great influence during the crop growing season. Maintaining the stability of corn production under climate variability becomes more and more important, as well as increasing the corn potential yield and narrowing the yield gap. Providing high-resolution weather-related agronomic information can help producers/researchers to make better field management decisions.

This research is part of the NSF-USDA Useful to Usable (U2U) project, which is described as: “Transforming Climate Variability and Change Information for Cereal Crop Producers, is an integrated research and extension project working to improve farm resilience and profitability in the North Central Region by transforming existing climate information into usable knowledge for the agricultural community (www.Agclimate4U.org).”

The major objectives of this sub-research are:

- Provide a high-resolution agro-meteorological database of the U.S. Corn Belt.
- Link the high-resolution meteorological data products with corn simulation models to estimate corn yield at a different spatial scale across the U.S. Corn Belt

- Evaluate the impacts of climate variability on the U.S. Corn Belt
- Combine future weather/climate predictions with corn yield simulation.

For this research, the Hybrid-Maize model was selected as the main crop model during the corn yield estimation process. The Land Data Assimilation System (LDAS) was used as the major reanalysis meteorological large raw data product. The El Niño–Southern Oscillation (ENSO) has been analyzed as the climate variability for possible impacts on corn yield. The hypotheses were:

1) The Hybrid-Maize model can provide reliable yield estimations at both the regional scale and field scale.
2) The meteorological data products from the Land Data Assimilation are reliable for applying in corn yield simulation.
3) The effects of the El Niño–Southern Oscillation (ENSO) on corn planting date and corn yield are significant.
4) The bias range of future yield prediction is acceptable.

This regional research will span over 30 years (1981-2010) and across 20 sites (Fig. 1.1) located in the U.S. Corn Belt, the site selections were based on the representative value of these sites and the date availability. Data sources are listed in Table 1.1. Detailed methodologies are presented in separate chapters. The main topic of each chapter are as follow:

Chapter 1 provides the overall research background.

Chapter 2 describes the process and results of the sensitivity analysis and model validation of the Hybrid-Maize model across the U.S. Corn Belt.

Chapter 3 presents the process of building a 4-km resolution agro-meteorological database based on the Land Data Assimilation System (LDAS) and High Resolution Land Data Assimilation System (HRLDAS). The process and results of running the crop model at gridded scale also described.

Chapter 4 reports the effects of the El Niño–Southern Oscillation (ENSO) on corn planting date and corn yield. A crop model-based study on the impacts of alternating planting date on corn yield under different ENSO phases is also included.

Chapter 5 discusses a preliminary study on simulating corn yield using climate model projected weather data.

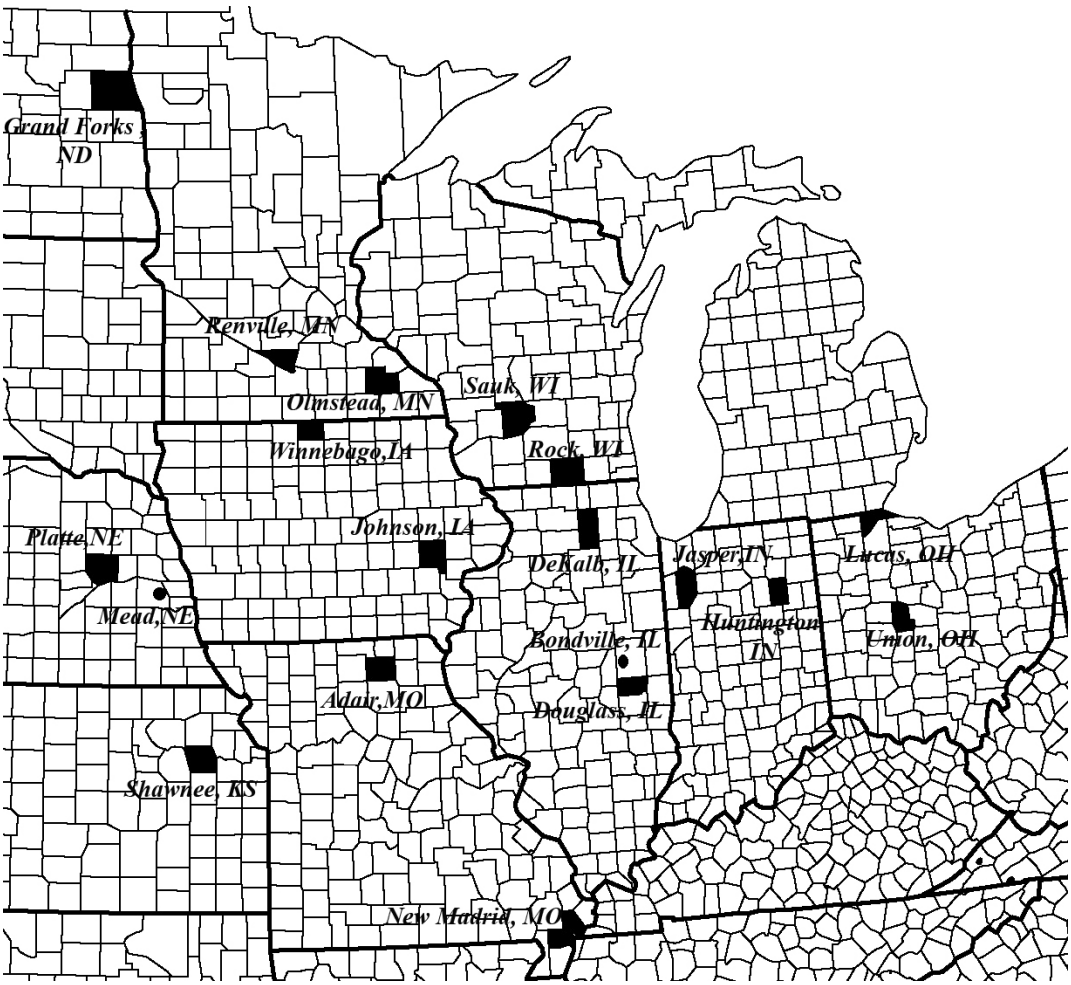

Figure 1.0 Research area: County-scale simulation sites and two field-scale sites (black dots): Bondville, IL (40.00°N, 88.29°W) and Mead, NE (41.18°N, 96.44°W)

Table 1.0 Major data sources used in this research (detailed information was provided in separate chapters)

Data	Source	Period
Reanalysis contemporary meteorological data (e.g., air temperature, solar radiation, precipitation, etc.)	Phase 2 of the North American Land Data Assimilation System (NLDAS-2)	1979-2012
On-site meteorological data of 18 county-level sites	National Climatic Data Center (NCDC)	1981-2010
On-site meteorological data of 2 field-level sites	Ameriflux	Mead, NE: 2002-2006 Bondville, IL: 1997-2007
Future projected data of Bondville, IL	North American Regional Climate Change Assessment Program (NARCCAP)	1979-2003
Corn yield of 18 county-level sites	National Agricultural Statistics Service（NASS）	1981-2010
Corn yield of 2 field-level sites	Ameriflux	2001, 2003, 2005
Planting date of 9 states in the U.S. Corn Belt	National Agricultural Statistics Service（NASS）	1994-2010

CHAPTER 1. INTRODUCTION TO CROP MODEL, LAND DATA ASSIMILATION SYSTEM (LDAS), AND EL NIÑO–SOUTHERN OSCILLATION (ENSO)

1.1 Introduction to the crop model

1.1.1 What is crop model

The model is a description of operations in a system structured by interacted objects (Soltani and Sinclair, 2012). Objects are elemental unit-based on the observations (Haefner, 2005). Depending on the way the systems are described, models can be classified into four groups (Haefner, 2005):

1. Conceptual or verbal models—describe the operations of a system in natural common language. For example, the paragraphs in a textbook or web page which describes the carbon cycle.
2. Diagrammatic models--- graphically describe the operations of a system. For example, the “box-and-arrow” diagrams of the carbon cycle.
3. Physical models --- physical mock-up of the system and the objects. For example, a car model or globe.
4. Mathematical models ---mathematically describe the operations and relations.

Crop growth is driven by carbon assimilation, plant development, respiration, and plant transpiration. Solar radiation influences the growth rate while temperature decides the growth duration (de Wit, 1978; Goudriaan and van Laar, 1994). Water stress and nitrogen stress limit leaf growth and biomass accumulation (Brisson et al., 2003). Crop models are mathematical models that use equations to describe the crop growth eco-physiological processes and development response to environmental variability and agricultural management.

To describe a model, there are several critical common terms (Soltani and Sinclair, 2012):

- Modelling: the process of developing a model.
- Simulation: "running" a model to get output values. For example, the process of running a crop model to obtain the yield output called "simulation".
- System analysis: analyze the output from the simulations and then draw conclusions.

Crop models, as well as other mathematical models, can be grouped into different categories (Soltani and Sinclair, 2012; Haefner, 2005):

- Process-oriented or descriptive: a process-oriented model has explicit representations of mechanistic processes while a descriptive model is more empirical. Process-oriented models can become descriptive models at lower organization levels.
- Static or dynamic: depends on whether the model accounts for the element of time. A dynamic model has an explicit representation of future system conditions.

- Continuous or discrete: a continuous model can take any value (e.g., 0.5 day) while a discrete model takes integers only (e.g., 5 days).
- Deterministic or stochastic: a stochastic model allows for random events and variables are described by probability distributions instead of unique values.

Generally, most crop models used in recent years are descriptive, dynamic, discrete, and deterministic.

Crop models are built by equations, which include the amount of variables. When describing a crop model, these variables can be grouped into three forms (Goudriaan and van, 1994; Brun et al., 2006; Soltani and Sinclair, 2012):

1. State variables: state variables illustrate the current status of the system. In a crop model, the state variables usually include yield, biomass, and leaf area index, etc. Equations in the crop model describe the evolution of state variables.
2. Parameters: variables represent the characteristics of a system, which usually keep constant values across simulations of interest. For example, in a crop model, the parameters include initial light use efficiency, growth respiration rate, and kernel filling rate, etc.
3. Explanatory variables: also known as “driving variables”, they enter into the equations to help calculate the state variables. They are usually environmental variables, and in crop models, they typically include temperature and solar radiation as well as the management variables.

For the commonly used crop models, explanatory variables are usually considered as "input" for the model, while state variables are considered as "output".

Rabbinge (1993) classifies the crop production into three situations --- Potential production: limited by solar radiation and temperature; Attainable production: adds influences from water, nitrogen, and phosphorus; Actual production: considers the possible yield reduction resulting from weeds, pests, and disease. Therefore, three themes of crop models were characteristic (Rabbinge and Kropff, 2008):

1. Basic biophysical, physiological processes of crop growth.
2. The influences of water-stress and nutrition-stress on crop growth.
3. The influence of weeds, pests, and diseases on crop yield.

Dynamic crop models were developed in the 1960s by de Wit (1966), and through more than 45 years of development, crop models have been used to support theoretical research, crop management, education, and policy analysis (Hammer et al., 2002). All crop models must simulate crop growth and development, biomass translocating from leaves to other organs, and yield (Yang et al., 2004)

Based on the target simulated crop species, crop models can be divided into generic crop models and specific crop models. Generic simulation models describe the crop growth regardless of the crop species, and then modifies to simulate the phonological and physiological traits of selected crops (Yang et al., 2004). Such models include SUCROS, WOFOST, and INTERCOM (Van Ittersum at al., 2003), STICS (Bryson et al., 2003), and EPIC (Jones et al., 1991). Specific crop models have been developed to simulate a

specific crop, such as DSSAT (Jones et al., 2003) and the Hybrid-Maize model (Yang et al., 2004, 2006).

Specific crop models and generic crop models are different in the theoretical development stage and model driving schemes. For example, CERES-MAIZE (Jones et al., 1986), a corn specific model, has five growth stages: emergence---end of juvenile stage, tassel initiation, silking, effective grainfilling, and maturity. INTERCOM (Kropff and van Laar, 1993) has only two phases: from emergence to anthesis and then from anthesis to maturity. Specific models are mainly driven by temperature and solar radiation while generic models are primarily driven by the availability of carbon assimilation.

Crop models were developed for different objectives. Some are for scientific research while others are more suited for decision support; therefore, some models are complex while other are relative simple. However, it is improper to say the complexity of a crop model represents the reliability of the simulation. The complexity of a crop model represents the amount of equations and parameters, which means collecting the data of parameters and driving variables is a major problem of crop models. Therefore, when selecting the model, it is important to consider the study objectives and the data availability.

Crop model simulations are usually constrained by collecting the input and calibrating the parameters, such as shortwave solar radiation, soil conditions, and kernel filling rate. Except for some controlled scientific research fields, those input data are difficult to obtain.

Since this research is designed to run a crop model at the regional scale, a relatively simpler model requiring less input data and fewer parameter calibrations is preferred. Therefore, for this research, the Hybrid-Maize model was selected.

1.1.2 The Hybrid-Maize Model

The Hybrid-Maize model was developed by combining the advantages of existing models. This model combined the attributes related to phenology from CERES-Maize (Jones et al., 1986) and the attributes related to organ growth from assimilated-driven models. The objective of developing this model is to simulate the potential corn yield and sensitivity to climatic variability (Yang et al., 2004).

The Hybrid-Maize model requires three groups of input data: crop and management, weather, and soil (Fig.1.2). Crop and management data include corn maturity (in total growing degree days, or GDD), plant date, and plant population. For simulations under optimal water management (i.e., non-water limiting) of yield potential, required weather data includes daily minimum and maximum air temperature (˚C), daily sum of global radiation (MJ/m^2) , and no soil data is required. For rainfed conditions, the model also requires daily precipitation (mm), daily average air humidity, and reference evapotranspiration (ET, mm), and basic soil information including texture of topsoil and subsoil, bulk density of topsoil, and soil moisture conditions at planting.

In past studies, the Hybrid-Maize model has demonstrated reliable performance in simulations and has shown considerable responsiveness to changing environmental conditions (Yang et al. 2004, 2006; Grassini et al., 2009).

1.2 Introduction to the Land Data Assimilation System (LDAS) and the NCAR High-Resolution Land Data Assimilation System (HRLDAS)

1.2.1 Land Data Assimilation System

Traditionally, crop models usually run using weather station data, which are accurate and easy to access. However, weather station data are not spatially continuous and lack the key input data for crop models --- solar radiation. In the U.S., most weather stations provide air temperature and precipitation while the solar radiation is only available from a small percentage of weather stations (Bristow and Campbell, 1984; Meinke et al., 1995; Goodin et al., 1999; Mahmood and Hubbard, 2002; Grant et al., 2004).

In previous modeling studies, solar radiation is usually estimated from a weather generator based on the location, precipitation, and temperature, such as Weather Generator (WGEN) (Richardson, 1981; Richardson and Wright, 1984), Simulation of Meteorological Variable (SIMMETEO; Geng et al., 1988), and the Weather Generator for Solar Radiation (WGRNR) (Hodges et al., 1985). However, some generators require detailed location-specific information which is not generally available (Grant et al., 2004) and data preparations are also time-consuming and require intensive computations for regional study.

Because of the limitations of using weather-station data in crop model simulations, the Land Data Assimilation System (LDAS, Fig. 1.3), which provides spatial and temporal continuous weather data including solar radiation, was selected for this study.

The Land Data Assimilation System (LDAS) consists of land-surface models (LSM) forced with precipitation gauge observations, satellite data, radar precipitation measurements, and output from numerical prediction models. The goal of LDAS is using the model results (e.g., soil moisture, evapotranspiration) to support water-resource applications, numerical weather prediction studies, etc. This system has been run at $1/8^{th}$-degrees resolution across central North America from January 1979 till near real-time (http://ldas.gsfc.nasa.gov/).

The land-surface models (LSM) in the Land Data Assimilation System (LDAS) including Mosaic (Koster and Suarez, 1992, 1996), Noah (Chen et al., 1996; Koren et al., 1999; Ek et al., 2003; Mitchell et al., 2004; Niu et al., 2011; Xia et al., 2012), Sacramento (SAC; Burnash et al., 1973; Anderson, 1973; Anderson et al., 2006) and Variable Infiltration Capacity (VIC; Liang et al., 1994,1996; Wood et al., 1997). The forcing data product in this system includes: Global Land Data Assimilation System (GLDAS) forcing, Phase 1 of the North American Land Data Assimilation System (NLDAS-1) forcing, and Phase 2 of the North American Land Data Assimilation System (NLDAS-2) forcing. (http://ldas.gsfc.nasa.gov).

Because of different characteristics in the four land-surface models mentioned previously, such as different model parameterizations, even though they used the same input forcing file, the outputs from each model are not the same. Dirmeyer et al. (2006) indicates that the means of output from multi-models are more representative than output from a single model.

This research uses Phase 2 of the North American Land Data Assimilation System (NLDAS-2) forcing file A, which was designed based on NLDAS-1 (Mitchell et al., 2004) forcing, providing gauge-based observed precipitation, bias-correcting shortwave radiation, and surface meteorology reanalyses at hourly temporal resolution, and 1/8th degree special resolution (Table 1.2; http://ldas.gsfc.nasa.gov/nldas/NLDASgoals.php).

Except for precipitation, other meteorological forcing fields of the NLDAS-2 File are mostly derived from NCEP North American Regional Reanalysis (NARR). The spatial resolution of NARR is 32-km and the temporal resolution is 3-hour. Forcing from NARR has been spatially interpolated and temporally disaggregated into NLDAS-2's hourly 1/8th –degree format. During interpolation, the surface downward longwave radiation, surface pressure, air temperature, and specific humidity have been adjusted vertically (Cosgrove et al., 2003).

The downward shortwave radiation (solar radiation) in NLDAS-1 is primarily from satellite-derived Geostationary Operational Environmental Satellite (GOES) imagery (Pinker et al., 2003). In NLDAS-2, a bias-modification was applied to the downward shortwave radiation from NARR with the GOES-based data (http://ldas.gsfc.nasa.gov/nldas/NLDAS2forcing.php). In a previous study about the validation of solar radiation from NARR, a strong agreement (r= 0.98) with the station measurements was observed (Schroeder et al., 2009).

The precipitation field in the NLDAS-2 File A is derived from hourly Doppler Stage II radar precipitation data (1996-present), PRISM topographical adjusted CPC daily CONUS gauge data (Daly et al., 1994; 1979-6 hourly CMORPH data 2002-present), and

3-hourly NARR precipitation data (1979-present) (http://ldas.gsfc.nasa.gov/nldas/NLDAS2forcing.php).

The validation of NLDAS-2 is still underway, and from published validation studies, Noah-based NLDAS-2 generally matched observed soil temperature at different soil layers (Xia et al., 2013). Compared with NARR, NLDAS-2 has higher resolution both spatially and temporally. The downward shortwave radiation had been bias-corrected and the precipitation is observation-based while precipitation in NARR is simulation-based. In the study by Mo et al. (2011), they indicated that NLDAS has a better ability for capturing partitioning between runoff and evapotranspiration. NLDAS-2 has also been applied in estimating evapotranspiration (Peters-Lidard et al., 2011), drought indices estimation (Mo et al., 2011), and climatology of rainfall (Matsui et al., 2010).

In this study, the hourly 1/8th degree-resolution NLDAS-2 forcing was used as the first-step input files. Because this research aims to provide a 4-km-resolution product, the next step goes to the NCAR High-Resolution Land Data Assimilation System, which can increase the resolution of the forcing data from NLDAS, drive the Noah-based land-surface model, and provide high-resolution meteorological and biophysical output.

1.2.2 NCAR High-Resolution Land Data Assimilation System

The High-Resolution Land Data Assimilation System (HRLDAS) was developed by the National Center for Atmospheric Research (NCAR, Chen et al., 2007). The goal of developing HRLDAS is to meet the increasing need of high-resolution meteorological data products (e.g., air temperature, solar radiation) and provide high resolution initial

soil conditions for numerical weather prediction models coupled with a land surface model (e.g., WRF/Noah).

Similar to LDAS, HRLDAS is also based on a land-surface model, namely the Noah Land Surface Model (Noah- LSM), which is driven by meteorological forcing files to simulate soil temperature, soil moisture, surface energy balance, surface water balance, etc.

Noah-LSM was developed on the diurnally dependent Penman potential evaporation approach (Mahrt and Ek, 1984), the multilayer soil model (Mahrt and Pan, 1984) and the primitive canopy model (Pan and Mahrt, 1987). Chen et al. (1996) extended this model by including the canopy resistance approach and Ek et al. (2003) added the formulation of bare soil.

Originally, Noah-LSM was developed to provide the land state for the NOAA/NCEP mesoscale Eta model (Betts et al., 1997; Chen et al., 1997; Ek et al., 2003). It has been included in LDAS, coupled with the Weather Research and Forecasting (WRF) regional atmosphere model, and is also used as the core in HRLDAS.

The running scheme of HRLDAS is presented in Fig.1.2. The input data for running Noah-LSM of HRLDAS includes three parts:

1) Initialized data (e.g., multiple-level soil temperature, canopy water content). Generally, initialized data is only required for the initial time.
2) Land-surface data, including geophysical information (e.g., latitude, longitude, terrain height, land use), soil texture, and vegetation category. In HRLDAS, the

land-surface data is produced by WRF processing. Because the WRF-grid has the same resolution as HRLDAS, data interpolation is not needed for land-surface data. The land-use input is based on 30-s U.S. Geological Survey (USGS) 24 categories (Loveland et al., 1995). Terrain height is based on USGS-derived 30-s topographical height data, soil texture is based on the U.S. General Soil Map, and green vegetation fraction is based on monthly satellite-derived green vegetation fraction (Gutman and Ignatov, 1998).

3) Meteorological forcing data, including near-surface air temperature, downward shortwave radiation, and precipitation. The meteorological forcing data can be prepared from different sources. For example, it can merge the temperature data from NLDAS-2 forcing, precipitation data from NCEP stage-IV, and downward solar radiation derived from GEOS.

Running HRLDAS has five steps (HRLDAS User's Guide, 2012):

1) Raw data preparation.
2) Raw data extraction and organization for forcing data.
3) Model grid configuration.
4) Forcing data interpolation (bilinear).
5) Noah-LSM simulations.

The output data of HRLDAS can be customized, but commonly, the output includes four-layer soil moisture, four-layer soil temperature, evapotranspiration, and meteorological data. The detailed information of input data and output data in this research is included in Chapter 3.

1.3 Studies Using LDAS/HRLDAS in Agricultural Applications

Over the past three decades, remote-sensing data has been integrated with crop models to estimate growth stage and yields. Several studies indicate that remote-sensing data can improve the overall performance of crop models (Maas, 1988a, b; Delecolle et al., 1992; Moulin et al., 1998; Plummer, 2000; Doraiswamy et al., 2004, 2005).

Doraiswamy et al. (2004, 2005) used MODIS-derived LAI to calibrate crop model parameters by adjusting the LAI simulated from the climate-based crop yield model. Using this method, the simulated yield was within 10% of county yields reported by the USDA National Agricultural Statistics Service (NASS). However, in Doraiwamy's studies, the meteorological input data are from 10 weather stations, and only three of them include solar radiation data. The limitation in collecting meteorological data limits the application of remote-sensing based crop simulations at the larger regional scale. Fang et al. (2008) also used MODIS-derived LAI to calibrate crop model parameters, differently than Doraiswamy et al. (2004, 2005). In Fang's study, meteorological data from NLDAS was used in model simulations, and results indicate that NLDAS offers reasonable inputs for simulating crop yield over a regional scale. McNider et al. (2011) developed a real-time gridded crop model for assessing spatial drought stress on crops in the southeastern U.S. using high-resolution radar-derived precipitation, GOES satellite-derived solar radiation, and NOAA Rapid Update Cycle RUC reanalysis temperature. However, in McNider's study, the crop model calibration was only based on three sites in Alabama, therefore the calibration may not be applicable to other areas.

Currently, there is no investigation on applying data from LDAS or HRLDAS as an integrated input for running crop models across the U.S. Corn Belt at a high-resolution regional scale. Based on preliminary studies, it is advantageous to use LDAS/HRLDAS in agricultural applications which also include future yield projections at regional scale. In this research, the Hybrid-Model was selected as the major crop model, which is also a multiple model option rather than only running DSSAT.

1.4 Introduction to El Niño–Southern Oscillation (ENSO)

1.4.1 Definition of El Niño–Southern Oscillation (ENSO)

Climate variability is the variability of climate records where the state of the climate system has no movement (Salinger et al., 2000), and where climate change has shifted the climate system because of internal changes of the system itself or external changes resulting from natural or anthropogenic factors (International Panel on Climate Change, IPCC 1996). Climate variability occurs at long-term and short-term scales and is one of the characteristics of the global climate system (Mavi and Tupper, 2004).

The El Niño-southern oscillation (ENSO) is the phenomenon resulting from the coupled interaction between the tropical oceans and atmosphere through changes in sea surface temperature (SSTs). ENSO is the major seasonal/interannual climate variability which has an influence throughout the world. ENSO includes three phases: El Niño years (Warm Events), La Niña years (Cold Events), and neutral years (Trenberth, 1997). There are different criterion to classify the ENSO years, the details of the ENSO classification in this research is included in Chapter 4.

