Spatial Quantification of the Gap between Farm Field and University Trial Maize Yields in the United States

A Thesis
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

Christine Forland

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE IN BIOPRODUCTS/BIOSYSTEMS SCIENCE
ENGINEERING AND MANAGEMENT

Dr. Jason Hill

August 2015

Acknowledgements

I thank Dr. Jason Hill for his guidance as my adviser and Dr. J. Coulter and B. Befort for their assistance. This work was supported in-part by a grant from CenUSA Bioenergy, Agriculture and Food Research Initiative Competitive Grant No. 2011-68005-30411 from the USDA National Institute of Food and Agriculture.

I also thank the following university variety trial personnel who provided correspondence and, in many cases, yield and irrigation data: William Batchelor and Edzard van Santen, Auburn University; Richard Bond and Donald Dombek, University of Arkansas; Carol Frate and Kent Brittan, University of California; Sally Sauer, Colorado State; Richard Taylor , James Hawk, and Tecle Weldekidan, University of Delaware; Jose Santos and Jerry Wasdin, University of Florida; Anton Coy and Don Day, University of Georgia; James L. Brewbaker, University of Hawaii at Manoa; Darin Joos, University of Illinois at Urbana-Champaign; Phil DeVillez, Purdue University; Jim Rouse, Iowa State University; Jane Lingenfelser, Kansas State University; Todd Pfeiffer, Chad Lee, Bill Bruening, and Jessica Cole, University of Kentucky; John S. Russin, John Kruse, and Rick Mascagni, Louisiana State University; Rick Kersbergen, University of Maine; Robert Kratochvil, University of Maryland; Masoud Hashemi, University of Massachusetts; Bill Widdicombe, Kurt Thelen, and Lori Williams, Michigan State University; Jeffrey Coulter, University of Minnesota; Brad Burgess, Mississippi State University; Howard Mason, University of Missouri; Kent McVay and Ken Kephart, Montana State University; Teshome Regassa, University of Nebraska-Lincoln; Michael O’Neill and Mark Marsalis, New Mexico State University; Margaret E. Smith and Judy Singer, Cornell University; Daryl Bowman, North Carolina State University; Joel Ransom and Marcelo Carena, North Dakota State University; Allen Bruce Geyer, Ohio State University; Rick Kochenower, Oklahoma State University; Greg Roth and John Shaffer, Pennsylvania State University; Dewey Barefield, Clemson University; Robert Hall, South Dakota State University; Fred Allen, University of Tennessee; Dennis Pietsch and Katrina Horn, Texas A&M University; Earl Creech and Mike Whitesides, Utah State University; Heather Darby, University of Vermont; Wade Thomason and Elizabeth Hokanson, Virginia Polytechnic Institute and State University; Barton Baker, West Virginia University; Joe Lauer, University of Wisconsin; and Stephen Herbert, Joanne Newcomb, and Jerry Nachtman, University of Wyoming.

Dedication

This thesis is dedicated to my parents in honor of their constant and unconditional love and support.

Abstract

Greater crop production will be required to support both an increase in biofuel use and a forecasted doubling of global food demand by 2050. An improved understanding of yield potential and realistic estimates of the magnitude and spatial variability of the gap between actual yield and yield potentials are critical to achieving maximum crop production. This study examines near-term yield potentials and gaps of maize (*Zea mays* L.) yield data over the years of 2006 to 2011 from two sources: university crop variety trials and the United States Department of Agriculture yield surveys. Yield potentials are analyzed across 32 states through a compiled database of 129,499 trial maize hybrid entries. From the database, 1,102 direct, irrigation-specific, year-to-year, county-to-county yield comparisons are made across 27 states. These 32 and 27 states comprise nearly all United States maize production—99% and 97%, respectively. Trial yield is calculated as the 90th percentile of hybrid yields in a given county in a given year, and farm yield is the USDA-reported county-level yield in that same trial-performing county in that same year.

Analysis of the median yield gap values in each state shows a yield gap of 13% to 53% in rainfed maize and a yield gap of 16% to 39% in irrigated maize. The magnitude of these differences between farming and trial yields indicates that maize yields in the United States, particularly rainfed, have considerable room for improvement. Additionally, the 40% range of median rainfed yield gap values and the 23% range of median irrigated yield gap values suggest that the yield gap varies greatly between states. The results of

this study are expected to support the production of more accurate biofuel crop projections and identify where yields might be increased, thereby avoiding further land conversion to cropland while reaching the goal of increasing biofuel production and sustaining ample food production.

Abbreviations: USDA-NASS, United States Department of Agriculture National Agricultural Statistics Service; YF, farm yield; YFi, irrigated farm yield; YFr, rainfed farm yield; YG, yield gap; YGi, irrigated yield gap; YGr, rainfed yield gap; YT, trial yield; YTi, irrigated trial yield; YTr, rainfed trial yield

Table of Contents

List of Tables

List of Figures

Introduction

Crop production will need to increase substantially if it is to meet both the forecasted global food demand increase of 100–110% from 2005 levels by 2050 (Tilman et al. 2011) and the increasing demand for biofuel crops (e.g., maize, soybean, and rapeseed) driven by the biofuel production mandates and the targets of current renewable energy policies (Johnston et al. 2009). To accomplish this feat, some combination of cropland area and crop yield, the two primary factors contributing to crop production, must also increase.

Land is a finite resource—a portion of which is used for crop production. The amount of cropland can only increase by converting lands currently dedicated to other purposes. Conversion of pastureland and natural landscapes to cropland may entail numerous negative