1.4.2 The Effects of ENSO on Weather Conditions in the U.S. Corn Belt

Cleaveland and Duvick (1992) showed that in Ohio, the El Niño phase correlates with higher probability of wet years while the La Niña phase is associated with drought years. Carlson et al. (1996) indicated that maximum temperatures in August are highly correlated to ENSO events in Iowa. Phillips et al. (1999) reported that in the Corn Belt, compared with neutral years, La Niña years tend to be warmer and drier in summer and El Niño years tend to be cooler and wetter.

1.4.3 The Effects of ENSO on Corn Yield

Many studies indicate that ENSO has a significant impact on crop yield (Garnett and Khandekar, 1992; Hammer et al., 2001; Podestá et al., 2002), including the southeastern U.S. (Garcia y Garcia et al., 2006; Hansen et al., 1998; Mavromatis et al., 2002) and the U.S. Corn Belt (Phillips et al., 1999; Hollinger et al.,). Hansen et al. (1998) indicate that in the southeastern U.S., the mean corn yield in La Niña years was 13.9% higher than the yield in neutral and El Niño years.

In the Midwest, Carlson et al. (1996) claimed that corn yield in the Midwest tended to be higher in El Niño years, and lower in La Niña years. Phillips et al. (1999) shows that ENSO explained 15% of inter-annual corn yield variability in the Corn Belt, positive corn yield anomalies were associated with El Niño years and negative corn yield anomalies associated La Niña years.

ENSO prediction can be used to help producers in making better crop management decisions to reduce farm risk (Cabera et al., 2006. Garcia y Garcia et al., 2010), such as

changes to the planting dates. Solow et al. (1998) estimate that the annual value of perfect ENSO prediction to U.S. agriculture is $323 million.

Based on these past studies, this research investigated the effects of ENSO on corn yields and planting dates in more sites across the U.S. Corn Belt, and also estimated the sensibility of crop models in response to ENSO events.

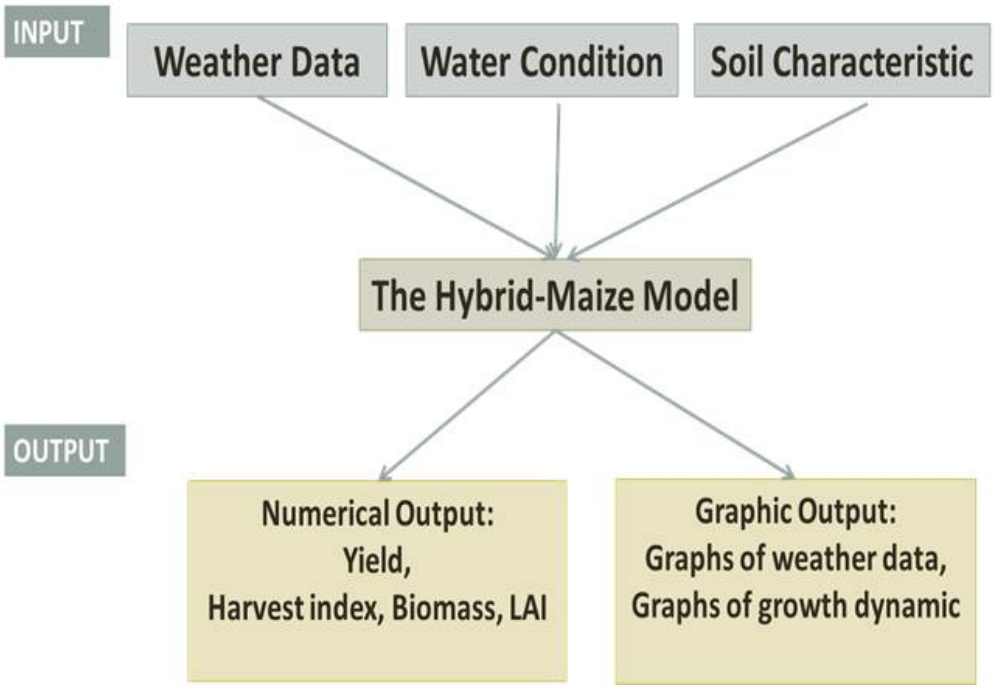

Figure 1.1 Framework of the Hybrid-Maize crop simulation model

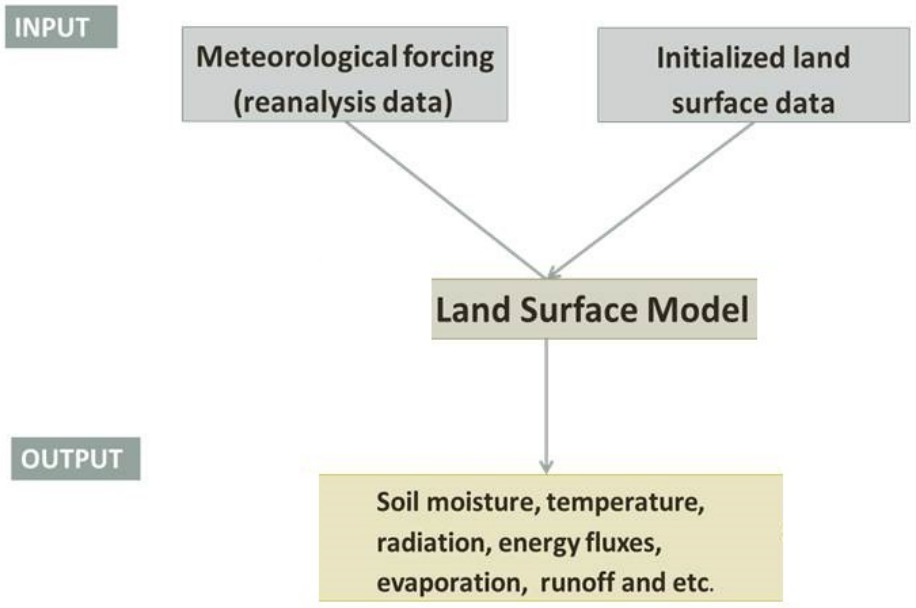

Figure 1.2 Framework of the Land Data Assimilation System (LDAS)

Table 1.1 Fields contained in NLDAS-2 forcing File A

Description	Units
U wind component at 10 m	m/s
V wind component at 10 m	m/s
Air temperature at 2m	K
Specific humidity at 3m	kg/kg
Surface pressure	Pa
Surface downward longwave radiation	W/m^2
Surface downward shortwave radiation	W/m^2
Precipitation hourly total	Kg/m^2
Fraction of total precipitation that is convective	No units
Convective Available Potential Energy(CAPE)	J/kg
Potential evaporation	Kg/m^2

References

Anderson, E. A. (1973). National Weather Service river forecast system--snow accumulation and ablation model. TECHNICAL MEMORANDUM NWS HYDRO-17, NOVEMBER 1973. 217 pp. Anderson, R. M., et al. (2006). Using SSURGO data to improve Sacramento Model a priori parameter estimates. Journal of Hydrology, 320(1–2): 103-116.

Betts, A. K., Chen, F., Mitchell, K. E., & Janjic, Z. I. (1997). Assessment of the land surface and boundary layer models in two operational versions of the NCEP Eta model using FIFE data. Monthly Weather Review, 125(11), 2896-2916.

Brisson, N., Gary, C., Justes, E., Roche, R., Mary, B., Ripoche, D., & Sinoquet, H. (2003). An overview of the crop model STICS. European Journal of Agronomy, 18(3), 309-332.Bristow, K. L., & Campbell, G. S. (1984). On the relationship between incoming solar radiation and daily maximum and minimum temperature. Agricultural and Forest Meteorology, 31(2), 159-166.Brun, F., Wallach, D., Makowski, D., & Jones, J. W. (2006). Working with dynamic crop models: evaluation, analysis, parameterization, and applications. Elsevier, 462 pp.

Burnash, R. J., Ferral, R. L., & McGuire, R. A. (1973). A generalized streamflow simulation system, conceptual modeling for digital computers. Technical Report, Joint Federal and State River Forecast Center, U.S. National Weather Service and California Department of Water Resources, Sacramento, 204 pp.

Brisson, N., Gary, C., Justes, E., Roche, R., Mary, B., Ripoche, D., & Sinoquet, H. (2003). An overview of the crop model STICS. European Journal of agronomy, 18(3), 309-332.

Cabrera, V. E., Fraisse, C., Letson, D., Podestá, G., & Novak, J. (2006). Impact of climate information on reducing farm risk by optimizing crop insurance strategy. Trans. ASABE, 49(4), 1223-1233.

Carlson, R. E., Todey, D. P., & Taylor, S. E. (1996). Midwestern corn yield and weather in relation to extremes of the Southern Oscillation. Journal of Production Agriculture, 9(3), 347-352.

Chen, F., Mitchell, K., Schaake, J., Xue, Y., Pan, H. L., Koren, V., Duan, Q.Y., M. Ek, M., & Betts, A. (1996). Modeling of land surface evaporation by four schemes and comparison with FIFE observations. Journal of Geophysical Research, 101(D3), 7251-7268.

Chen, F., Janjić, Z., & Mitchell, K. (1997). Impact of atmospheric surface-layer parameterizations in the new land-surface scheme of the NCEP mesoscale Eta model. Boundary-Layer Meteorology, 85(3), 391-421.

Chen, F., Manning, K. W., LeMone, M. A., Trier, S. B., Alfieri, J. G., Roberts, R., Tewari, M., Niyogi, D., Horst, T.W., Oncley, S.P., Basara, J.B., & Blanken, P. D. (2007). Description and evaluation of the characteristics of the NCAR high-resolution land data assimilation system. Journal of Applied Meteorology and Climatology, 46(6), 694-713.

Cleaveland, M. K., & Duvick, D. N. (1992). Iowa climate reconstructed from tree rings, 1640–1982. Water Resources Research, 28(10), 2607-2615.

Cosgrove, B. A., Lohmann, D., Mitchell, K. E., Houser, P. R., Wood, E. F., Schaake, J. C., Robock, A., Marshall, C., Sheffield, J., Duan, Q., Luo, L., Higins, R. W., Pinker, R., Tarpley, J. D., & Meng, J. (2003). Real-time and retrospective forcing in the North American Land Data Assimilation System (NLDAS) project. Journal of Geophysical Research, 108(D22), 8842.

Daly, C., Neilson, R. P., & Phillips, D. L. (1994). A statistical-topographic model for mapping climatological precipitation over mountainous terrain. Journal of Applied Meteorology, 33(2), 140-158.

de Wit, C. T. (1966). Photosynthesis of leaf canopies. Centre for Agricultural Publications and Documentation.

de Wit, C. T. (1978). Simulation of assimilation, respiration and transpiration of crops. Wiley, 140 pp.

Delécolle, R., Maas, S. J., Guerif, M., & Baret, F. (1992). Remote sensing and crop production models: present trends. ISPRS Journal of Photogrammetry and Remote Sensing, 47(2), 145-161.

Dirmeyer, P. A., Koster, R. D., & Guo, Z. (2006). Do global models properly represent the feedback between land and atmosphere? Journal of Hydrometeorology, 7(6), 1177-1198.

Doraiswamy, P. C., Hatfield, J. L., Jackson, T. J., Akhmedov, B., Prueger, J., & Stern, A. (2004). Crop condition and yield simulations using Landsat and MODIS. Remote Sensing of Environment, 92(4), 548-559.

Doraiswamy, P. C., Sinclair, T. R., Hollinger, S., Akhmedov, B., Stern, A., & Prueger, J. (2005). Application of MODIS derived parameters for regional crop yield assessment. Remote Sensing of Environment, 97(2), 192-202.

Ek, M. B., Mitchell, K. E., Lin, Y., Rogers, E., Grunmann, P., Koren, V., & Tarpley, J. D. (2003). Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model. Journal of Geophysical Research: Atmospheres, 108(D22), 8851, doi:10.1029/2002JD003296.

Fang, H., Liang, S., Hoogenboom, G., Teasdale, J., & Cavigelli, M. (2008). Corn-yield estimation through assimilation of remotely sensed data into the CSM-CERES-Maize model. International Journal of Remote Sensing, 29(10), 3011-3032.

Garcia y Garcia, A., Hoogenboom, G., Guerra, L. C., Paz, J. O., & Fraisse, C. W. (2006). Analysis of the inter-annual variation of peanut yield in Georgia using a dynamic crop simulation model. Transactions of the ASABE, 49(6), 2005-2015.

Garcia y Garcia, A., Persson, T., Paz, J. O., Fraisse, C., & Hoogenboom, G. (2010). ENSO-based climate variability affects water use efficiency of rainfed cotton grown in the southeastern USA. Agriculture, Ecosystems & Environment, 139(4), 629-635.

Garnett, E., & Khandekar, M. L. (1992). The impact of large-scale atmospheric circulations and anomalies on Indian monsoon droughts and floods and on world grain yields—a statistical analysis. Agricultural and Forest Meteorology, 61(1), 113-128.

Geng, S. (1988). A Program to Simulate Meteorological Variables: Documentation for SIMMETEO, University of California, Agricultural Experiment Station.

Goodin, D. G., Hutchinson, J. M. S., Vanderlip, R. L., & Knapp, M. C. (1999). Estimating solar irradiance for crop modeling using daily air temperature data. Agronomy Journal, 91(5), 845-851.

Goudriaan, J., & Van Laar, H. H. (1994). Modelling potential crop growth processes: textbook with exercises (Vol. 2). Springer Netherlands, 239 pp.

Grant, R. H., Hollinger, S. E., Hubbard, K. G., Hoogenboom, G., & Vanderlip, R. L. (2004). Ability to predict daily solar radiation values from interpolated climate records for use in crop simulation models. Agricultural and Forest Meteorology, 127(1), 65-75.

Grassini, P, Yang, H. S., and Cassman, K. G. (2009). Limits to maize productivity in Western Corn-Belt: A simulation analysis for fully irrigated and rainfed conditions. Agricultural and Forest Meteorology, 149, 1254-1265.

Gutman, G., & Ignatov, A. (1998). The derivation of the green vegetation fraction from NOAA/AVHRR data for use in numerical weather prediction models. International Journal of Remote Sensing, 19(8), 1533-1543.

Haefner, J. W. (2005). Modeling biological systems: principles and applications. Springer, 476 pp.

Hammer, G. L., Hansen, J. W., Phillips, J. G., Mjelde, J. W., Hill, H., Love, A., & Potgieter, A. (2001). Advances in application of climate prediction in agriculture. Agricultural Systems, 70(2), 515-553.

Hammer, G. L., Kropff, M. J., Sinclair, T. R., & Porter, J. R. (2002). Future contributions of crop modelling—from heuristics and supporting decision making to understanding genetic regulation and aiding crop improvement. European Journal of Agronomy, 18(1), 15-31.

Hansen, J. W., Hodges, A. W., & Jones, J. W. (1998). ENSO influences on agriculture in the southeastern United States. Journal of Climate, 11(3), 404-411.

HRLDAS User's Guide, 2012: Version 0.5. Available online at: http://www.rap.ucar.edu/research/land/technology/lsm/HRLDAS_USERS_GUIDE_34.pdf

Hodges, T., French, V., & LeDuc, S. (1985). Estimating solar radiation for plant simulation models (No. N-85-21751; E-85-10089; NASA-CR-175524). National Oceanic and Atmospheric Administration, Columbia, MO (USA).

Hollinger, S. E., Ehler, E. J., & Carlson, R. E. (2001). Midwestern United States corn and soybean yield response to changing El Niño-Southern Oscillation conditions during the growing season. Chapter in Impacts of El Niño and Climate Variability on Agriculture, Cynthia Rosenzweig (ed.), 31-54.

IPCC. 1996. Climate change 1995: the science of climate change. Cambridge University Press, Cambridge, UK.

Jones, C. A., Dyke, P. T., Williams, J. R., Kiniry, J. R., Benson, V. W., & Griggs, R. H. (1991). EPIC: an operational model for evaluation of agricultural sustainability. Agricultural Systems, 37(4), 341-350.

Jones, C. A., Kiniry, J. R., & Dyke, P. T. (1986). CERES-Maize: A simulation model of maize growth and development. Texas A&M University Press, 194 pp.

Jones, J.W., G. Hoogenboom, C.H. Porter, K.J. Boote, W.D. Batchelor, L.A. Hunt, P.W. Wilkens, U. Singh, A.J. Gijsman, and J.T. Ritchie (2003). The DSSAT cropping system model. European Journal of Agronomy, 18, 235–265

Koren, V., Schaake, J., Mitchell, K., Duan, Q. Y., Chen, F., & Baker, J. M. (1999). A parameterization of snowpack and frozen ground intended for NCEP weather and climate models. Journal of Geophysical Research: Atmospheres, 104(D16), 19569-19585.

Koster, R. D., & Suarez, M. J. (1992). Modeling the land surface boundary in climate models as a composite of independent vegetation stands. Journal of Geophysical Research: Atmospheres, 97(D3), 2697-2715.

Koster, R. D., & Suarez, M. J. (1996). Energy and water balance calculations in the Mosaic LSM. NASA Tech. Memo, 104606(9), 59.

Kropff, M.J., and. van Laar, H.H. (1993). Modelling crop-weed interactions. CABI International, in association with the International Rice Research Institute, 274 pp., Wallingford, Oxon, UK.

Liang, X., Lettenmaier, D. P., Wood, E. F., & Burges, S. J. (1994). A simple hydrologically based model of land surface water and energy fluxes for general circulation models. Journal of Geophysical Research: Atmospheres, 99(D7), 14415-14428.

Liang, X., Wood, E. F., & Lettenmaier, D. P. (1996). Surface soil moisture parameterization of the VIC-2L model: Evaluation and modification. Global and Planetary Change, 13(1), 195-206.

Loveland, T. R., Merchant, J. W., Brown, J. F., Ohlen, D. O., Reed, B. C., Olson, P., & Hutchinson, J. (1995). Seasonal land-cover regions of the United States. Annals of the Association of American Geographers, 85(2), 339-355.

Maas, S. J. (1988). Use of remotely-sensed information in agricultural crop growth models. Ecological Modelling, 41(3), 247-268.

Maas, S. J. (1988). Using satellite data to improve model estimates of crop yield. Agronomy Journal, 80(4), 655-662.

Mahmood, R., & Hubbard, K. G. (2002). Effect of time of temperature observation and estimation of daily solar radiation for the Northern Great Plains, USA. Agronomy Journal, 94(4), 723-733.

Mahrt, L., & Ek, M. (1984). The influence of atmospheric stability on potential evaporation. Journal of Applied Meteorology, 23:2, 222-234.

Mahrt, L., & Pan, H. (1984). A two-layer model of soil hydrology. Boundary-Layer Meteorology, 29(1), 1-20.

Matsui, T., Mocko, D., Lee, M. I., Tao, W. K., Suarez, M. J., & Pielke, R. A. (2010). Ten□year climatology of summertime diurnal rainfall rate over the conterminous US. Geophysical Research Letters, 37(13).

Mavi, H. H. S., & Tupper, G. J. (2004). Agrometeorology: principles and applications of climate studies in agriculture. The Haworth Press, Inc., New York, 364 pp.

Mavromatis, T., Jagtap, S. S., & Jones, J. W. (2002). El Nino-Southern Oscillation effects on peanut yield and nitrogen leaching. Climate Research, 22(2), 129-140.

McNider, R. T., Christy, J. R., Moss, D., Doty, K., Handyside, C., Limaye, A., Garcia Y Garcia, A., & Hoogenboom, G. (2011). A real-time gridded crop model for assessing spatial drought stress on crops in the Southeastern United States. Journal of Applied Meteorology and Climatology, 50(7), 1459-1475.

Meinke, H., Carberry, P. S., McCaskill, M. R., Hills, M. A., & McLeod, I. (1995). Evaluation of radiation and temperature data generators in the Australian tropics and sub-tropics using crop simulation models. Agricultural and Forest Meteorology, 72(3), 295-316.

Mitchell, K. E., Lohmann, D., Houser, P. R., Wood, E. F., Schaake, J. C., Robock, A., Cosgrove, B.A., Sheffield, J., Duan, Q., Luo, L., Higgins, R.W., Pinker, R.T., Tarpley, J.D., Lettenmaier, D.P., Marshall, C.H., Entin, J.K., Pan, M., Shi, W., Koren, V., Meng, J., Ramsay, & Bailey, A. A. (2004). The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system. Journal of Geophysical Research, 109(D7), D07S90.

Mo, K. C., Long, L. N., Xia, Y., Yang, S. K., Schemm, J. E., & Ek, M. (2011). Drought indices based on the Climate Forecast System Reanalysis and ensemble NLDAS. Journal of Hydrometeorology, 12(2), 181-205.

Moulin, S., Bondeau, A., & Delecolle, R. (1998). Combining agricultural crop models and satellite observations: from field to regional scales. International Journal of Remote Sensing, 19(6), 1021-1036.

Niu, G. Y., Yang, Z. L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., Kumar, A., Manning, K., Niyogi, D., Rosero, E., Tewari, M., & Xia, Y. (2011). The community Noah land surface model with multiparameterization options (Noah MP): 1. Model description and evaluation with local scale measurements. Journal of Geophysical Research: Atmospheres, 116(D12), DOI:10.1029/2010JD015139.

Pan, H. L., & Mahrt, L. (1987). Interaction between soil hydrology and boundary-layer development. Boundary-Layer Meteorology, 38(1-2), 185-202.

Peters Lidard, C. D., Kumar, S. V., Mocko, D. M., & Tian, Y. (2011). Estimating evapotranspiration with land data assimilation systems. Hydrological Processes, 25(26), 3979-3992.

Phillips, J., Rajagopalan, B., Cane, M., & Rosenzweig, C. (1999). The role of ENSO in determining climate and maize yield variability in the US cornbelt. International Journal of Climatology, 19(8), 877-888.

Pinker, R. T., Tarpley, J. D., Laszlo, I., Mitchell, K. E., Houser, P. R., Wood, E. F., ... & Higgins, R. W. (2003). Surface radiation budgets in support of the GEWEX Continental Scale International Project (GCIP) and the GEWEX Americas Prediction Project (GAPP), including the North American Land Data Assimilation System (NLDAS) project. Journal of Geophysical Research: Atmospheres (1984–2012), 108(D22).

Plummer, S. E. (2000). Perspectives on combining ecological process models and remotely sensed data. Ecological Modelling, 129(2), 169-186.

Podesta, G., Letson, D., Messina, C., Royce, F., Ferreyra, R. A., Jones, J., Hansen' J., Llovet, I., Grondona, M., & O'Brien, J. J. (2002). Use of ENSO-related climate information in agricultural decision making in Argentina: a pilot experience. Agricultural Systems, 74(3), 371-392.

Rabbinge, R. (1993). The ecological background of food production. In Ciba Foundation Symposium 177-Crop Protection and Sustainable Agriculture. John Wiley & Sons, Ltd.

Rabbinge, R. & Kropff, M.J. (2008). Introduction to 40 Years Theory and Model in Wageningen. In: 40 Years Theory and Model, Wageningen UR, 1-4, http://www.icasa.net/news/40Years.pdf

Richardson, C. W. (1981). Stochastic simulation of daily precipitation, temperature, and solar radiation. Water Resources Research, 17(1), 182-190.

Richardson, C. W., & Wright, D. A. (1984). WGEN: A model for generating daily weather variables. ARS.

Salinger, M. J., Stigter, C. J., & Das, H. P. (2000). Agrometeorological adaptation strategies to increasing climate variability and climate change. Agricultural and Forest Meteorology, 103(1), 167-184.

Schroeder, T. A., Hember, R., Coops, N. C., & Liang, S. (2009). Validation of solar radiation surfaces from MODIS and reanalysis data over topographically complex terrain. Journal of Applied Meteorology and Climatology, 48(12), 2441-2458.

Solow, A. R., Adams, R. F., Bryant, K. J., Legler, D. M., O'Brien, J. J., McCarl, B. A., ... & Weiher, R. (1998). The value of improved ENSO prediction to US agriculture. Climatic change, 39(1), 47-60.

Soltani, A., and Sinclair, T. R. (2012). Modeling physiology of crop development, growth and yield. CABI, 322 pp.

Trenberth, K. E. (1997). The definition of el nino. Bulletin of the American Meteorological Society, 78(12), 2771-2777.

Van Ittersum, M. K., Leffelaar, P. A., Van Keulen, H., Kropff, M. J., Bastiaans, L., & Goudriaan, J. (2003). On approaches and applications of the Wageningen crop models. European Journal of Agronomy, 18(3), 201-234.

Wood, E. F., Lettenmaier, D., Liang, X., Nijssen, B., & Wetzel, S. W. (1997). Hydrological modeling of continental-scale basins. Annual Review of Earth and Planetary Sciences, 25(1), 279-300.

Xia, Y., Mitchell, K., Ek, M., Sheffield, J., Cosgrove, B., Wood, E., Luo, ' L., Alonge, C., Wei, ' H., Meng, J., Livneh, B., Lettenmaier, D., Koren, V., Duan,, Q., Mo, K., Fan' Y., & Mocko, D. (2012). Continental□scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS□2): 1. Intercomparison and application of model products. Journal of Geophysical Research: Atmospheres, 117(D3).

Xia, Y & Coauthors (2013). Validation of Noah-simulated soil temperature in the North American Land Data Assimilation System Phase 2. Journal of Applied Meteorology and Climatology, 52, 455–471.

Yang, H.S., Dobermann, A., Lindquist, J.L., Walters, D.T., Arkebauer, T.J., and Cassman, K.G. (2004). Hybrid-Maize: A maize simulation model that combines two crop modeling approaches. Field Crops Research 87:131–154.

Yang, H.S., Dobermann, A, Cassman, K.G., and Walters, D.T. (2006). Features, applications, and limitations of the Hybrid-Maize simulation model. Agronomy Journal, 98, 737-748.

CHAPTER 2. SENSITIVITY ANALYSIS AND VALIDATION OF THE HYBRID-MAIZE SIMULATION MODEL OVER THE U.S CORN BELT

The Hybrid Maize is a crop simulation model that estimates corn yields using agronomic and climatic information. This model has been used in prior studies but a long-term, regional analysis over the U.S. Corn Belt was lacking. In this chapter, such an assessment has been undertaken, including sensitive analysis and model validation. The study was conducted at two scales: county scale and field scale. The county-scale study is based on 30-year daily weather data and the National Agricultural Statistics Service (NASS) survey corn yield data for 18 sites across the Midwest. The field-scale study is based on 3-year daily weather data and measured corn yield data from two Ameriflux sites at Bondville, IL and Mead, NE. The overall scheme flowchart is provided in Fig. 2.1. The hypothesis in this chapter is: the Hybrid-Maize model can provide reliable yield estimations at both the regional scale and field scale.

2.1 Materials and methods

2.1.1 Data resources and locations

In this research, validations of the Hybrid Maize model were applied at two scales – the county scale and field scale. The county-scale study included 18 counties across the Corn Belt (Fig. 1.1). These counties display representative values for corn yield and

climatic conditions, data availability, and accessibility and this plays an important role in selecting these counties.

Thirty years of (1981-2010) daily weather data (minimum temperature, maximum temperature, and rainfall) were collected from the NOAA Summary of the Day Data Set for a representative station site within the county of interest. Due to the non-availability of downward shortwave radiation data in those selected weather stations, in this research, solar radiation was generated with the WeatherMan utility from the DSSAT crop simulation model package (Pickering et al. 1994). County corn yield data were collected from the National Agricultural Statistics Service (NASS) (http://www.nass.usda.gov/) annual survey.

The field-scale study included two AmeriFlux sites: Bondville, IL (40,00°N, 88.29°W) and Mead, NE (41.18°N, 96.44°W)(Fig. 1.1). Hourly weather data (2001 ~2006), and yield data were collected for both sites from the AmeriFlux site and data exploration system (http://ameriflux.ornl.gov/). The data were analyzed, paired, and checked for consistency. They were also analyzed for outliers and for any missing periods.

2.1.2 Crop model configuration

The Hybrid-Maize model requires three groups of input data: crop and management, weather, and soil. Crop and management data include corn maturity (in total growing degree day, or GDD), plant date, and plant population. For simulations under optimal water management (i.e., non-water limiting) of yield potential, required weather data

include daily minimum and maximum air temperature (˚C), daily sum of global radiation (MJ/m^2). No soil data is required.

In this research, the model was run under optimal water conditions, which means no water stress was present. For the county-scale study, the planting date was set as May 1 and the plant population was set to 78*100ha (31, 600/acre), the corn maturity condition is GDD 2500 (50F based) and the genetic parameters were set as model default. For the field-scale studies at Bondville, IL and Mead, NE, three years of corn planting data and corn yield are presented in Table 2.1. The breeding brand is Pioneer and the potential number of kernels per ear was set as 550. The soil nitrogen condition was set as optimal at both sites

2.1.3 Sensitivity analysis scheme

The initial sensitivity analysis was conducted based on 30-year weather data for 18 county-scale sites across the Midwest and use a one-at-a-time (OAT) approach in sensitivity analysis. Based on the model settings, there are three groups of 29 parameters tested, with parameter changes set at ±10%, ±20%, and ±30% of the default values. For the upper temperature cutoff for GDD accumulation, the changes in daily maximum temperature were ±3˚, ±7, ˚and ±10˚ (Table 2.2). Every change in the parameters resulted in changes in simulated yields. There are a total of 94,500 simulations for the 30 years (1981-2010) of 18 county-scale sites. Besides using relative percentage change of simulated yield to indicate model sensitivity, Sensitivity Index (SI) was also used to assess model sensitivity, and was derived as:

$$SI=\left|\frac{(O-O_{BC})}{(I-I_{BC})} * \frac{I_{BC}}{O_{BC}}\right| \quad (1)$$

Where *O* is the output value, O_{BC} is the output value for the baseline scenario which uses the default parameter values, *I* is the input value, and I_{BC} is the original input value of the baseline scenario. The larger the SI parameter, the more sensible the yield output is for a parameter.

Due to the limitation of the OAT method in reflecting the interaction between parameters, in the second-step sensitive analysis, a global sensitivity analysis (Niyogi et al. 1997) was conducted based on 30-year weather data in Johnson County, IA. Since the focus was on parameters that can possibly be calibrated from remote sensing data and other methods at the regional scale, five parameters were selected based on the results of initial sensitivity analysis: K (light extinction coefficient), UT (upper temperature cutoff for growing degree days accumulation), TL (Threshold LAI above which leaf senescence due to light competition occurs), LUE (initial light use efficiency), and GRG (GDD10C requirement for germination). The 10 interaction groups are K+UT, K+TL, K+LUE, K+GRG, UT+TL, UT+LUE, UT+GRG, TL+LUE, TL+GRG, and LUE+GRG. For every interaction running, two parameters were changed each time. There were a total of $2^5 * 30 = 960$ factorial design simulations conducted for the five parameters. Sensitivity indices were calculated as:

$$Y_{i+j} = Y_d + \alpha_i + \alpha_j + \alpha_{ij} \quad (2)$$

Y_d is the result using default parameter values, α_i and α_j are the main effect of each parameter, and α_{ij} is the interaction effect between two parameters. Take *K* and *LUE* for

example; Y_{K+LUE} is the simulated result when both parameters *K* and *LUE* were changed, $Y_{K+LUE} = f(K, LUE)$, Y_K is the simulated result when only parameter *K* was changed, $Y_K = f(K)$. Y_{LUE} is the simulated result when only *LUE* was changed, $Y_{LUE} = f(LUE)$. $\alpha_K = Y_K - Y_d$ is the main effect from parameter K. $\alpha_{LUE} = Y_{LUE} - Y_d$ is the main effect from parameter *LUE*. $\alpha_{K+LUE} = Y_{K+LUE} - Y_d - \alpha_K - \alpha_{LUE}$ is the interaction effect between *K* and *LUE*.

$$V_T = \sum_i V_i + \sum_{i<j} V_{ij} \quad (3)$$

V_T is the total variability of the 960 simulations, V_i is the sum of squares on the main effect of parameter i, V_{ij} is the sum of squares on the interaction effect between parameters.

Main effect sensitivity index $S_i = \frac{V_i}{V_T}$ (4)

Interaction effect sensitivity index $S_{ij} = \frac{V_{ij}}{V_T}$ (5)

Total effect sensitivity index $S_{iT} = \frac{V_{i+V_{ij}}}{V_T}$ (6)

For parameter *LUE*, $S_{LUE} = \frac{V_{LUE}}{V_T}$, the interaction effect sensitivity index between *LUE* and *K* is $S_{K+LUE} = \frac{V_{K+LUE}}{V_T}$, the total effect sensitivity index of *LUE* is $S_{LUE\,T} = \frac{V_{LUE} + V_{LUE+K} + V_{LUE+UT} + \cdots}{V_T}$.

The calculations were conducted in the Excel spreadsheets.

2.1.4 Model validation and regression analysis scheme

This research also validated the simulated yield data against actual yield data. The validations were conducted at two scales. For the county-scale study, we validated the 30-year simulated yield output with NASS survey data. For the field-scale study, the 6-year simulated yields were validated against field observations from Ameriflux at two field sites at Bondville, IL and Mead, NE. The difference between simulated yields and observed data were quantified using the mean absolute error (*MAE*):

$$D_i = Y_s - Y_a \tag{7}$$

Where Y_s is simulated yield data and Y_a is the actual data. *MAE* was calculated as

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|D_i| \tag{8}$$

The advantage of using *MAE* is that it is convenient and has the same units as the yield (Wallach et al. 2006).

Since the Hybrid-Maize model was developed to simulate the potential yield without yield losses from water stress, nutrient deficiencies, diseases, pests and insects, multiple regression analysis was used to quantify the gap between the modeled potential yield data and the actual yield data. In order to obtain an averaged multiple coefficient, the constant in regression analysis was set to zero. This procedure allowed calibration of the model results to account for other environmental and agronomic as well as management decisions that were not available as input to the model.

2.2 Results

2.2.1 Sensitivity analysis results

The sensitivity index (Fig. 2.2) indicates that the five most sensitive parameters are G2 (potential number of kernels per ear), G5 (potential kernel filling rate), LUE (initial light use efficiency), UT (Upper temperature cutoff for growing degree days) accumulation, and RG (growth respiration coefficient of grain). According to the relative change in yield simulation (Fig. 2.3), changes in G2 and G5 have the largest impact on yield simulation, and they have equal influence on the model. For the general parameters, the model is most sensitive to UT and it is noted that the model is much more sensitive to decreases in the UT value than increases in UT. Among the respiration and photosynthesis parameters, LUE is the dominant one that most influences the model results. The sensitivity index of the yield simulation was significantly stable across the 30 years of weather data with relatively small variations. Therefore the variation of climate in different years would have a moderate impact on the sensitivity analysis results for the optimum parameter conditions set in the model.

In this paper, there were 29 parameters tested, however, when running the model under optimal water conditions, the model was not sensitive to nine of the parameters, meaning the relative change in model prediction of yield is non-significant when changing the parameters. The nine parameters include: FT (fraction of leaf biomass that can be translocated as carbohydrate to grain each day), MF (maximum fraction of leaf biomass at silking that can be translocated as carbohydrate to grain), EF (efficiency of

carbohydrate translocation from stem of leaf to grain), LF (senescent leaf area at maturity as a fraction of maximum LAI achieved at silking), EP (empirical parameters that determine the relative contribution of a soil layer to water uptake), LWS (leaf water suction at a permanent wilting point in cm), RTT (resistance of plants to transpiration in cm), MDE (maximum days allowed form planting to emergence), and MRG (maintenance respiration coefficient for grain).

The results of the OAT sensitivity analysis indicate it is important to validate and calibrate the G2, G5, LUE, UT, and RG parameters. However, since the model will be applied across the Corn Belt at the regional scale and aim for future climate scenarios, it is difficult to collect genetic parameters for the whole domain. Therefore, based on the OAT sensitivity analysis results, an additional global sensitivity analysis based on factorial design was conducted. Five parameters: K, UT, TL, LUE, and GRG, whose information could be potentially obtained through remote sensing data and other methods at the regional scale, were selected. In Fig. 2.4, sensitivities smaller than 1% were ignored, and LUE had the largest sensitivity index. In Fig. 2.5, LUE contributes the most to the total sensitivity index. Therefore, calibrating LUE will be helpful for future regional applications.

2.2.2 Model validation at county-scale and field-scale

In order to apply the Hybrid-Maize model in the Midwest, the model was validated at 18 county-scale sites across 30 years. The results (Fig. 2.6) show that there is bias but generally similar trends between the model-simulated yield and the NASS survey yield.

The MAE for the 18 sites is 5.4 Mg/ha (86 bu/acre). There are two limitations which can explain the bias between the model-simulated yield and the NASS survey yield: (1) the Hybrid-Maize model was developed to simulate the potential yield under optimal conditions; (2) the NASS survey data is the average yield data which includes different varieties of corn and different agronomic management. However, the overall similar trends between simulated yield and survey yield indicate that application of a regression analysis can help to narrow the gap between simulations and observations. At field level, Table 2.3 shows that the 3-year average simulated yield in Mead, NE is 8.54 Mg/ha (136 bu/acre) while the 3-year average measured yield is 8.67 Mg/ha (138 bu/acre). The 3-year average simulated yield in Bondville, IL is 10.30 Mg/ha (164 bu/acre), which is slightly lower than the 3-year average measured yield data of 10.99 Mg/ha (175 bu/acre). MAE of these two field sites is 0.63 Mg/ha (10 bu/acre).

The bias between simulated and measured yield at field scale is narrower than the bias at the county scale. This could be because the two field sites are under better agronomic management than average producers, which helps the actual yield to approach the potential yield.

2.2.3 Regression analysis

After conducting the yield estimation through regression analysis, the bias between simulated and census yield was reduced (Fig. 2.7). The MAE of the yield data after regression analysis is 1.32 Mg/ha (21 bu/acre), which is much lower than the MAE before regression analysis. In order to obtain a multiple coefficient which can be applied regionally, during the regression analysis, the constant was set as 0. The averaged

multiple regression coefficient of the 18 site county-scale study is 0.6 with a variance of 0.007. Therefore, if the Hybrid-Maize model is applied in predicting county average corn yield, it is possible that the model-simulated yield can be used by multiplying 0.6 to decrease the bias between the simulated potential yield and actual survey yield. Since the agronomic management of the two field sites is appropriate to help the yield approach the potential yield, there is no need to conduct a regression analysis at the field scale.

2.3 Summary

According to the results of two different sensitivity analyses, it was shown that yield simulations are sensitive to the genetic parameters: for instance, G2 (potential number of kernels per ear) and G5 (potential kernel filling rate). Also, the model is highly sensitive to LUE (initial light use efficiency) and is useful in calibrating those parameters. However, since the objective is to widely apply the Hybrid-Maize model across the Corn Belt, it is difficult to collect genetic parameters for the whole area. Calibrating the LUE is a possible way to improve model performance in future studies. The validation results indicate the Hybrid-Maize model performs well in simulating yield at field scale where there is appropriate agronomic management. Although when validating the model at the county scale, there is a gap between the simulated and actual survey yield, and after regression analysis, the gap can be narrowed down by a multiple of 0.6 with the original simulated results. The study has several key limitations: 1) the model was running under optimal water conditions during the study period, and 2) the lack of soil characteristics that can influence the model simulation.

Table 2.1 The planting date and plant density for Bondville, IL and Mead, NE

Sites	Year	Planting date	Plant density (per ha)
Bondville, IL	2001	April 19	78,000
	2003	April 16	78,000
	2005	April 22	78,000
Mead, NE	2001	May 14	62,236
	2003	May 13	66,108
	2005	April 27	60,358

Table 2.2 Parameter variations for the one-at-a-time approach

Parameter	Description	default value	-30%	-20%	-10%	+10%	+20%	+30%
genetic parameters								
G2	Potential number of kernels per ear	675	472.5	540	607.5	742.5	810	877.5
G5	potential kernel filling rate	8.7	6.09	6.96	7.83	9.57	10.44	11.31
k	Light extinction coefficient	0.55	0.385	0.44	0.495	0.605	0.66	0.715
general growth parameters								
FT	Fraction of leaf biomass that can be translocated as carbohydrate to grain each day	0.0050	0.0035	0.0040	0.0045	0.0055	0.0060	0.0065
MF	Maximum fraction of leaf biomass at silking that can be translocated as carbohydrate to grain	0.15	0.105	0.12	0.135	0.165	0.18	0.195
EF	Efficiency of carbonhydrate translocation from stem of leaf to grain	0.26	0.182	0.208	0.234	0.286	0.312	0.338
RD	Daily root death(turnover) rate in fraction of total root biomass	0.0050	0.0035	0.0040	0.0045	0.0055	0.0060	0.0065
SDC	Stay-green coefficient for controlling leaf senescence after silking	4	2.8	3.2	3.6	4.4	4.8	5.2
LF	Senescent leaf area at maturity as a fraction of maximum LAI achieved at silking	0.7	0.49	0.56	0.63	0.77	0.84	0.91
UT	Upper temperature cutoff(degree C) for GDD accumulation	34	24	27	31	37	41	44
TL	Threshold LAI above which leaf senescence due to light competition occurs	4	2.8	3.2	3.6	4.4	4.8	5.2
BAC	Biomass alloction coefficient for root at emergence	0.35	0.245	0.28	0.315	0.385	0.42	0.455
DS	Development stage(0 to 2 scale with silking at 1) at which root system stops growing	1.15	0.805	0.92	1.035	1.265	1.38	1.495
EP	Empirical parameters that determines the relative contribution of a soil layer to water uptake	4	2.8	3.2	3.6	4.4	4.8	5.2
LWS	Leaf water suction at permanent wilting point(cm)	17000	11900	13600	15300	18700	20400	22100
RTT	Resistance of plant to transpiration(cm)	9690	6783	7752	8721	10659	11628	12597
GRG	GDD10C requirement for GERMINATION	15	10.5	12	13.5	16.5	18	19.5
GRE	GDD10C requirement for emergence per cm depth	6	4.2	4.8	5.4	6.6	7.2	7.8
MDE (*)	Maximum days allowed form planting to emergence	25	18	20	23	28	30	33
resp & photosyn parameters								
RL	Growth respiration coefficient of leaf	0.470	0.329	0.376	0.423	0.517	0.564	0.611
RS	Growth respiration coefficient of stem	0.520	0.364	0.416	0.468	0.572	0.624	0.676
RR	Growth respiration coefficient of root	0.450	0.315	0.360	0.405	0.495	0.540	0.585
RG	Growth respiration coefficient of grain	0.490	0.343	0.392	0.441	0.539	0.588	0.637
MRL	Maintenance respiration coefficient for leaf	0.0100	0.0070	0.0080	0.0090	0.0110	0.0120	0.0130
MRS	Maintenance respiration coefficient for stem	0.0060	0.0042	0.0048	0.0054	0.0066	0.0072	0.0078
MRR	Maintenance respiration coefficient for root	0.0050	0.0035	0.0040	0.0045	0.0055	0.0060	0.0065
MRG	Maintenance respiration coefficient for grain	0.0050	0.0035	0.0040	0.0045	0.0055	0.0060	0.0065
MSR	Maximum (photosynthetic) assimilation rate	7.0	4.9	5.6	6.3	7.7	8.4	9.1
LUE	Initial light use efficiency	12.5	8.8	10.0	11.3	13.8	15.0	16.3

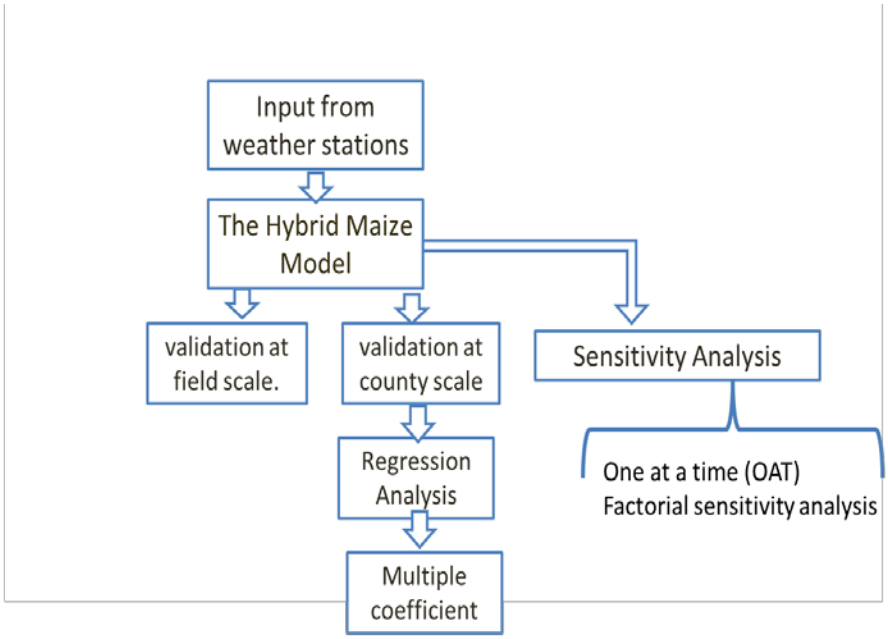

Figure 2.1 Methodology flowchart

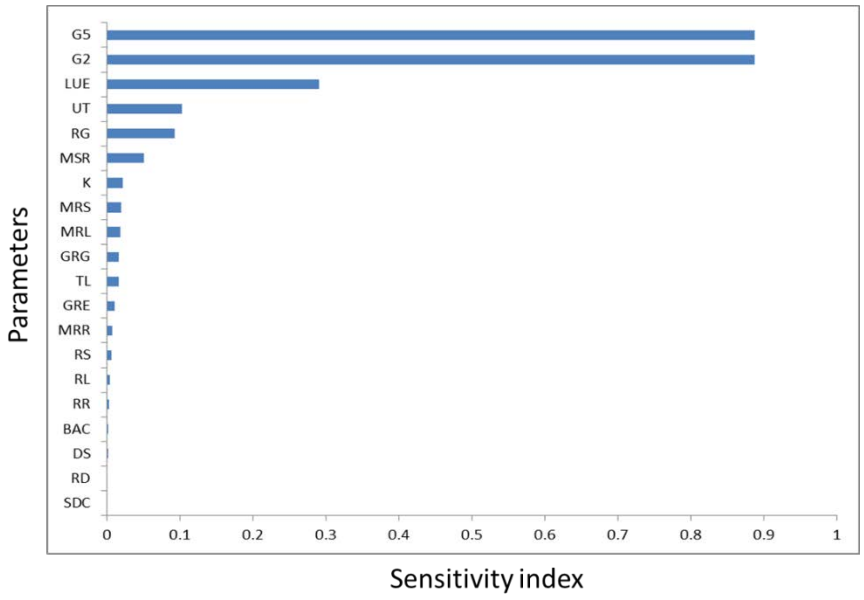

Figure 2.2 Grain yield sensitivity index of parameters in the Hybrid-Maize model based on the one-at-a-time (OAT) approach

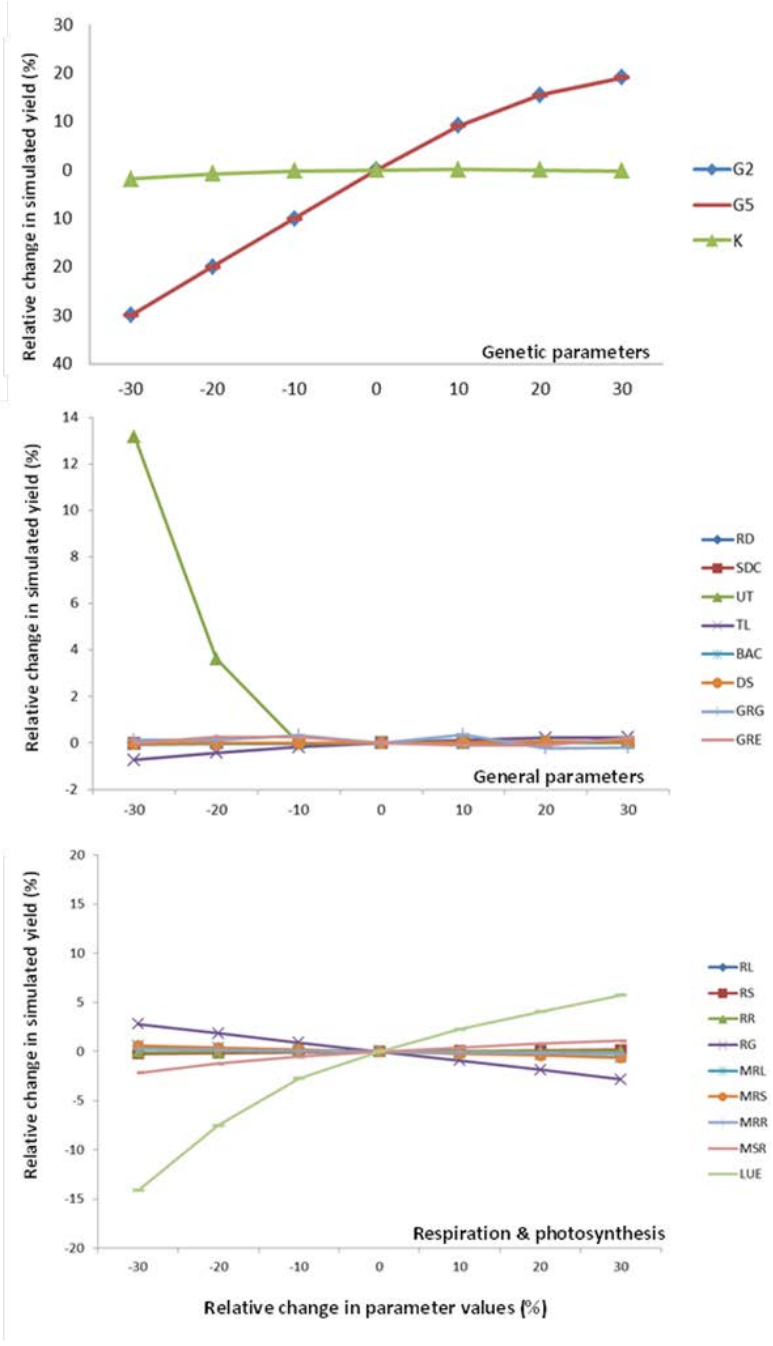

Figure 2.3 The average relative change in model prediction reflects the relative change in parameter values of the Hybrid-Maize model across 18 counties in the Corn Belt through 30 years (1981-2010).

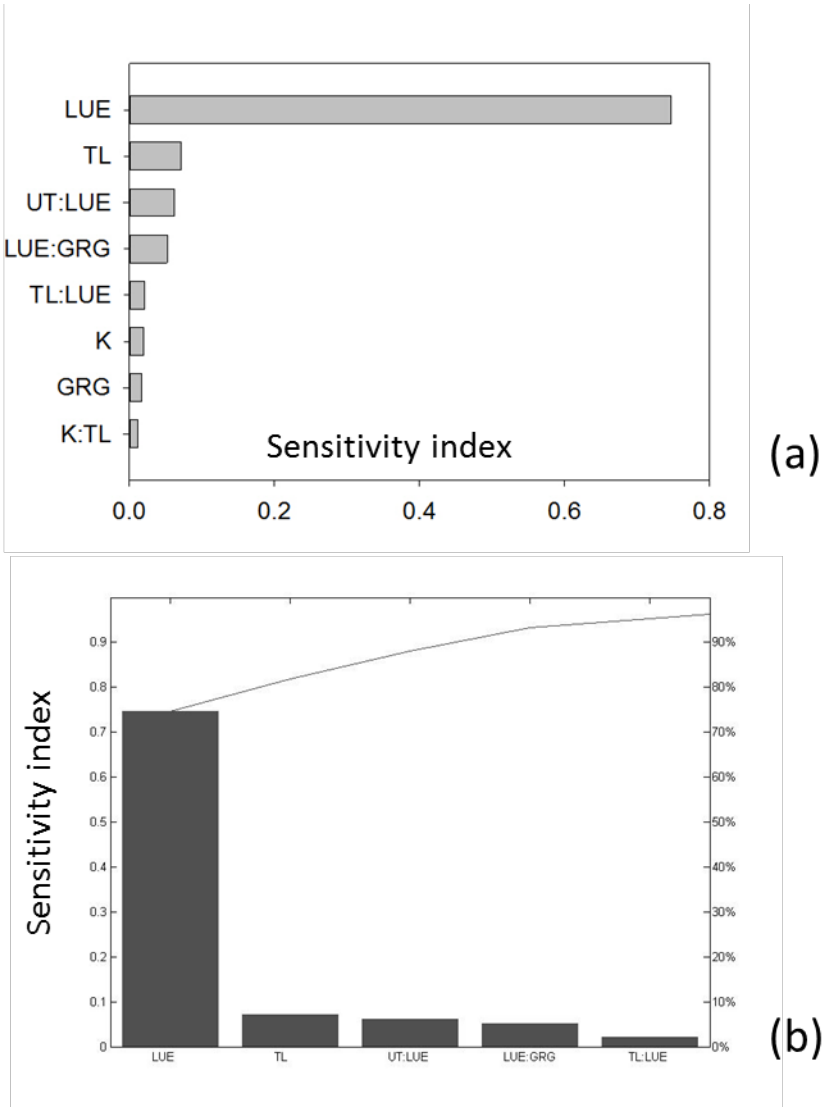

Figure 2.4 The eight largest factorial sensitivity indices based on (a) the factorial design and (b) the Pareto plot for the five largest factorial sensitivity indices

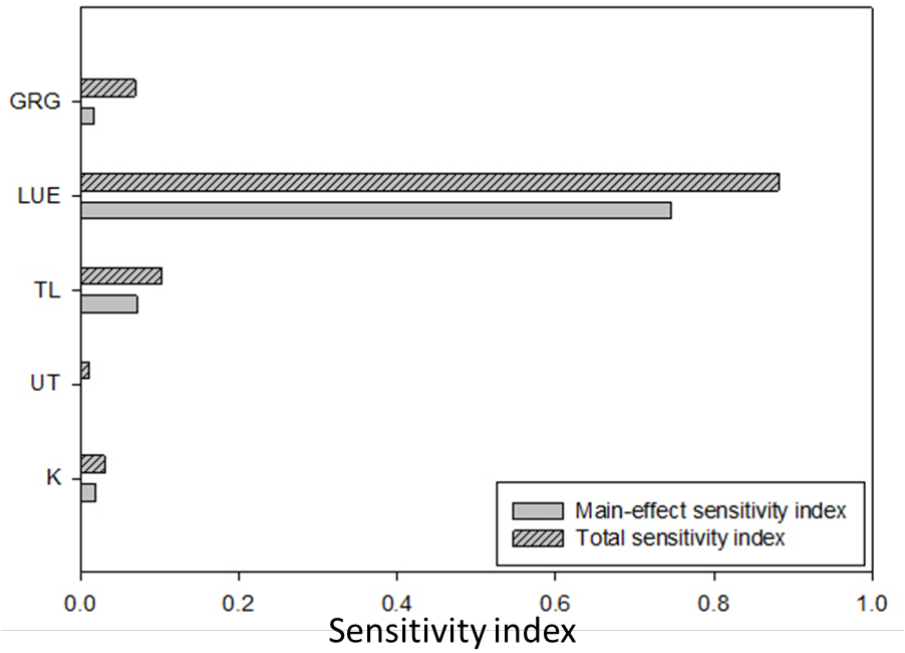

Figure 2.5 The main-effect and total sensitivity indices based on the factorial design

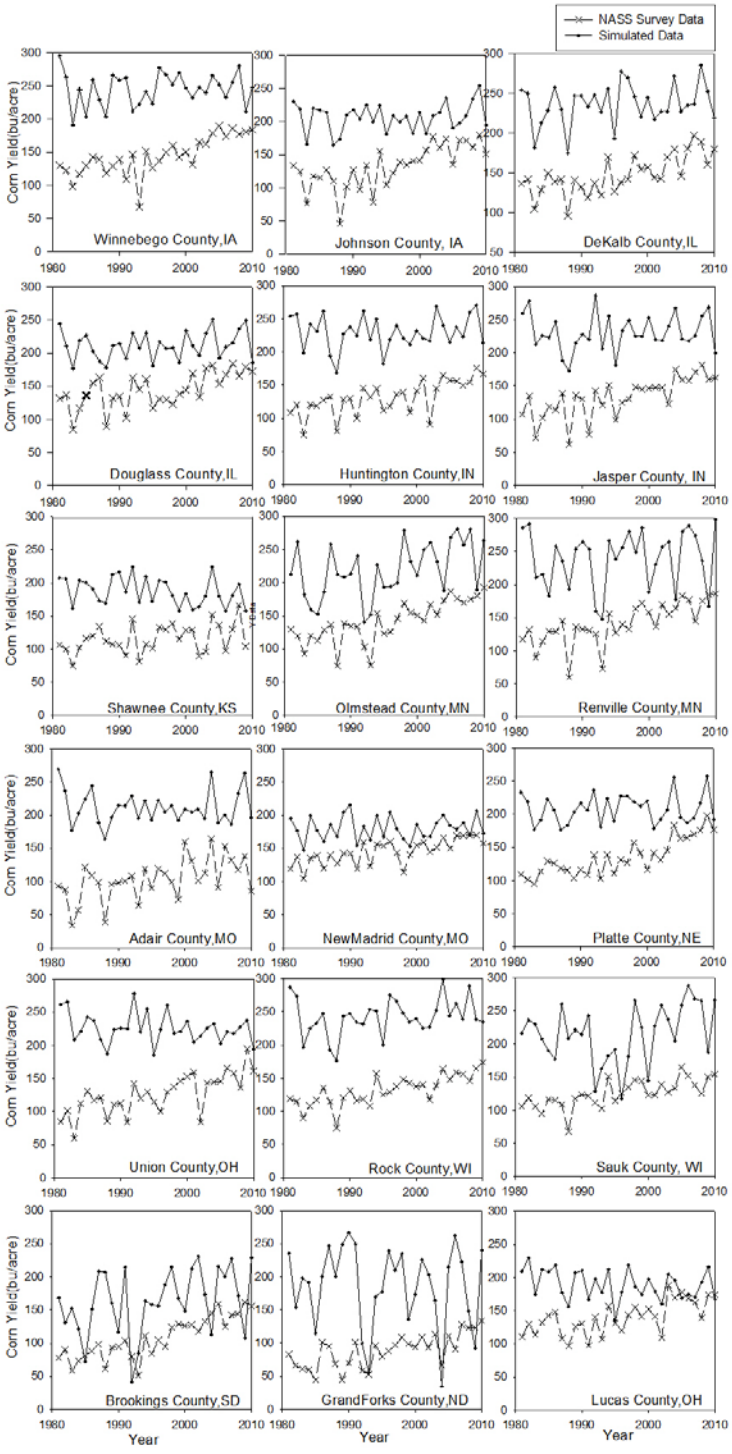

Figure 2.6 The Hybrid-Maize model validations at county scale for 18 sites across 30 years

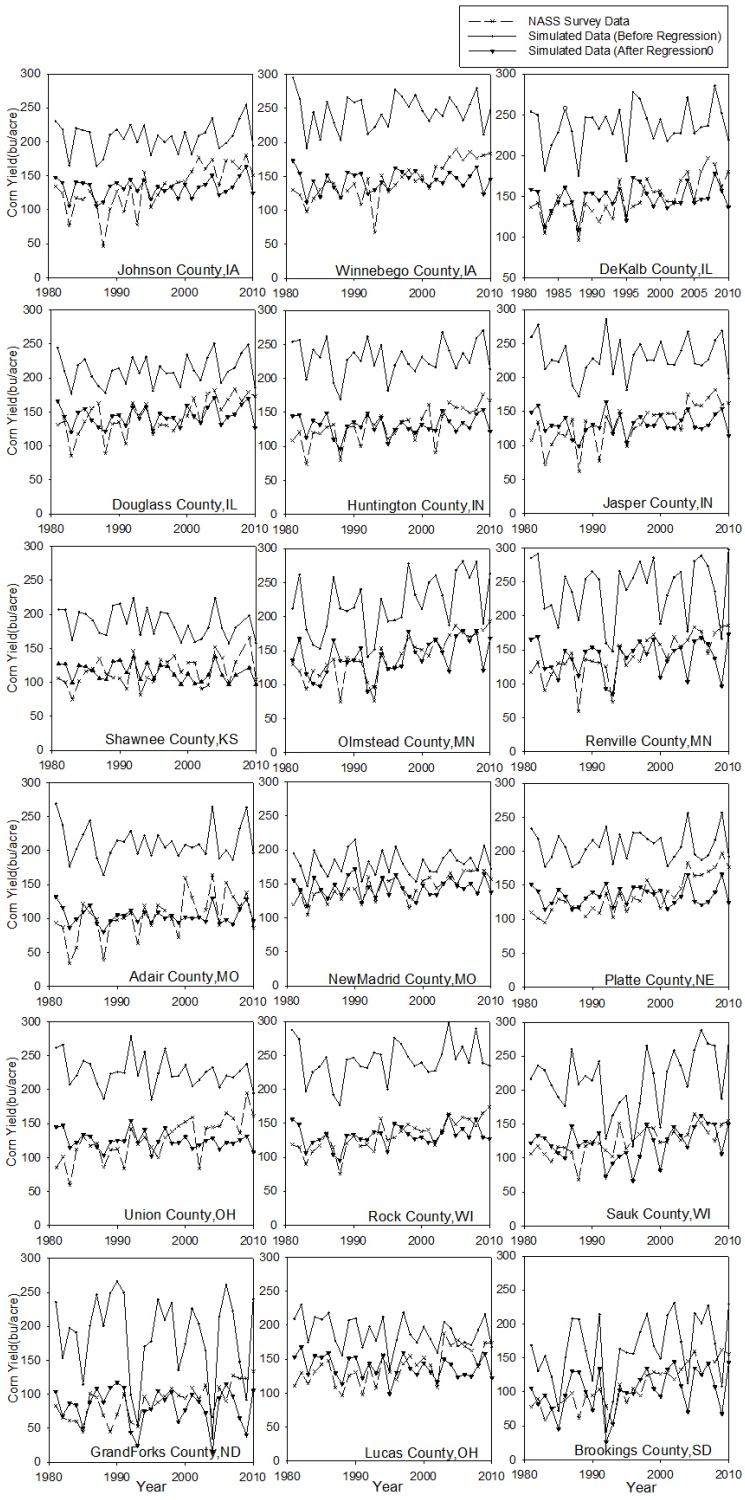

Figure 2.7 The simulated yield after regression with the survey data.

References

Niyogi, D.S., Raman, S., & Alapaty, K. (1997). A dynamic statistical experiment for atmospheric interactions. Environmental Modeling and Assessment, 2, 209 - 225.

Pickering, N. B., Hansen, J. W., Jones, J. W., Wells, C. M., Chan, V. K., & Godwin, D. C. (1994). WeatherMan: a utility for managing and generating daily weather data. Agronomy Journal, 86(2), 332-337.

Wallach, D., Makowski, G., & Jones, J. W., Eds. (2006). Working with dynamic crop models: Evaluation, analysis, parameterization, and applications. Elsevier Science, 462pp

CHAPTER 3. BUILDING A HIGH-RESOLUTION AGRO-METEOROLOGICAL DATABASE AND ESTIMATING CORN YIELDS REGIONALLY ACROSS THE US CORN BELT

Regional agro-meteorological applications are often constrained by the spatially discontinuous meteorological data from regular weather stations. Also, the application of crop models is often limited by the uncertainties of input hydro-meteorological data, such as solar radiation, soil moisture, soil temperature, evaporation/transpiration, and precipitation. These variables are routinely not available from weather stations except for specific experimental fields. Therefore, in this research, an approach has been developed which uses the Land Data Assimilation System (LDAS)/ High Resolution Land Data Assimilation System (HRLDAS) to build a high-resolution agro-meteorological database and then assimilate it into a crop growth model.

Research objectives are to build a high resolution agro-meteorological database and estimate corn yield regionally over the U.S. Corn Belt at grid scale. Developing such a high resolution database and modeling framework is expected to provide answers that are needed for agricultural/climatic regional impact assessments and decision support tools.

The hypotheses were:

1) This reanalysis agro-meteorological database can replace weather stations in regional agronomic applications.
2) Solar radiation from this agro-meteorological database has stronger agreement with observations than when developed from weather generators.
3) By providing such information, the performance of the crop model will be superior when applied at a regional scale.

To that end, this research validated the reanalysis meteorological data with site-measured data and validated model-simulated crop yield (driven by reanalysis meteorological data) with available NASS data for 20 sites across the Midwestern United States (Fig. 1.1). Figure 3.1 provides the overall methodology flowchart.

3.1 Data Resources and locations

As presented in the Fig. 3.1, in this research the meteorological data were collected from hourly NLDAS-2 forcing-A files in the Land Data Assimilation System (LDAS) from 1981-2012, across the Corn Belt at 1/8 degree resolution. Each file includes air temperature, downward shortwave radiation, precipitation, etc. (Table 1.2). In order to validate the agro-meteorological database, 30-years (1981-2010) of measured temperature data for 18 counties (Fig. 1.1) were collected from the National Climatic Data Center (NCDC), and 10-years (1997-2007) measured solar radiation data for Bondville, IL were collected from Ameriflux. 30-years corn yield of 18 counties (Fig. 1.1) were collected from the National Agricultural Statistics Service (NASS) annual survey.

3.2 Agro-metrological database building

3.2.1 HRLDAS running procedures

In running the HRLDAS, the first step was to collect the raw meteorological data from NLDAS-2 (32-km resolution) and land-surface initialized data (e.g., soil temperature, soil moisture, and canopy water content) from EDAS, then extract the required parameters separately into Grib files. Functions in WRF defined the model grids and provided the land use categories, terrain height, soil texture, and green vegetation fraction to HRLDAS. The second step was to downscale the raw meteorological data from 1/8 degreee resolution to 4-km resolution by running the consolidation module in HRLDAS. This step provided basic high-resolution meteorological data of every hour and initialized land-surface conditions for the first hour of each year, which are the input for running the last-step model. The "input" data across the Corn Belt contain a total of 419×530 = 222,070 grids. The parameters included in each grid are listed in Table 3.1. In this research, these basic hourly 4-km resolution meteorological data were grouped as "Database 1".

The last step was applying the 4-km resolution meteorological data to drive the Noah LSM to simulate the soil conditions (e.g., soil moisture, soil temperature), ET (evapotranspiration), etc. The "spin-up" time for Noah LSM in this research is 24 months (1979.01-1980.12).

The final outputs from HRLDAS are hourly and at 4-km resolution. In this research, the outputs across the Corn Belt contain a total of 222,070 grids. The parameters included in

each grid are listed in Table 3.2. Figure 3.2 presents the overall process of running HRLDAS. The hourly 4-km resolution output data are grouped as “Database 2”.

This research used temperature, precipitation, and solar radiation from Database 1 and soil temperature and soil moisture from Database 2. Figures 3.3 and 3.4 present the sample parameter images from Database 2.

3.2.2 Data extraction and origination

One of the goals in this research is to build a high-resolution agro-meteorological database which can be easily applied to crop models and other agronomic decision tools. The minimum requirements of meteorological inputs for crop models (e.g., the Hybrid Maize model) include daily minimum temperature, daily maximum temperature, solar radiation, and precipitation. Therefore, to meet the needs of crop models, data extraction from the hourly database into daily data was required. In this research, NCAR command Language (NCL) was the major programming language in the data extraction process. Unit conversion was also applied during the extraction process.

The data extraction from hourly to daily has been applied with air temperature, precipitation, solar radiation, soil moisture, soil temperature, etc. (Fig. 3.5). “Database 3” is compiled using these daily data (Table 3.3). Figure 3.6 presents the sample parameter image from Database 3.

It is notable that the time zone of Database1 and Database 2 is Universal Time Zone. Because the daily meteorological data will be applied at local time, some bias might exist caused by gaps in the time zones. Here the data hasn’t been corrected to local time

because: 1) the research area crossed 3 time zones (Eastern Time Zone, Central Time Zone, and Mountain Time Zone), and is complicated to correct in this preliminary study. 2) It is a high possibility that daily maximum and minimum temperature are not influenced by the time zone gap. For example, at Universal Time Zone, today is defined from 00:00 to 00:00, while converted to the Eastern Time Zone the local time is from yesterday's 19:00 to today's 19:00. The daily maximum and minimum temperature is usually included during this time period.

Simply said, when accessing this agro-meteorological database (Database 3), users just need to provide the location's coordinates, and then the system will extract the requested data of that specific location.

3.3 Meteorological data validation

3.3.1 Temperature validation

In order to test the reliance of this agro-meteorological database (Database 3), several validations were applied in this study, for example, daily maximum temperature and daily minimum temperature from Database 3 of Johnson County, IA was compared with the site observations. The validation results (Fig. 3.7) indicate that the reanalysis daily maximum and minimum temperature in Database 3 have strong agreement with the observations, for maximum temperature, r^2= 0.97, for minimum temperature, r^2 = 0.95.

3.3.2 Solar radiation validation

As mentioned in Chapter 1, crop models are often limited by the lack of solar radiation data. A major part of this agro-meteorological database (Database 3) provides daily solar

radiation data, which can be used by not only crop models, but also other agronomic decision tools.

The solar radiation data from Database 3 was compared with the observed solar radiation data of Bondville, IL which were collected from Ameriflux. The validation results (Fig. 3.8) indicate that the reanalysis solar radiation data from Database 3 fit well with the measured real data ($r^2 = 0.81$). This study also validates the solar radiation from the weather generator (WeatherAid, Yang et al. 2005) where the r^2 between generated solar radiation and measured observations is 0.67 (Fig. 3.9), which indicates that solar radiation data from this agro-meteorological database (Database 3) are better than solar radiation generated by the weather generator.

3.4 Gridded crop model running system --- estimating corn yield regionally across Corn Belt with the agro-meteorological database

After the meteorological data validations, it has been shown that the meteorological data from the Agro-meteorological database (Database 3) are reliable. In this section, a process of estimating corn yield regionally at 4-km resolution will be illustrated.

3.4.1 Validation of simulated corn yield at county scale.

In Chapter 2, the Hybrid-Maize model validations were driven by weather station meteorological data. In that chapter, new validations of the crop model were driven by meteorological data from reanalysis data based on the agro-meteorological database (Database 3). These new validations were performed to ensure the reliability of reanalysis

database use in crop modeling, which is important for applying this this database in regional corn yield estimations.

In these new validations, except for the meteorological input data, other model settings (e.g., water condition, planting date) are kept the same as those used for the validations in chapter 2; it helps to clarify whether this agro-meteorological database is superior to traditional station data.

Based on the regression analysis results in Chapter 2, each simulated corn yield in this chapter has been rescaled by a factor of 0.6. The averaged MAE for 18 sites using meteorological input from the weather station is 1.25 Mg/ha while the averaged MAE derived using meteorological input from Database 3 is 1.27 Mg/ha (Table 3.3). The one-way ANOVA tests between simulated corn yield driven by meteorological input from weather station data and 30-years simulated corn yield driven by reanalysis meteorological input from Database 3 (Table 3.4) report that except for Olmstead County, MN and Sauk County, WI, the P-Values of the other 16 counties are larger than 0.05. This means at the 95% confidence interval, for most counties, there is no significant difference between the two driven scenarios. The results indicate this reanalysis agro-meteorological database (Database 3) has great potential when expanding to regional corn yield simulations.

3.4.2 Estimating corn yield across Corn Belt at 4-km resolution

The research domain contains a total of 222,070 4 km ×4 km grids. Consequently, running the crop model across the whole area needs to create 222,070 weather input files.

However, not all the grids are cropland. In order to extract the non-cropland grids, a mask file based on USGS land-use categories has been created. Therefore, in this research, the regional corn yield simulations were only applied on the cropland, which is more precise than simulating every single grid. The total cropland 4 km ×4 km grids total around 85,000 across the Corn Belt.

After creating the input files, the Hybrid-Maize model will run automatically using a script. Because this is preliminary research, the management settings (e.g., planting date and plant density) of the Hybrid-Maize model are the same for every cropland grid. Other parameter settings are the same as the county-scale simulations in Chapter 2. Figure 3.10 depicts the overall process while the sample figure of gridded yield output is presented in Fig. 3.11. The frequency histogram (Fig.3.12 (a)) of NASS surveyed corn yield of the U.S. Corn Belt (2003) shows the highest frequencies of yield are located between 120(bu/acre) ~ 160(bu/acre). The histogram of the estimated yield (Fig.3.12 (b)) illustrates the highest frequencies of yield are distributed between100 (bu/acre) ~120(bu/acre).The result indicates at regional scale, the model was under estimated the corn yield, and the histogram also shows the model cannot catch the extreme events (extremely high or low) of the corn yield. The possible reason is during this preliminary gridded crop running, the planting date was set as May 1st for all the grid points.

3.5 Case study based on the gridded yield estimation system --- the impacts of planting date on corn yield

In order to optimize corn yield and make corn replanting decisions, it is important to know the corn yield response to different planting dates (Nafziger1994). Based on the

gridded corn yield estimation process, the Hybrid-Maize model was running under different planting dates in 2003: April 1st, May 1st, and June 1st. MultiSpec (V 3.3. Biehl and Landgrebe 2002) was applied in analyzing the gridded yield outputs.

The yield frequency histogram of the 85,000 cropland grids across the Corn Belt (Fig. 3.13) indicates that under each planting date, the highest frequency of the yield was 100-120 bu/acre. When model running under planting on Aril 1st, the estimated yield data show a higher frequency of reach to 120-140 bu/acre and 140-160 bu/acre than planting on May 1st and June 1st. Planting on June 1st can bring the highest frequency to 160-180 bu/acre, but it also results in the highest frequency in low yield (50-100 bu/acre) demonstrating that late planting is acceptable for some areas but can hurt the yield in other areas. Although the overall performance in the histogram shows planting on April 1st is better than the other two dates, it is still improper to conclude that based on the model estimations, April 1st is the best planting date.

In MultiSpec, the three gridded yield image outputs of different planting dates had been combined into a single multispectral image file with three 3 channels (channel 1: Planting on April 1st, channel 2: Planting on May 1st, channel 3: Planting on June 1st). Channel 2 minus channel 1 is the model-simulated yields responding to a change in planting dates from April 1st to May 1st. Similarly, channel 3 minus channel 2 is the model-simulated yields responding to a planting date change from May 1st to June 1st. The results (Fig. 3.14) indicate that the impact of planting date on corn yield is varied for different areas. For example, in Fig. 3.14 (a), when the planting date changes from April 1st to May 1st, the estimated corn yield data for the majority of Iowa are decreased while the yield data

of Michigan are increased, indicating that the best planting date for Michigan is later than Iowa. However, it doesn't mean that for Michigan, the later planting date can bring better results. In Fig. 3.14(b), when the planting date changes from May 1st to June 1st, the estimated yield data of Michigan are no longer increased. On the contrary, the yield data are decreased. It can be concluded that in 2003, the best planting date for the majority of Michigan is a day or several days during May 1st to June 1st. For further applications, this gridded crop model running system can test every single planting date to pick up the "best planting date".

It also notable that in Fig. 3.14, the yield varied range of changing planting date from May 1st to June 1st is doubled of changing planting date from April 1st to May 1st. This result indicates that late planting dates bring more uncertainty or risk regarding the corn yield, and caution is needed when making late planting decisions.

The most important advantage of this gridded crop model system is to provide a regional perspective on the impacts of meteorological factors in the simulation of crop growth. Although there are several limitations of this preliminary gridded model system, such as ignored soil moisture, soil type and non-dynamic plant density, it still has great potential for wide use in agronomic and agro-economic applications, and in future studies, soil moisture will be added.

3.6 Extended application of the Agro-meteorological database---Growing degree days map

Growing degree days (GDD), as heat units, are often used to describe and predict crop growth stages (Miller et al. 2001; Swan et al. 1987). The basic equation of daily GDD is:

$$\text{GDD} = \frac{(T_{MAX}+T_{MIN})}{2} - T_{BASE} \quad (1)$$

Where T_{MAX} is the daily maximum temperature and T_{MIN} is the daily minimum temperature, T_{BASE} is the base temperature for plant growth and plant growth will be limited when the temperature is below T_{BASE} (McMaster and Wilhelm 1997). Different plant species have different T_{BASE} (Wang 1960). Because this research focuses on corn, the T_{BASE} of corn is 10˚C (Cross and Zuber 1972). The GDD for corn is often calculated with the upper temperature threshold (T_{UT}). In this research, T_{UT} was set at 34˚C; the default value of the Hybrid-Maize model. To calculate the GDD, the methods used in this research are: (1) If $\frac{(T_{MAX}+T_{MIN})}{2}$ is less thanT_{BASE}, then GDD = 0; (2) If T_{MAX} is larger than T_{UT}, then Tmax = T_{UT}.

An NCL script was used to calculate the daily GDD of every single grid in the whole research domain. The total GDD had been accumulated from the planting date. In this particular study, the planting date was set as May 1st. Because in the Hybrid Maize model, GDD (T_{BASE} =10˚C) = 1389 was considered as maturity for corn, GDD = 1389 is the reference line in the GDD maps. In order for ease of use by U.S. corn producers, the temperature unit was converted to Fahrenheit, so the reference line is GDD (T_{BASE} =

50˚F) = 2500. Sample GDD maps are listed in Fig. 3.15, which can be the reference for estimating corn harvest date.

3.7 Summary

The goal of building this high- resolution agro-meteorological database is to bring available reanalysis meteorological information from the Land Data Assimilation System (LDAS) to usable agronomic applications, such as crop models. Through interpolating data from 32-km into 4-km and running Noah-LSM by the High- Resolution Land Data Assimilation System, an hourly database was created. To meet with the needs of most agronomic applications, finally a daily database of 32 years (1981- 2012) was built, which includes daily maximum temperature, daily minimum temperature, solar radiation, precipitation, etc. The validations of meteorological parameters in the agro-meteorological database show a strong agreement between the reanalysis data and site observations. Data from this database are a better fit with observed data especially for solar radiation when compared with weather generator data. Validations of estimated corn yield show that there is no significant difference between the crop model driven by meteorological inputs from this database and from weather stations. These results give confidence to widely apply this high-resolution agro-meteorological database in agronomic applications, which not only can save time in data collecting , but the spatially continuous dataset can also help to understand how meteorological factors influence crop growth at the regional scale.

Based on this high-resolution agro-meteorological database, a gridded crop model system has been developed. The preliminary gridded yield outputs provide a regional perspective

on corn yield simulation which can also be applied in related studies, such as the impact of planting date on corn yield.

This agro-meteorological database has great potential for wide application in agronomic and agro-economic areas and is not limited to combining it with a crop model to estimate regional corn yield at grid-scale or for developing GDD maps. Moreover, owing to similar formats between datasets from LDAS and the North American Regional Climate Change Assessment Program (NARCCAP), it will be time efficient to combine future projections with this database.

References

Biehl, L., & Landgrebe, D. (2002). MultiSpec—a tool for multispectral–hyperspectral image data analysis. Computers & Geosciences, 28(10), 1153-1159.

Cross, H. Z., & Zuber, M. S. (1972). Prediction of flowering dates in maize based on different methods of estimating thermal units. Agronomy Journal, 64(3), 351-355.

McMaster, G. S., & Wilhelm, W. W. (1997). Growing degree-days: one equation, two interpretations. Agricultural and Forest Meteorology, 87(4), 291-300.

Miller, P., Lanier, W., & Brandt, S. (2001). Using growing degree days to predict plant stages. Montana State University, USA. Extension Service.

Nafziger, E. D. (1994). Corn planting date and plant population. Journal of production agriculture, 7(1), 59-62

Swan, J. B., Schneider, E. C., Moncrief, J. F., Paulson, W. H., & Peterson, A. E. (1987). Estimating corn growth, yield, and grain moisture from air growing degree days and residue cover. Agronomy journal, 79(1), 53-60.

Wang, J. Y. (1960). A critique of the heat unit approach to plant response studies. Ecology, 41(4), 785-790.

Yang, HS; Dobermann, A; Cassman KG and Walters, DT. (2005).WeatherAid: A Software for Weather Data Management. Universityof Nebraska – Lincoln

Table 3.1 Parameters in HRLDAS input files (Database 1).

Name	Unit	Description
Included in each hourly file		
T2D	K	Temperature at 2 m
Q2D	kg kg{-1}	Specific Humidity at 2 m
U2D	m s-1	Horizontal wind speed at 10 m
V2D	m s-1	Vertical wind speed at 10 m
PSFC	Pa	Surface Pressure
RAINRATE	mm	Rainrate
SWDOWN	W m{-2}	Downward short-wave radiation flux
LWDOWN	W m{-2}	Downward long-wave radiation flux
Included in the first-hour file of each day		
WEASD	kg m{-2}	Water equivalent snow depth
VEGFRA	%	green vegetation fraction
Included in the first-hour file of each year		
SMOIS_1	kg m-3	Soil Moist 0-10 cm below ground layer
SMOIS_2	kg m-3	Soil Moist 10-40 cm below ground layer
SMOIS_3	kg m-3	Soil Moist 40-100 cm below ground layer
SMOIS_ 4	kg m-3	Soil Moist 100-200 cm below ground layer
STEMP_1	K	Soil temperature 0-10 cm below ground layer
STEMP_2	K	Soil temperature 10-40 cm below ground layer
STEMP_3	K	Soil temperature 40-100 cm below ground layer
STEMP_4	K	Soil temperature 100-200 cm below ground layer
CANWAT	kg m-2	Plant Canopy Surface Water
GVFMIN	%	Minimum green vegetation fraction
GVFMAX	%	Maximum green vegetation fraction
Z2D	m	

Table 3.2 Parameters in HRLDAS hourly output files (Database 2)

Name	Unit	Description
IVGTYP	category	Dominant vegetation category
ISLTYP	category	Dominant soil category
SKINTEMP	K	Skin temperature
CANWAT	mm	Canopy water content
SOIL_T (4-layers)	K	soil temperature
SOIL_M (4-layers)	m{3} m{-3}	volumetric soil moisture
SOIL_W (4-layers)	m{3} m{-3}	liquid volumetric soil moisture
SOIL_MX	mm	total column soil moisture
SFCRNOFF	mm	Accumulated surface runoff
UGDRNOFF	mm	Accumulated underground runoff
INTRFLOW	mm	Accumulated interflow runoff
SFCEVP	mm	Accumulated evaporation from surface
ETAKIN	mm	Evapotranspiration
CANEVP	mm	Accumulated canopy evaporation
EDIRX	mm	Accumulated direct soil evaporation
ETTX	mm	Accumulated plant transpiration
ALBEDX	fraction	Albedo
WEASD	m	Water equivalent snow depth
ACRAIN	mm	Accumulated precipitation
ACSNOM	mm	Accumulated snow melt
ESNOW	mm	Accumulated evaporation of snow
DRIP	mm	Accumulated canopy drip
DEWFALL	mm	Accumulated dewfall

Table 3.2 Continued

SNODEP	m	Snow depth
VEGFRA	fraction	Green vegetation fraction
Z0	m	Roughness length
HFX	W m{-2	Upward surface sensible heat flux
QFX	W m{-2}	Upward surface latent heat flux
GRDFLX	W m{-2}	Ground heat flux at surface
SW	W m{-2}	Downward shortwave radiation flux
LW	W m{-2}	Downward longwave radiation flux
FDOWN	W m{-2}	Radiation forcing at the surface
XLAI	dimensionless	Leaf area index
SNOTIME	s	Snow age
EMBRD	s	Background Emissivity
SNOALB	fraction	Maximum albedo over deep snow
NOAHRES	W m{-2}	Residual of surface energy balance
CH		Heat Exchange Coefficient

Table 3.3 Parameters in daily files (Database 3).

Name	Unit	Description
Tmax	C	Daily maximum temperature
Tmin	C	Daily minimum temperature
SR	MJ m{-2}	Daily solar radiation
Prep	mm	Daily precipitation
Soil_M	m{3} m{-3}	Daily averaged soil moisture
Soli_T	k	Daily averaged soil temperature
ET	mm	Daily evapotranspiration

Table 3.4 30-year mean absolute error (MAE) of corn yield simulations

County	MAE1 (model driven by station input)	MAE2 (model driven by reanalysis input)
Johnson, IA	1.05	1.02
Winnebago,IA	1.03	1.07
DeKalb, IL	0.90	1.13
Douglass, IL	1.16	1.18
Huntington,IN	0.79	0.81
Jasper,IN	0.86	0.85
Shawnees, KS	0.92	1.02
Olmstead, MN	1.24	0.97
Renville, MN	1.22	1.08
Adair, MO	1.68	1.58
NewMadrid, MO	2.28	2.39
Platte, NE	0.86	1.36
Union, OH	1.04	1.11
Rock, WI	1.03	0.84
Sauk,WI	1.38	1.32
GrandForks , ND	2.11	2.70
Lucas , OH	0.83	0.88
Brookings, SD	2.12	1.57
average	1.25	1.27

Table 3.5 -Value from one-way ANOVA test between 30-years simulated corn yield driven by meteorological input and reanalysis meteorological input form Database 3

County	P-Value between two different simulation-driven scenarios
Johnson, IA	0.13
Winnebago,IA	0.47
Dekalb, IL	0.51
Douglass, IL	0.73
Huntington,IN	0.7
Jasper,IN	0.61
Shawnees, KS	0.09
Olmstead, MN	**0.005**
Renville, MN	0.4
Adair, MO	0.07
NewMadrid, MO	0.48
Platte, NE	0.06
Union, OH	0.5
Rock, WI	0.14
Sauk,WI	**0.01**
GrandForks, ND	0.04
Lucas, OH	0.29
Brookings, SD	0.96

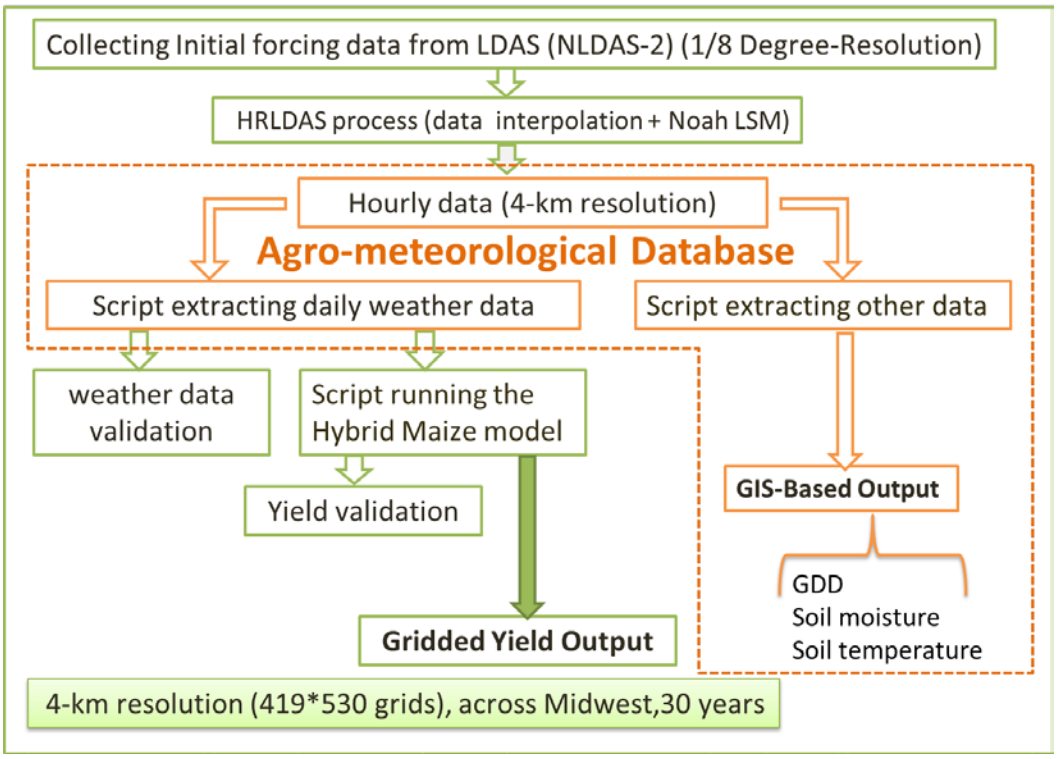

Figure 3.1 Methodology flowchart for chapter 3

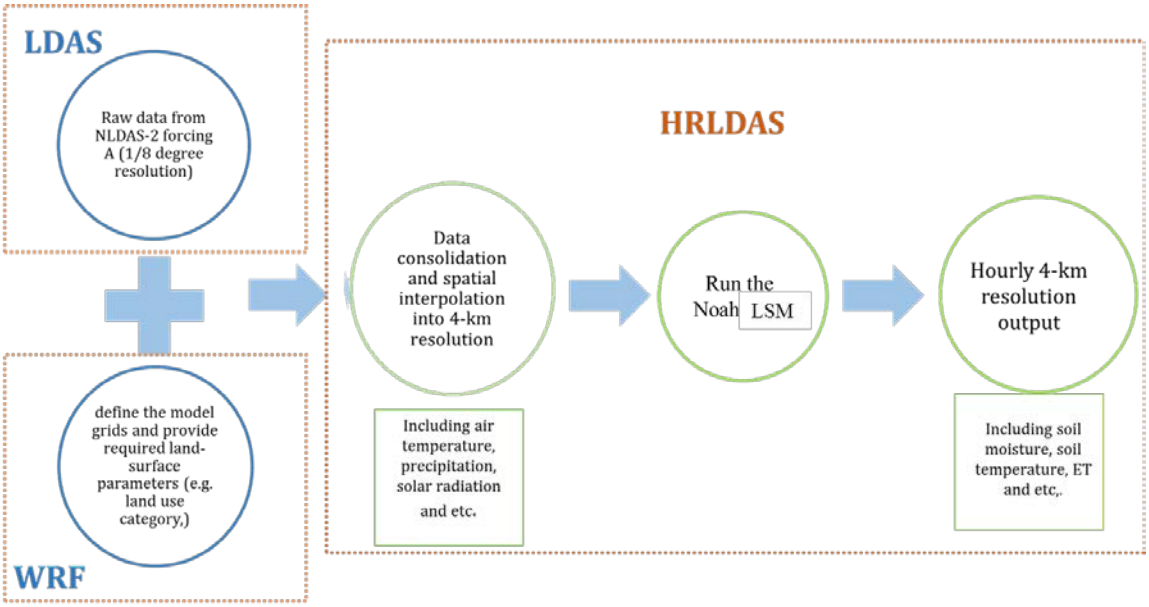

Figure 3.2 The overall process of running the HRLDAS

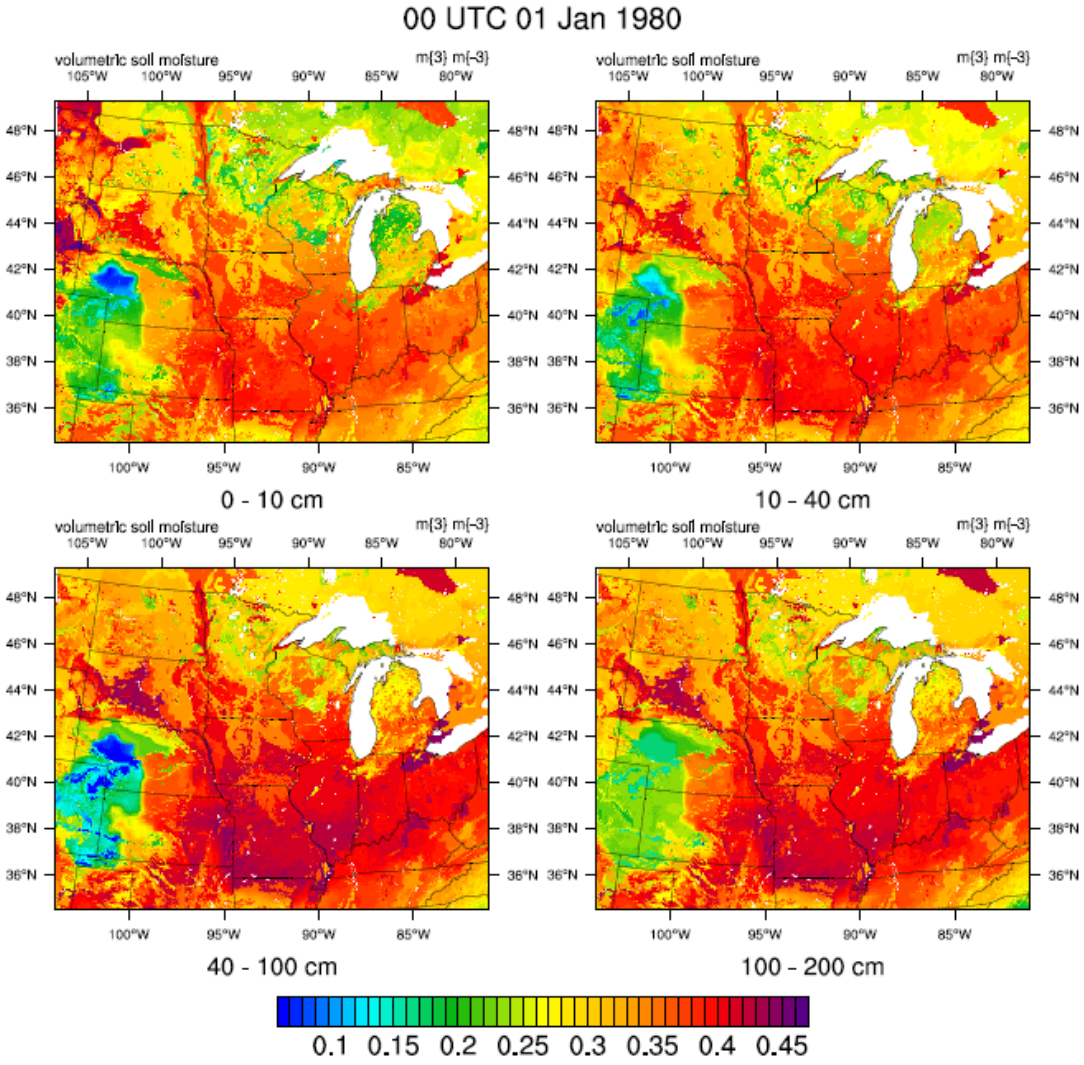

Figure 3.3 Sample image of 4-layer soil moisture from Database 2

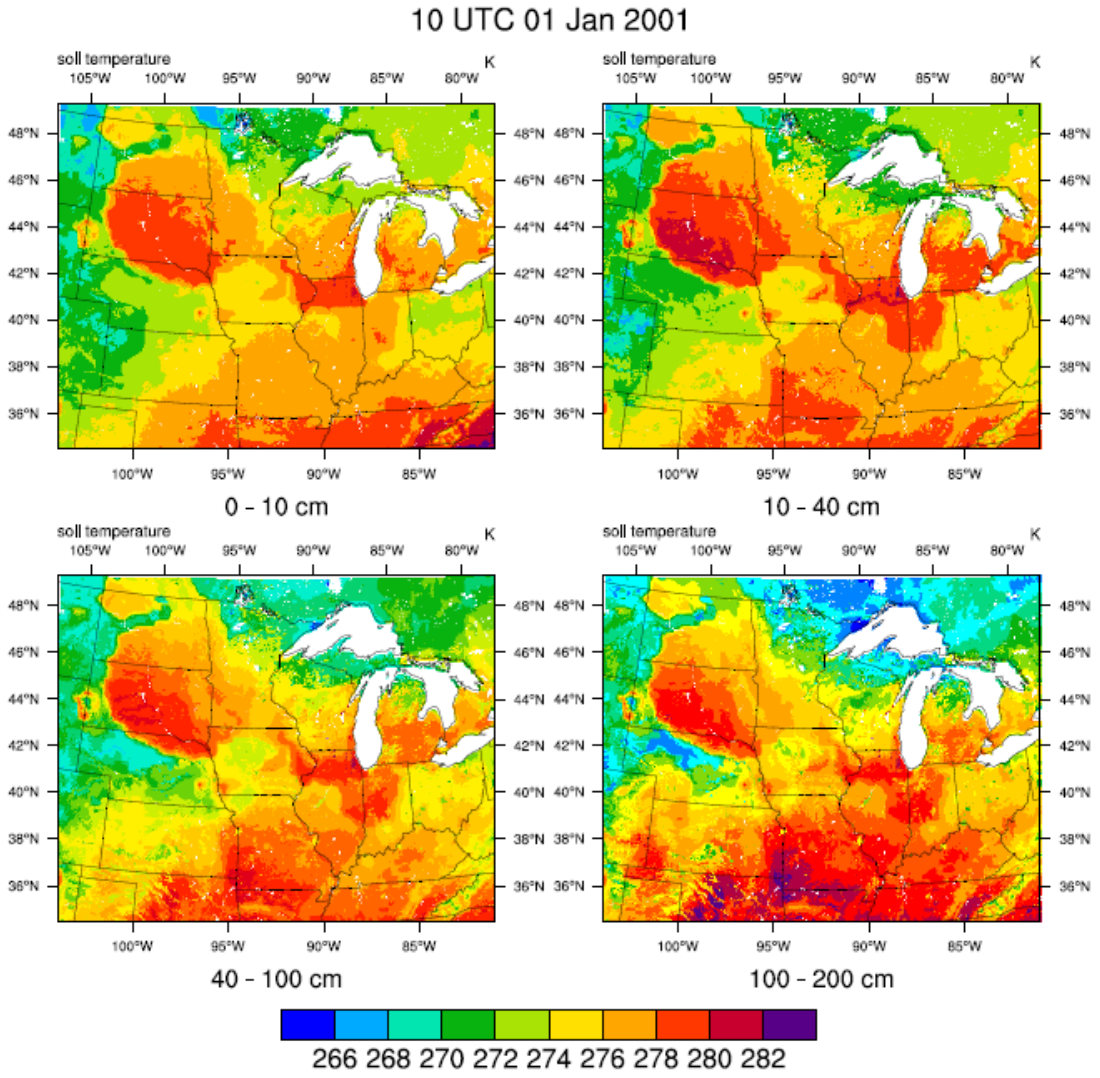

Figure 3.4 Sample image of 4-layer soil moisture from Database 2

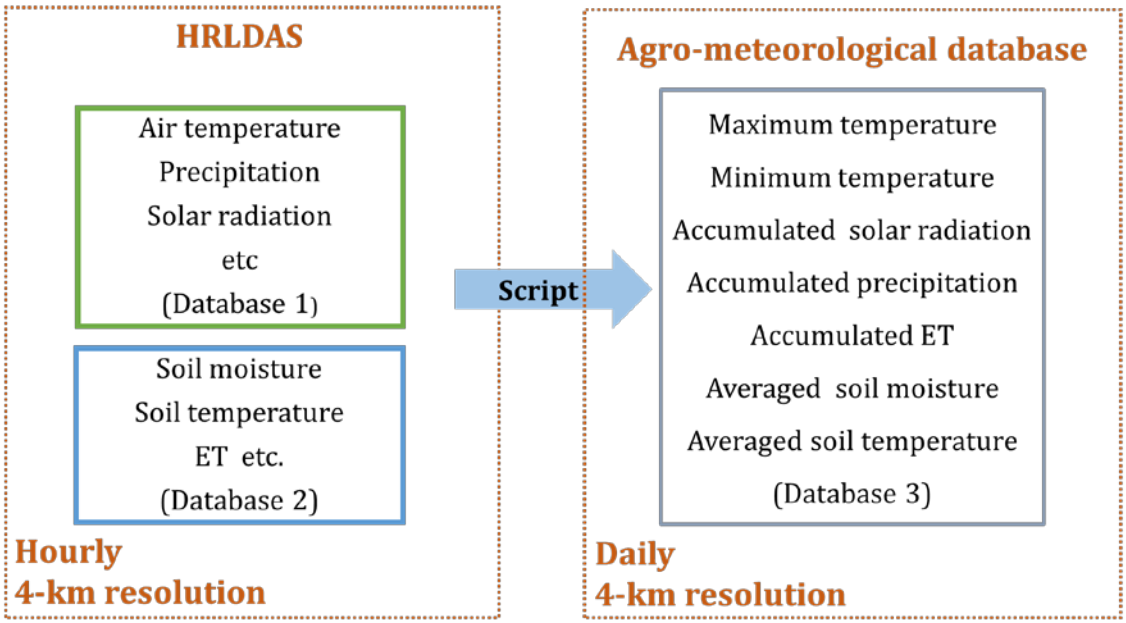

Figure 3.5 Building an agro-meteorological database (Database 3) from HRLDAS

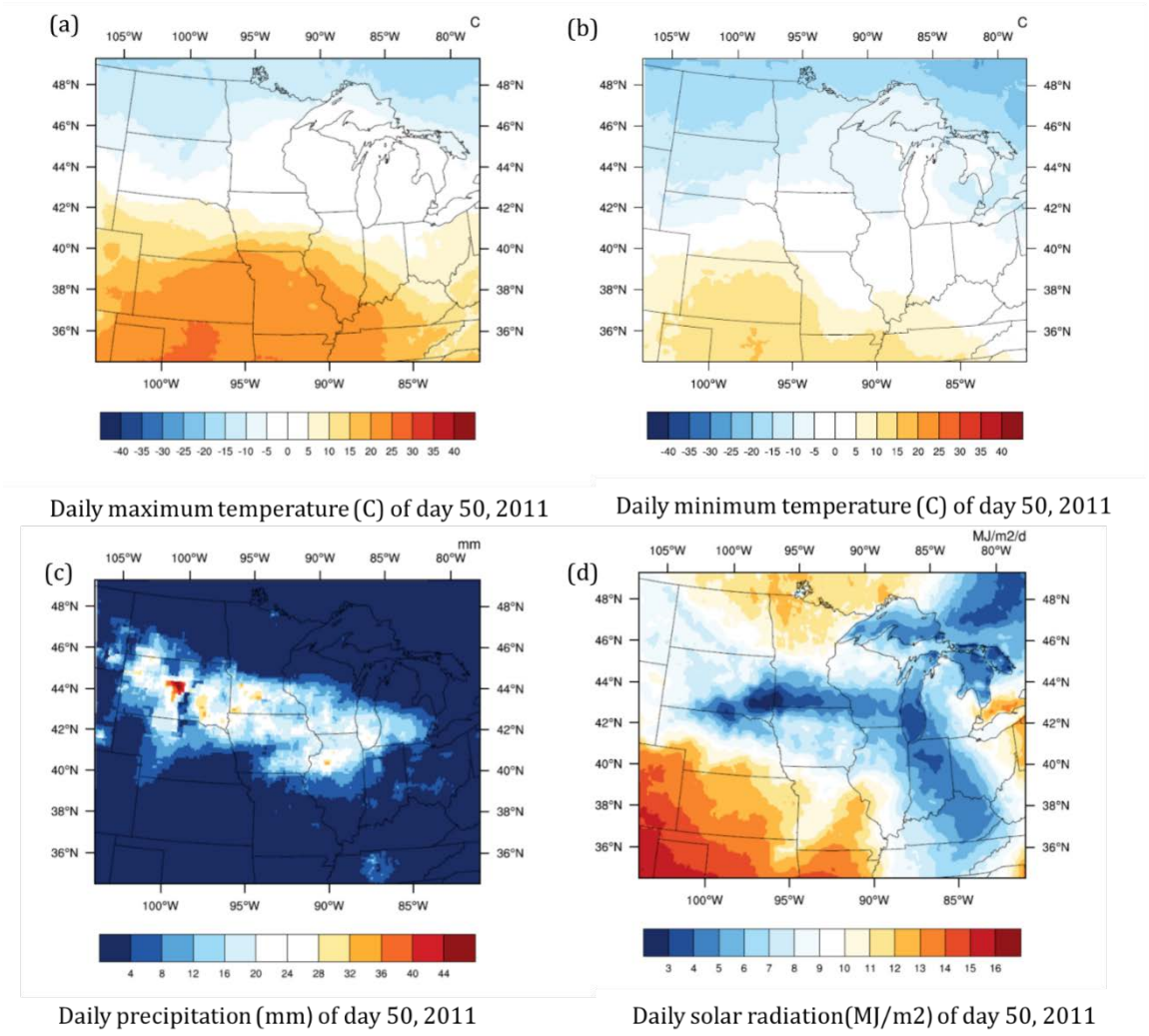

Figure 3.6 Sample images of parameters in Database 3

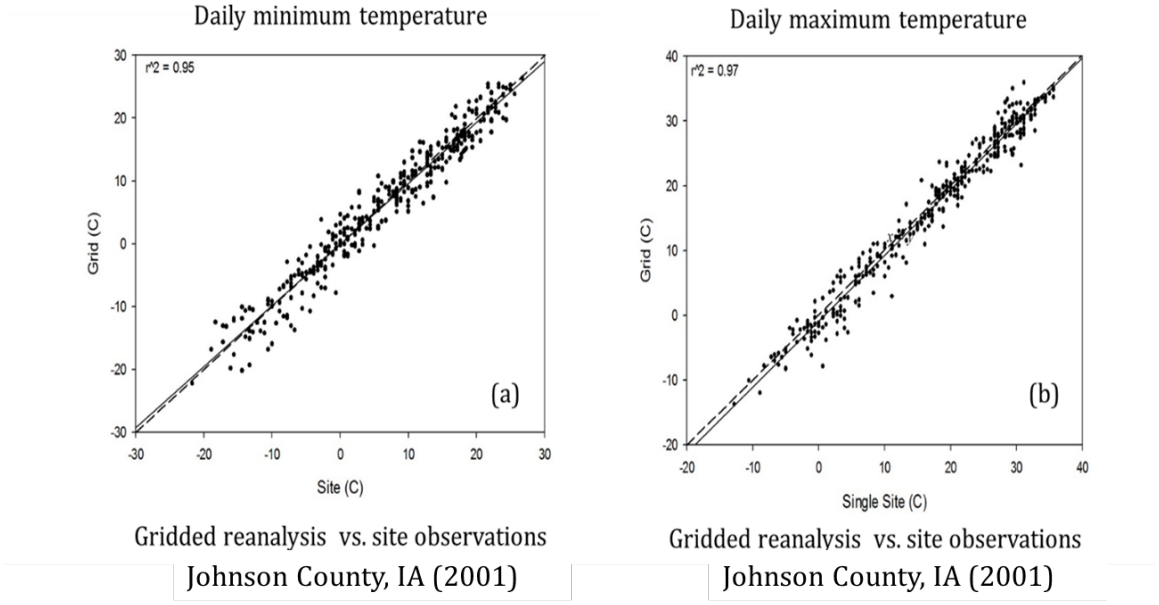

Figure 3.7 Validations of daily maximum and minimum temperature from Database 3.

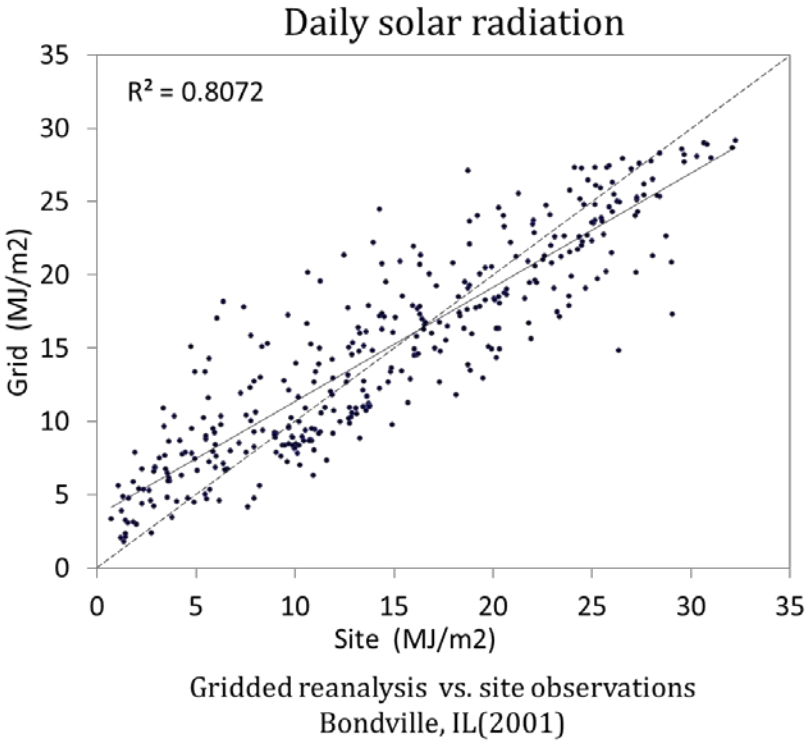

Figure 3.8 Validations of solar radiation from Database 3.

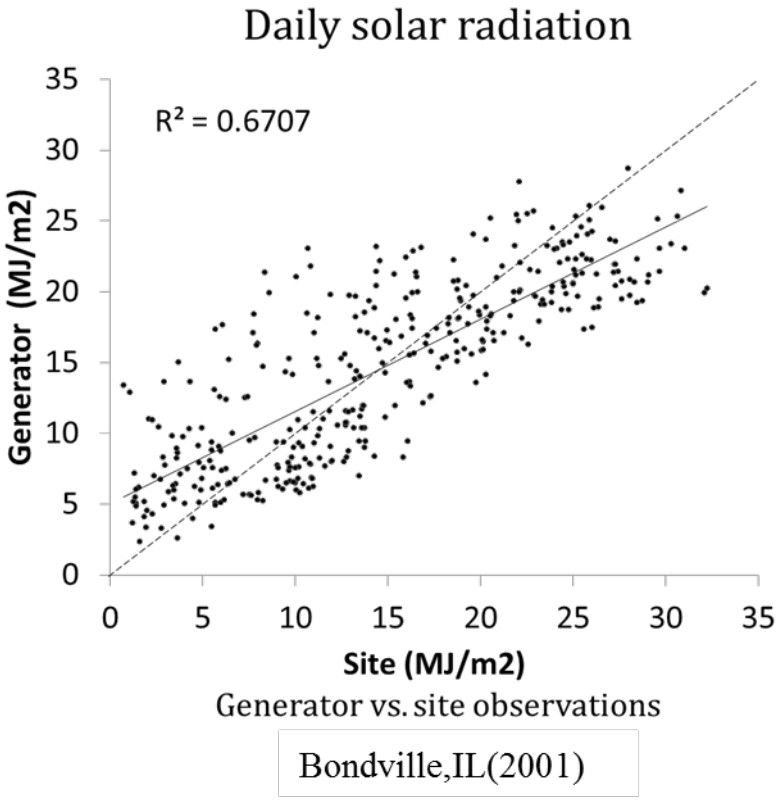

Figure 3.9 Validations of solar radiation from the solar radiation generator.

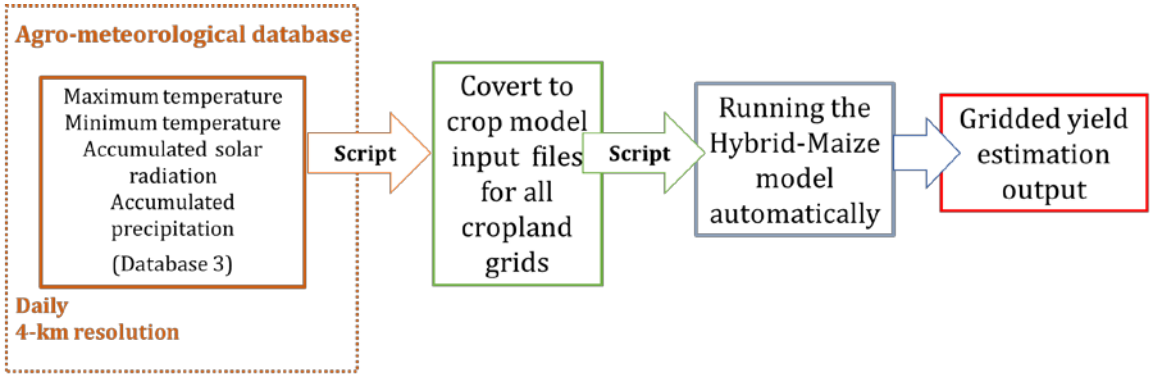

Figure 3.10 Process of running the Hybrid-Maize model at grid scale.

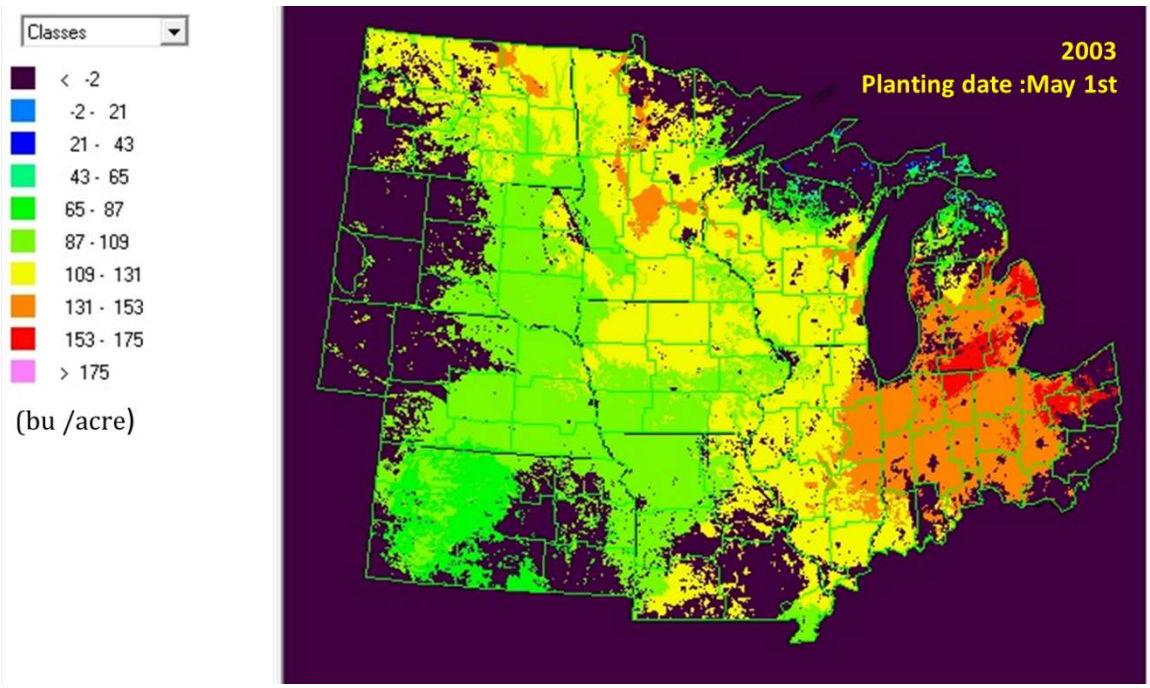

Figure 3.11 Sample image of gridded yield (bu/acre) estimation output

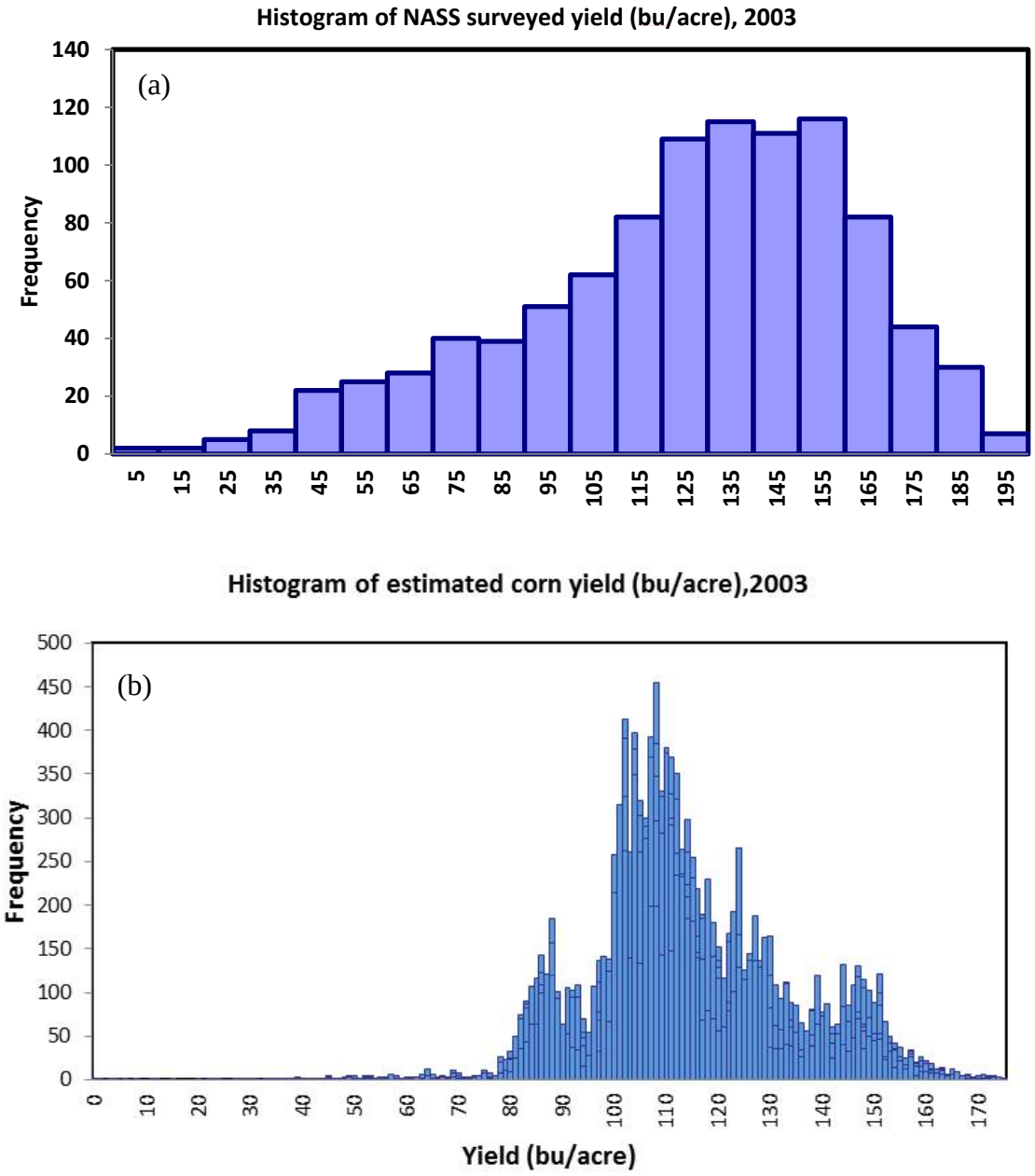

Figure 3.12 (a) Histogram of NASS surveyed yield across the U.S. Corn Belt (2003). (b) Histogram of grid-scale estimated corn yield (2003, planting date May 1st)

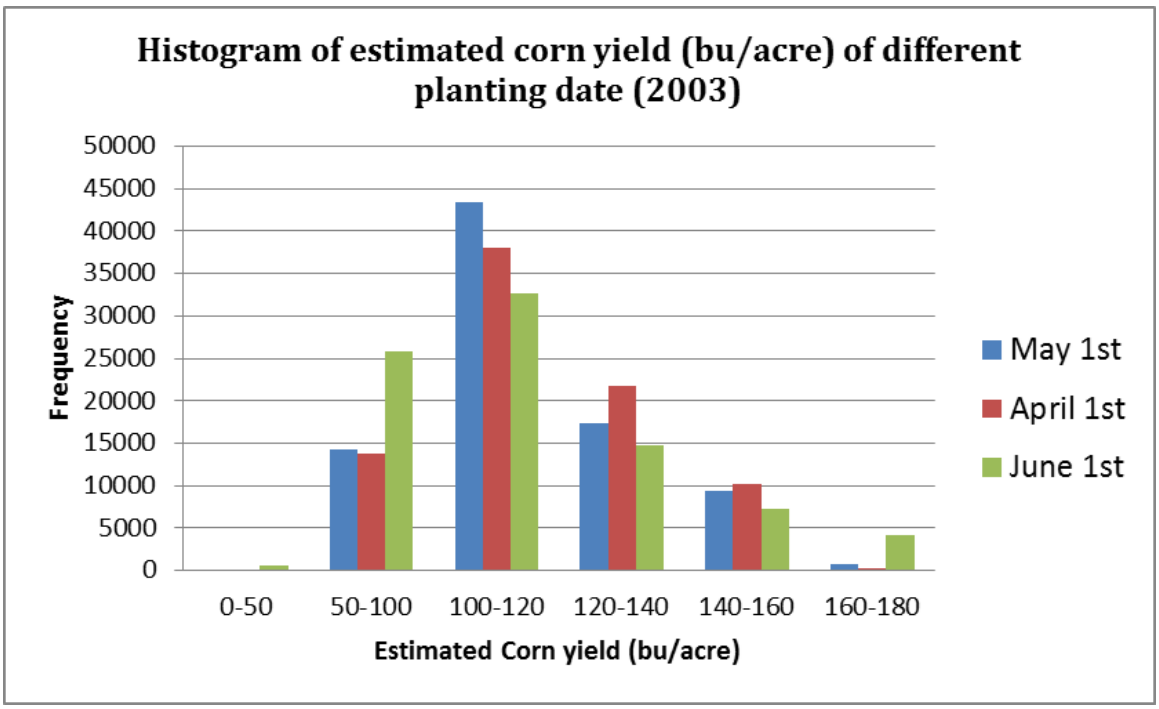

Figure 3.13 Histogram of grid-scale estimated corn yield under different planting date

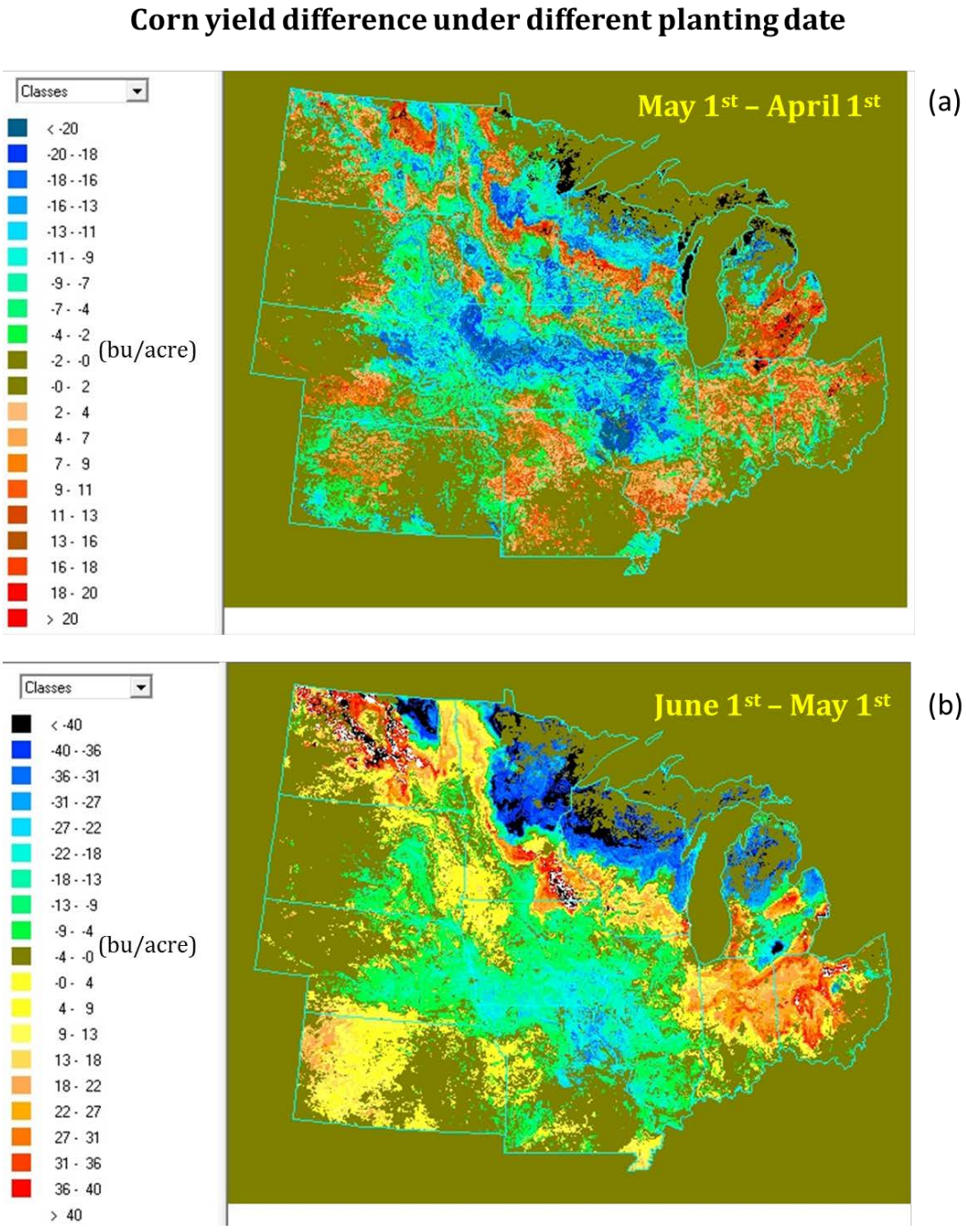

Figure 3.14 The estimated corn yield difference between different planting dates. (a) Planting on April 1st, 2003 compared with planting on May 1st, 2003. (b) Planting on May 1st, 2003 compared with planting on June 1st, 2003.

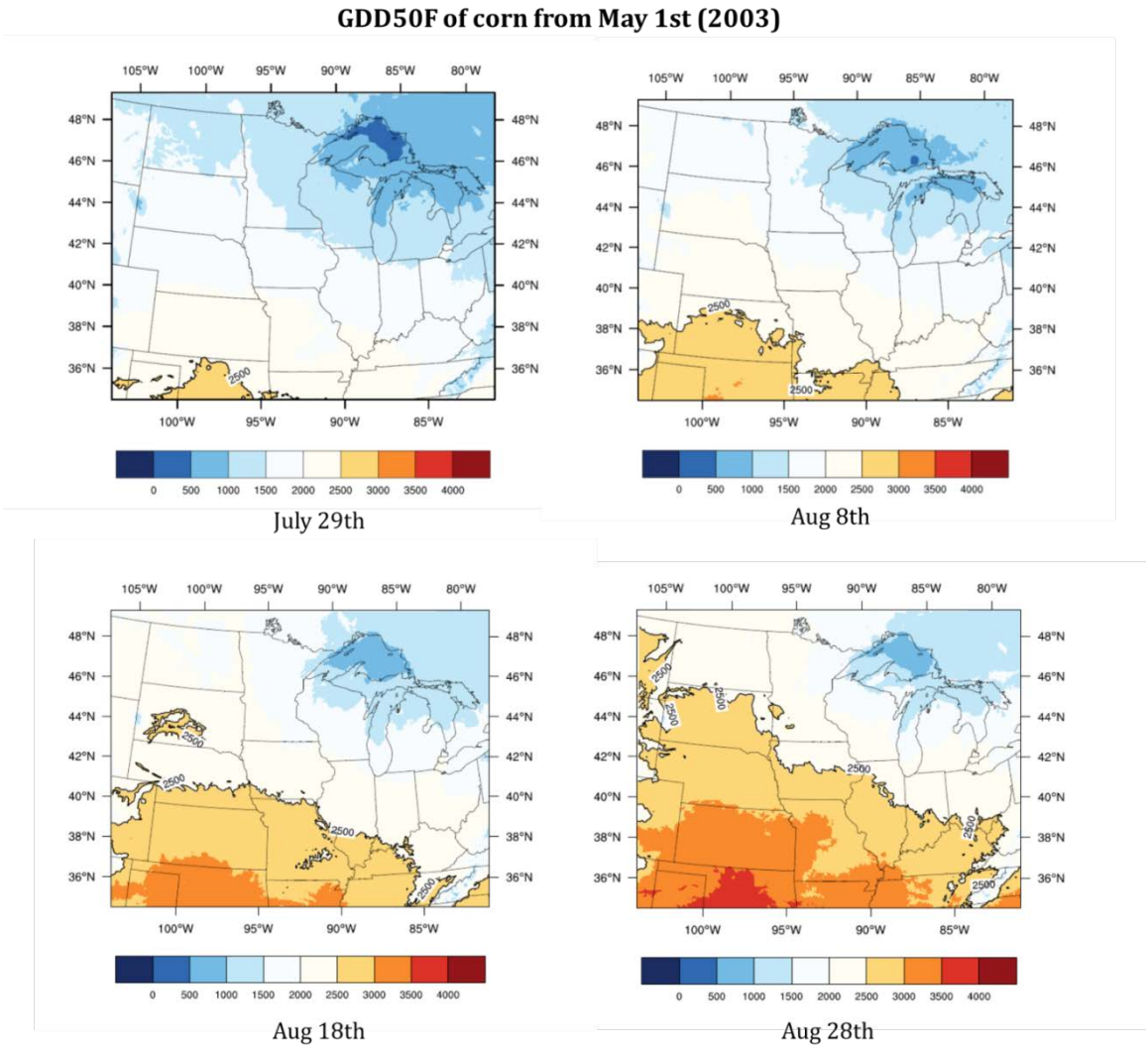

Figure 3.15 GDD50F map of the U.S. Corn Belt from May 1, 2003.

CHAPTER 4. EL NIÑO–SOUTHERN OSCILLATION (ENSO) WITH CORN AND CORN SIMULATION MODEL IN U.S. CORN BELT

As mentioned in Chapter 1, many studies indicate that ENSO has a significant impact on crop yield. In this chapter, three topics will be discussed: (1) The impact of the El Niño–Southern Oscillation (ENSO) on corn yield and corn planting date. (2) Can the crop model capture ENSO climate variability? (3) A model-based study to evaluate yields as influenced by planting date under different ENSO phases. Based on previous related studies, the hypotheses are: (1) El Niño has a positive influence on corn yield while La Niña has negative impacts on corn yield. (2) Planting dates are significantly different under different ENSO phases. (3) The crop model –Hybrid-Maize model can catch the ENSO climate variability well, and the model driven by reanalysis data will have a much stronger ENSO feedback than onsite data.

4.1 Data Resources and locations

In this chapter, 18 counties across the U.S. Corn Belt (Fig. 1.1) were selected. Thirty years of corn yield data were collected from the National Agricultural Statistics Service (NASS) (http://www.nass.usda.gov/) annual survey. Because this study focuses on the impacts of climate variability, which needs decrease the influences from new techniques on corn yield, the original surveyed data were detrended to the 30 – year (1981-2010)

averaged yield. Planting dates of nine states in the U.S. Corn Belt were collected from the NASS report (1994 – 2010) and include NE, KS, MN, MO, WI, IL, IN, IA, and OH. The meteorological data of these 18 counties were collected from the National Climatic Data Center (NCDC) and the agro-meteorological database described in Chapter 3.

4.2 ENSO years classification

To classify ENSO years, in this study, the annual JMA-based "ENSO year" index (COAPS, 2010) was used. This index is based on the 5-month running mean of sea surface temperature (SST) anomalies, which are spatially averaged across the tropical Pacific (4°S-4°N, 150°W-90°W). When index data are equal or larger than 0.5°C for six consecutive months, the year starting in October through the following September is classified as an El Niño year. If index data are equal or less than -0.5°C for six consecutive months, of the year starting in October through the following September is classified as La Niña years, while others are classified as Neutral years (COAPS 2010; Japan Meteorological Agency 1991). Based on this criterion, 30 years (1981-2010) were grouped into three ENSO phase, eight years are classified as El Niño years, 17 years are classified as Neutral years, and five years are classified as La Niña years (Table 4.1)

4.3 The impacts of the El Niño–Southern Oscillation (ENSO) on corn yield in U.S. Corn Belt

Based on the classification of the ENSO years, the detrended surveyed data from 18 counties were grouped by the ENSO phases (Table 4.2). The ratios between yield data in El Niño years and yield data in Neutral years indicate that 13 counties obtained higher

yield in El Niño years. The ratios between yield data in La Niña years and yield data in Neutral years show that for 11counties, the yield data were decreased during La Niña years. The overall summary of these 18 counties (Table 4.3) also shows that El Niño events have a positive influence (ratio = 1.03) on corn yield while La Niña events have a negative impact (ratio = 0.96). When running an ANOVA test for the total yield data of 18 sites, the results report the negative impacts of the La Niña phase on corn yield is significant at the 99% level of confidence (p-Value = 0.0055) while the positive impacts of the El Niño phase on corn yield is not significant at the 95% level of confidence (p – Value = 0.06).The results of negative impacts from La Niña are similar in previous studies which were reviewed in Chapter 1. The reason for lower yield in La Niña years could be the summers tend to be warmer and drier in La Niña years than Neutral years in the Corn Belt. Additionally, cooler temperatures and higher rainfall rates in El Niño years might lead to yield improvement in some counties (Phillips et al. 1999). It is notable that the spatial pattern of ENSO impacts is not homogeneous, and more detailed regional studies are preferred in the future.

4.4 The impacts of the El Niño–Southern Oscillation (ENSO) on corn planting date in U.S. Corn Belt

Based on the NASS report, the active planting dates of the nine states are from April 16th to June 4th, and the most active planting dates are varied in different states. For most states, the most active planting dates are from April 30th to May 14th. For some states, such as Missouri (MO), the most active planting dates are in early April (Fig. 4.1). The corn weekly percentage planted data were grouped into three ENSO phases (Table 4.4).

In Fig. 4.2, it can be seen that in most states, the majority of the active planting dates in La Niña years are later than in Neutral years. The influence pattern in El Niño years is not clear. Figure 4.3 displays the summary of weekly percentage of corn planted for the nine states, and clearly indicates that the peak of percentage planted in La Niña years is one-week later than in Neutral years (p-Value = 0.0014). In El Niño years, the weekly percentage planted data are significantly different from the data in Neutral years (p-Value = 0.0026), while the data from El Niño years are more normally distributed. The peak under the El Niño phase is 19% while the peak under the La Niña phase is 23%, with a 22% peak under the Neutral phase.

4.5 Can crop model capture the climate variability?

The ability of crop models to capture climate variability is very important when conducting climate change impact studies. To investigate whether crop models can capture the impacts of El Niño / La Niña, in this research, the Hybrid-Maize model was run using two meteorological input datasets: onsite data from NCDC versus regional reanalysis data from the agro-meteorological database. The hypothesis being that even if the onsite data may have a limited ENSO signature, the reanalysis data will have a much stronger ENSO feedback embedded within. The Hybrid-Maize model simulated 30-years (1981-2010) of corn yield from 18 counties (Fig. 1.1) across the Corn Belt. The model's setting and running scheme are the same as described in Chapter 2 and Chapter 3.

4.5.1 Crop model running with onsite meteorological data

Table 4.5 listed the 30-year averaged simulated corn yield for 18 counties. The overall summary of these 18 counties (Table 4.6) also shows that El Niño events have a positive influence (ratio = 1.04) on corn yield while La Niña events have a slight negative impact (ratio = 0.99). When applying ANOVA tests for all the simulated yield data from the 18 sites, the results indicate that the negative impacts of the La Niña phase on corn yield is not significant at the 95% level of confidence (p-Value = 0.8). The positive impacts of the El Niño phase on corn yield is not significant at the 95% level of confidence (p – Value = 0.05). The averaged MAE (Table 4.7) of the simulated yield show that during El Niño years, MAE is larger than in Neutral years with a significant difference at the 95% level of confidence (p-Value = 0.04). The MAE difference between La Niña years and Neutral years is not significant at 95%.

These results indicate that when running the Hybrid-Maize model with onsite meteorological data, the model cannot capture the impacts of ENSO on corn yield (at 95% level). The MAE during El Niño years is significantly larger than Neutral years, and indicates that the simulations in El Niño years have more bias than simulations in the other two phases.

4.5.2 Crop model running with reanalysis meteorological data

Table 4.8 listed the 30-year averaged simulated corn yield from 18 counties. The overall summary of these 18 counties (Table 4.9) show both El Niño events have a slight influence (ratio = 0.99) on corn yield while La Niña events have a strong negative impact

(ratio = 0.90). When applying ANOVA tests for all the simulated yield data from the 18 sites, the results indicate that the negative impacts of the La Niña phase on corn yield is significant at the 99% level of confidence (p-Value < < 0.0001), and the impacts of the El Niño phase on corn yield is not significant at the 90% level of confidence (p –Value = 0.05). The difference between averaged MAE (Table 4.10) of the simulated yield and detrended observed yield under different ENSO phases is not significant: for El Niño years the P-value = 0.26, and for La Niña years the P-value = 0.59.

These results show that when running the Hybrid-Maize model with reanalysis meteorological data, the model can capture the impacts of ENSO on corn yield, especially the negative influence from La Niña (at 99% level). The MAE data under three ENSO phases were not significantly different.

Through running the Hybrid-Maize crop model with two meteorological datasets, it can be concluded that when the model is running with the reanalysis dataset, the impacts of ENSO on corn yield can be captured.

4.6 A model-based study ---Corn yields as influenced by planting date under different ENSO phases

It was discussed that corn yield and planting dates are influenced by ENSO phases, but there is no study to explore the impacts of planting date on corn yield under different ENSO phases. The planting date is one of the key management factors which highly relates to the corn yield. It is of great importance to know whether the planting date effects will vary by ENSO events. In this study, the Hybrid-Maize model was selected to

simulate 30-year (1981-2010) corn yield data from 18 counties (Fig. 1.1) for eight planting dates from April 16th to June 4th. Simulated yield data were grouped by ENSO phases.

Figure 4.4 lists the simulated corn yield for different planting dates and the data has been grouped into three ENSO phases. It can be seen that the impact of alternating planting dates under different ENSO phases is varied by county. For instance, in Dekalb, IL, Huntington, IN, and Japer, IN, under the La Niña phase, the simulated yields are increased when the planting dates change from May 8th to June 4th. While under El Niño and Neutral phases, the yields are decreased with a change to later planting dates. For these counties, based on the model, choosing a late planting date can mitigate the negative impacts of La Niña on corn yield. This result could explain the findings in Chapter 4.4 where the planting dates in La Niña years are later than Neutral years. However, it is also notable that in Olmstead, MN, Renville, MN, and Grand Forks, ND, the corn yield decreased with a change in planting date under all three ENSO phases. For these counties, the planting date decisions are not influenced by ENSO phases; earlier planting dates can bring higher yields. Table 4.10 listed the mean simulated yield, yield standard deviation, and yield range of eight planting dates. When changing the research scale from county to regional, no significant difference was found when the ANOVA test was applied to the total 18 counties data , which means when exploring the combined (ENSO + planting date) impacts on corn yield, the county scale is preferred.

4.7 Summary

The main findings in the chapter are: (1) The La Niña phase has significant negative impacts on corn yield and during La Niña years, the planting dates are significantly later than Neutral years. (2) The Hybrid-Maize model can capture the ENSO impacts when running with reanalysis meteorological data. (3) Based on this model study, in some counties, late planting can mitigate negative impacts from the La Niña phase. More detailed studies will be applied at different spatial levels in the future.

References

COAPS (Center for Ocean-Atmospheric Prediction Studies),(2010). ENSO Indexaccording to JMA SSTA. http://coaps.fsu.edu/jma.shtml. Florida State University.28 September 2010

Japan Meteorological Agency, M.D., (1991). Climate Charts of Sea Surface Temperatures of the Western North Pacific and the Global Ocean. Japan Meteorological Agency, Tokyo, p. 51

Phillips, J., Rajagopalan, B., Cane, M., & Rosenzweig, C. (1999). The role of ENSO in determining climate and maize yield variability in the US cornbelt. International Journal of Climatology, 19(8), 877-888.

Table 4.1 Annual JMA-based classifications of years (1981- 2010) into ENSO phases

El Niño	Neutral		La Niña
1982	1981	1994	1988
1986	1983	1995	1998
1987	1984	1996	1999
1991	1985	2000	2007
1997	1989	2001	2010
2002	1990	2003	
2006	1992	2004	
2009	1993	2005	
		2008	

Table 4.2 Observed average corn yield (1981-2013) of 18 counties grouped into ENSO phases

County	Yield(kg/ha)			Yield Ratio	
	El Niño	La Niña	Neutral	El Niño/Neutral	La Niña/Neutral
Johnson County, IA	8930	7296	8298	**1.08**	**0.88**
Winnebago County, IA	9359	9183	8919	**1.05**	1.03
DeKalb County, IL	9233	9448	9282	0.99	1.02
Douglass County, IL	9232	8284	9173	**1.01**	**0.90**
Huntington County, IN	8222	7495	8470	0.97	**0.88**
Jasper County, IN	8390	7896	8442	0.99	**0.94**
Shawnee County, KS	7300	7221	7246	**1.01**	1.00
Olmstead County, MN	9415	8726	8673	**1.09**	1.01
Renville County, MN	9495	8345	8733	**1.09**	**0.96**
Adair County, MO	7122	4804	6698	**1.06**	**0.72**
NewMadrid County, MO	9006	8427	9290	0.97	**0.91**
Platte County, NE	8520	8742	8379	**1.02**	1.04
Union County, OH	7907	7887	7939	**1.00**	**0.99**
Rock County, WI	8389	8002	8269	**1.01**	**0.97**
Sauk County,WI	8211	7658	7753	**1.06**	**0.99**
GrandForks County, ND	6076	6131	5158	**1.18**	1.19
Lucas County, OH	8518	8574	8930	0.95	**0.96**
Brookings, SD	6764	6565	6498	**1.04**	1.01

Table 4.3 Total observed average corn yield (1981-2013) of 18 counties grouped into ENSO phases

ENSO phase	Yield (kg/ha)	Yield Ratio (event years/neutral years)
El Niño	8338	**1.03**
La Niña	7816	**0.96**
Neutral	8119	

Table 4.4 Averaged (1994 – 2010) percentage corn planted of every week from April 16th to Jun 4th grouped into ENSO phases.

State	ENSO	Planting date							
		16-Apr	23-Apr	30-Apr	07-May	14-May	21-May	28-May	04-Jun
NE	El Niño	1	6	19	25	28	15	4	2
	La Niña	0	4	10	22	29	22	9	4
	Neural	1	5	16	27	20	19	7	3
KS	El Niño	11	12	17	19	17	14	9	2
	La Niña	4	10	20	15	22	16	6	5
	Neural	12	14	20	20	14	11	4	5
MN	El Niño	0	4	26	28	21	13	6	2
	La Niña	0	3	28	39	19	6	3	2
	Neural	1	3	17	28	16	19	9	4
MO	El Niño	24	14	12	11	11	10	9	5
	La Niña	9	17	17	8	10	14	9	11
	Neural	25	16	15	10	6	8	6	3
WI	El Niño	0	2	8	15	20	19	16	9
	La Niña	0	2	9	18	27	22	13	8
	Neural	0	0	7	15	32	18	13	5
IL	El Niño	3	13	19	15	11	5	18	8
	La Niña	1	13	17	19	22	13	11	4
	Neural	7	15	20	21	11	5	8	0
IN	El Niño	1	3	13	15	10	8	21	13
	La Niña	1	6	13	14	24	23	12	4
	Neural	3	8	17	23	15	7	9	1
IA	El Niño	1	10	26	25	23	10	4	1
	La Niña	0	7	15	25	31	13	5	3
	Neural	1	7	24	28	16	13	7	1
OH	El Niño	2	3	18	21	10	8	19	8
	La Niña	2	3	15	15	28	16	14	6
	Neural	2	7	12	28	18	8	9	7

Table 4.5 Simulated average corn yield (1981-2013) of 18 counties driven by onsite meteorological data grouped into ENSO phases

County	Yield(kg/ha)			Yield Ratio	
	El Niño	La Niña	Neutral	El Niño/Neutral	La Niña/Neutral
Johnson, IA	7831	7291	7911	0.99	**0.92**
Winnebago, IA	9299	9270	9202	**1.01**	1.01
DeKalb , IL	9209	8270	8966	**1.03**	**0.92**
Douglass, IL	7789	7335	8234	0.95	**0.89**
Huntington, IN	8950	7821	8772	**1.02**	**0.89**
Jasper, IN	8888	7903	8828	**1.01**	**0.90**
Shawnee, KS	6961	6388	7303	0.95	**0.87**
Olmstead, MN	8845	9363	7675	**1.15**	1.22
Renville, MN	9564	9782	8551	**1.12**	1.14
Adair, MO	8292	7183	8063	**1.03**	**0.89**
NewMadrid, MO	6638	6370	6986	0.95	**0.91**
Platte, NE	7871	7532	7923	0.99	**0.95**
Union, OH	8804	7829	8588	**1.03**	**0.91**
Rock, WI	9151	8539	9296	0.98	**0.92**
Sauk ,WI	8619	9290	7534	**1.14**	1.23
GrandForks, ND	7679	8018	6989	**1.10**	1.15
Lucas, OH	9202	8500	9003	**1.02**	**0.94**
Brookings, SD	6743	7873	5391	**1.25**	1.46

Table 4.6 Total simulated average corn yield (1981-2013) of 18 counties driven by onsite meteorological data grouped into ENSO phases

ENSO phase	Yield (kg/ha)	Yield Ratio (event years/neutral years)
El Niño	8352	1.04
La Niña	8031	0.99
Neutral	8068	

Table 4.7 Mean absolute error (MAE) between simulated corn yields driven by onsite meteorological data with detrended observed data (1981-2010, 18 counties)

	El Niño	La Niña	Neutral
MAE (kg/ha)	1370.94	1185.83	1253.55
Std.Dev of bias	985.17	948.69	1055.55

Table 4.8 Simulated average corn yield (1981-2013) of 18 counties driven by reanalysis meteorological data grouped into ENSO phases

County	Yield(kg/ha)			Yield Ratio	
	El Niño	La Niña	Neutral	El Niño/Neutral	La Niña/Neutral
Johnson, IA	7822	7247	7962	0.98	0.91
Winnebago, IA	9049	8156	9272	0.98	0.88
DeKalb , IL	8905	7961	8885	1.00	0.90
Douglass, IL	7945	7426	8255	0.96	0.90
Huntington, IN	8694	7688	8773	0.99	0.88
Jasper, IN	8615	7788	8765	0.98	0.89
Shawnee, KS	6193	6294	6780	0.91	0.93
Olmstead, MN	9594	8930	9406	1.02	0.95
Renville, MN	9004	7406	8979	1.00	0.82
Adair, MO	7385	6838	7752	0.95	0.88
NewMadrid, MO	6669	6258	6810	0.98	0.92
Platte, NE	6910	6899	7704	0.90	0.90
Union, OH	8712	7736	8907	0.98	0.87
Rock, WI	8915	7948	8856	1.01	0.90
Sauk ,WI	9217	8887	9068	1.02	0.98
GrandForks, ND	8597	7847	8226	1.05	0.95
Lucas, OH	8861	8014	8864	1.00	0.90
Brookings, SD	8335	7180	8235	1.01	0.87

Table 4.9 Total simulated average corn yield (1981-2013) of 18 counties driven by reanalysis meteorological data grouped into ENSO phases

ENSO phase	Yield (kg/ha)	Yield Ratio (event years/neutral years)
El Niño	8301	**0.99**
La Niña	7583	**0.90**
Neutral	8417	

Table 4.10 Mean absolute error (MAE) between simulated corn yields driven by reanalysis meteorological data with detrended observed data (1981-2010, 18 counties)

	El Niño	La Niña	Neutral
MAE (kg/ha)	1370.94	1185.83	1253.55
Std.Dev of bias	985.17	948.69	1055.55

Table 4.11 Mean simulated yield, yield standard deviation and yield range of 8 planting date (18 counties, 1981-2010)

County	Mean yield(Mg/ha)			Yield std			Yield range		
	El Niño	Neutral	La Niña	El Niño	Neutral	La Niña	El Niño	Neutral	La Niña
Johnson, IA	8.0353	7.93924	7.2662	0.2927	0.07255	0.1314	0.8085	0.208588	0.4044
Winnebago, IA	9.0082	8.6621	9.3153	0.4715	0.7048	0.2205	1.42875	1.952824	0.6564
DeKalb, IL	9.11756	8.9270	8.5457	0.09096	0.1731	0.3370	0.27375	0.5982	0.9264
Douglass, IL	8.1245	8.2701	7.35120	0.3735	0.1074	0.08887	0.9975	0.2958	0.24960
Huntington, IN	9.0871	8.8742	8.1209	0.2045	0.1081	0.3924	0.4972	0.2876	1.1316
Jasper, IN	8.9228	8.9922	8.2934	0.1955	0.1788	0.3176	0.5392	0.4761	0.9903
Shawnee, KS	7.1782	7.33350	6.3945	0.2308	0.06106	0.1131	0.6097	0.19129	0.3312
Olmstead, MN	7.945	6.8821	8.8999	1.091	0.9235	0.5257	3.049	2.5708	1.5192
Renville, MN	9.1042	8.0060	9.3483	0.7483	0.7795	0.5234	2.1563	2.1681	1.5348
Adair, MO	8.5151	8.2747	7.3181	0.3152	0.2541	0.1530	0.7762	0.6776	0.4848
NewMadrid,MO	6.7357	6.9339	6.2369	0.1218	0.1101	0.1608	0.3833	0.3187	0.4080
Platte, NE	8.1360	8.0253	7.58595	0.2933	0.1248	0.09460	0.7163	0.3191	0.30000
Union, OH	8.9487	8.7457	8.0844	0.2154	0.2055	0.3189	0.5587	0.5548	0.8568
Rock, WI	8.9160	8.9179	8.9439	0.3716	0.4316	0.4804	1.0133	1.2494	1.3116
Sauk,WI	7.876	6.9469	9.0003	1.101	0.9112	0.4842	3.123	2.5761	1.4184
GrandForks, ND	7.162714	6.07916	7.502229	0.8739	0.9425	0.8759	2.4255	2.550706	2.4444
Lucas, OH	9.0688	8.9802	8.7825	0.2167	0.1785	0.4181	0.5445	0.5619	1.1508
Brookings, SD	7.766	6.3596	9.4810	1.104	0.9265	0.9119	2.996	2.5422	2.4852

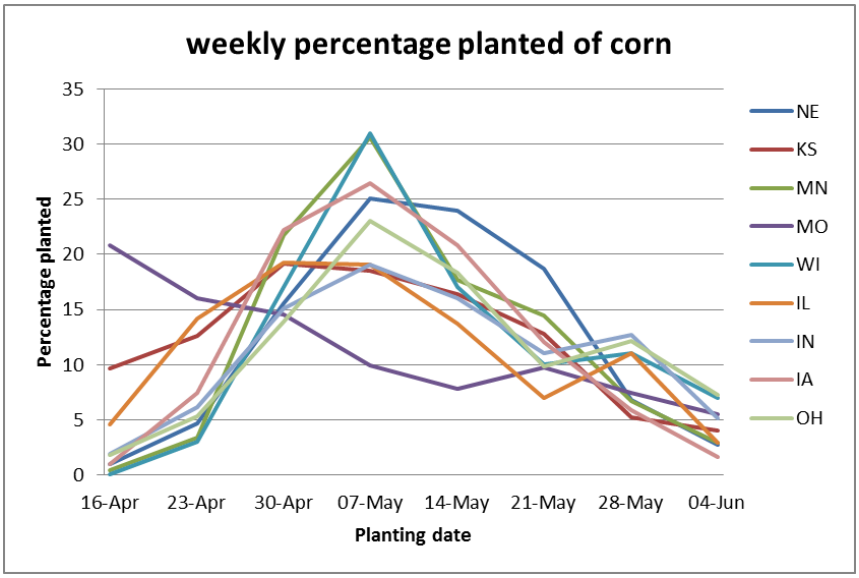

Figure 4.1 Averaged (1994 – 2010) weekly corn percentage planted for 9 states in U.S Corn Belt.

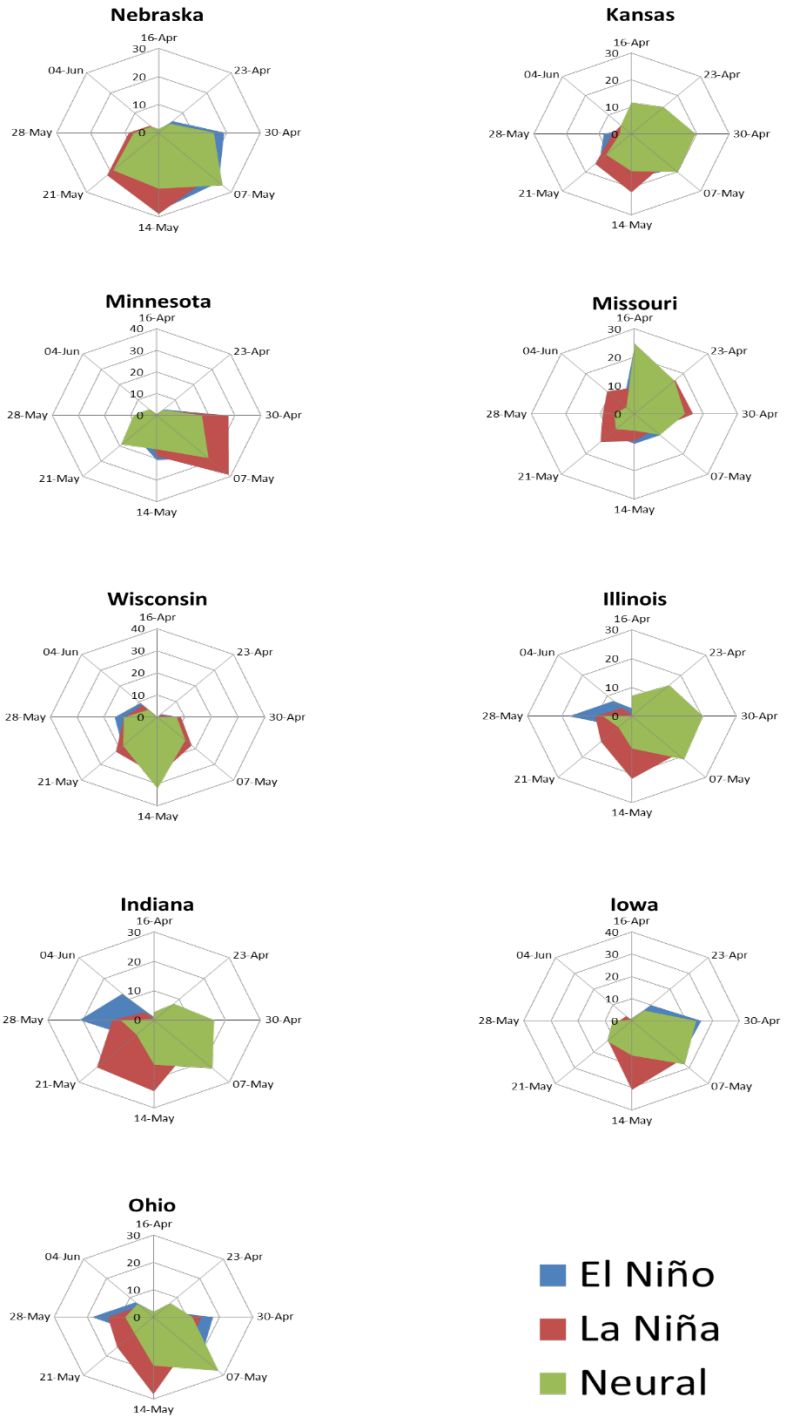

Figure 4.2 Averaged (1994 – 2010) weekly corn percentage planted of 9 states under different ENSO phase.

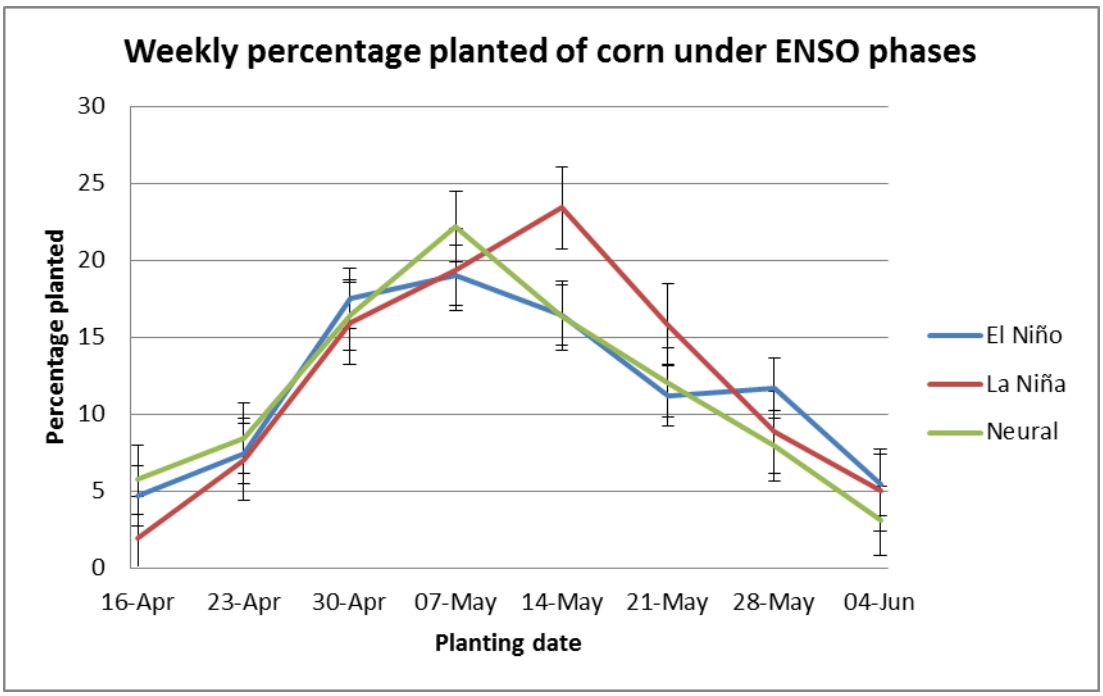

Figure 4.3 9-sates averaged (1994 – 2010) weekly corn percentage planted under different ENSO phase (significantly different at 99% level of confidence).

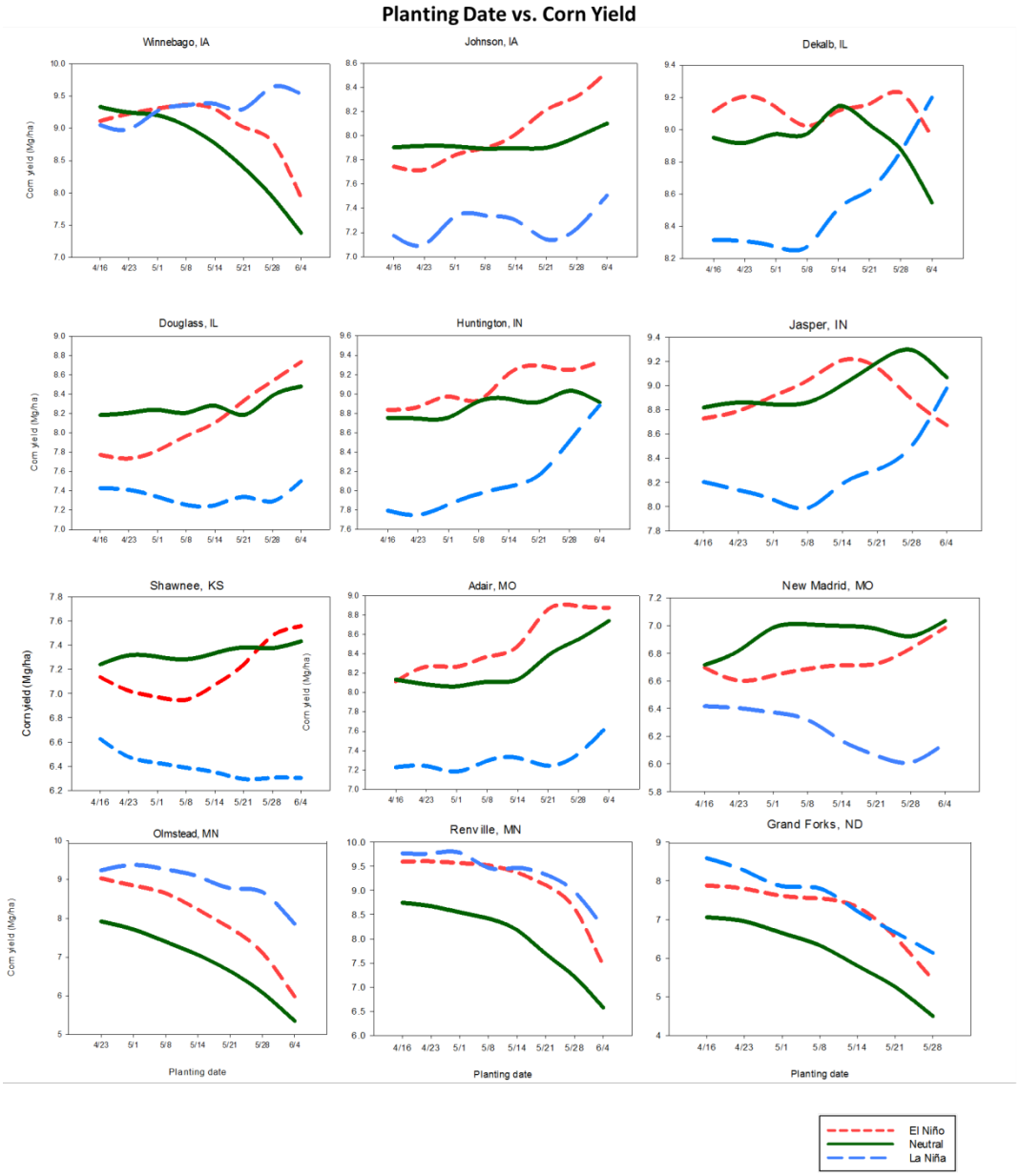

Figure 4.4 Corn yield simulations with different planting dates grouped into three ENSO Phases

CHAPTER 5. RUNNING CROP MODEL WITH FUTURE CLIMATE PROJECTION

The most important application of a crop model is projecting the future yield. Based on the reliability of crop models evaluated in the previous chapters, here a preliminary test of running the crop model (the Hybrid-Maize) with regional climate models (RCMs) will be discussed.

5.1 Data source and research location

In this study, the research site is Bondville, IL (40.00°N, 88.29°W). Climate model-simulated meteorological data were collected from the North American Regional Climate Change Assessment Program (NARCCAP, Mearns et al. 2009). Measured meteorological data were collected from AmeriFlux. NARCCAP data-driven simulated yields were compared with onsite meteorological data-driven simulated yields. The test period is 1981-2003.

5.2 NARCCAP meteorological data validations

To apply the climate model-simulated meteorological data with the crop model, it is necessary to validate the data reliability. In this research, the validations were conducted for daily maximum temperature (˚C), daily minimum temperature (˚C), and daily accumulated solar radiation (MJ/m^2). The validation of daily maximum temperature (Fig. 5.1) shows that the NARCCAP climate model-simulated maximum temperature slightly

underestimated lower values (< 10°C) and overestimated higher values (generally > 10 °C). Validation of daily minimum temperature (Fig. 5.2) shows a similar pattern as the maximum temperature where higher values (generally >10 °C) were overestimated and lower values (<10°C) are underestimated. Validation of daily solar radiation indicates that the climate model overestimated solar radiation values (Fig. 5.2). The overall agreement between NARCCAP-simulated values and Ameriflux observed data are acceptable for application in crop simulations.

5.3 Running crop model with NARCCAP meteorological model

The Hybrid-Maize model was used to simulate corn yields from 1981 to 2003 with two meteorological data scenarios: (1) NARCCAP data, and (2) Onsite data. The purpose of using these two data sources is to evaluate the performance of running the crop model with NARCCAP climate model-simulated data. If the bias between the results of the scenarios is not significant or can be rescaled, it means that the crop model can be driven by the NARCCAP future projected climate data. The model settings of these two scenarios were unified. Simulated corn yield driven by NARCCAP data shows a similar trend with simulated corn yield driven by onsite data (Fig. 5.4). Mean simulated yield bias between the two crop model-driven schemes are 17.5 bu/acre (1.1 Mg/ha). Applying regression analysis to rescale the NARCCAP-driven data will decrease the bias to 7.8 bu/acre (0.5 Mg/ha) (Fig. 5.5).

Based on meteorological data validations and simulated yield data validations in Bondville, IL, NARCAAP climate model-simulated meteorological data shows good potential when applied with corn yield simulations. However, since this preliminary

study only evaluated one site, more tests are needed before applying future projections from NARCCAP at a regional scale.

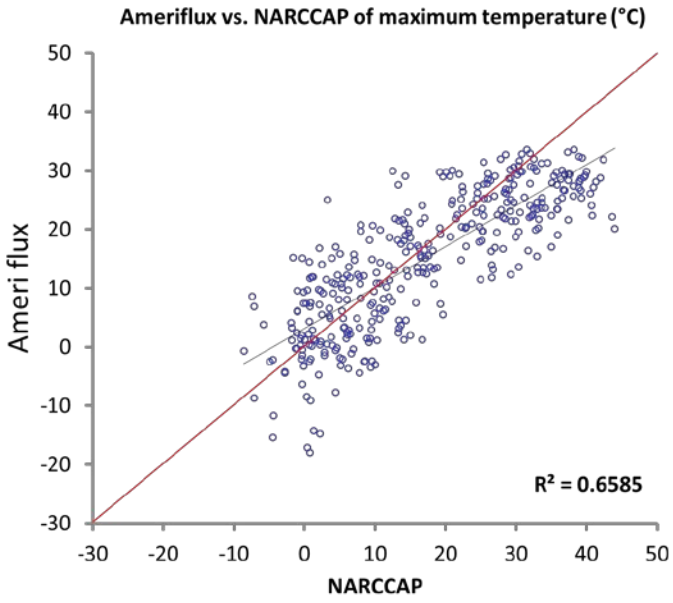

Figure 5.1 Validation of daily maximum temperature from NARCCAP (Bondville, IL)

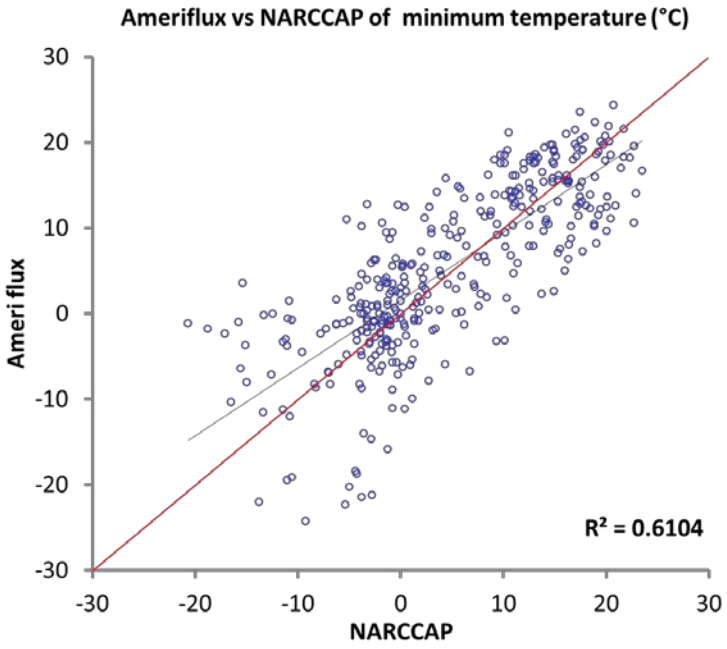

Figure 5.2 Validation of daily minimum temperature from NARCCAP (Bondville, IL)

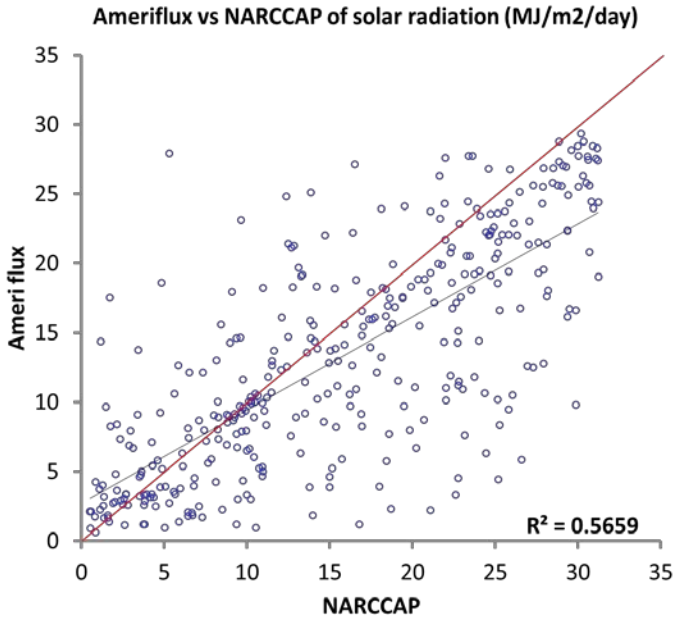

Figure 5.3 Validation of daily accumulated solar radiation from NARCCAP (Bondville, IL)

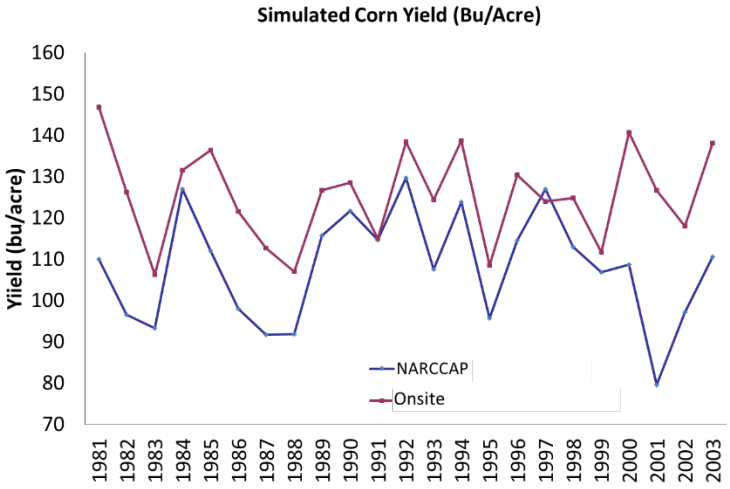

Figure 5.4 Validation of daily accumulated solar radiation from NARCCAP (Bondville, IL)

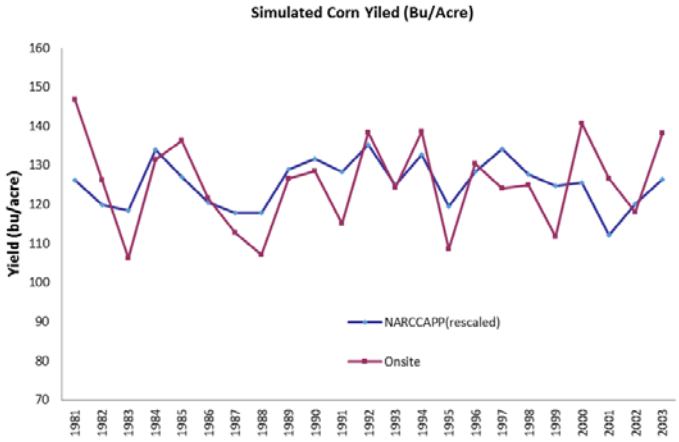

Figure 5.5 Rescaled simulated corn yield with NARCCAP meteorological input and Onsite meteorological model input.

References

Mearns, L. O., W. J. Gutowski, R. Jones, L.-Y. Leung, S. McGinnis, A. M. B. Nunes, and Y. Qian, (2009). A regional climate change assessment program for North America. *EOS*, Vol. 90, No. 36, pp. 311-312.

CONCLUSION

In this research, a high-resolution reanalysis agro-meteorological database across the U.S. Corn Belt has been compiled with raw data from the Land Data Assimilation System (LDAS) which includes daily maximum/minimum temperature, precipitation, solar radiation, etc. Validations of meteorological data show strong agreement between this reanalysis database and observed data, which gives confidence for wide use of this agro-meteorological database with agro-related applications at different spatial scales. Spatially-continuous daily solar radiation data are available in this database which provides a solution to the problem of sparse historical solar radiation in crop model related research.

A gridded crop model running system has been developed based on this agro-meteorological database and the Hybrid-Maize crop model. After validation, this system shows good potential for estimating regional corn yield at a gridded scale under different scenarios (e.g., different planting dates).

La Niña in ENSO phases show significant negative impacts on corn yield and are a factor in the relatively late planting dates in the Corn Belt. The Hybrid-Maize crop model can capture the impacts of ENSO on corn yield when the model is driven by reanalysis data from the agro-meteorological database.

The key limitations of this research include: (1) lack of soil information and standardized management variables when running the crop model across the Corn Belt. (2) Only one classification method and one ENSO index were used in this research. (3) Eighteen county-scale sites are relative small sample size for regional studies.

Future study based on this research will focus on: (1) evaluating hydrological parameters in the agro-meteorological database. (2) The addition of dynamic soil information and field management information to the gridded crop model. (3) Classifying ENSO years with ensemble methods and the ENSO index. (4) Expansion of the validation sample size.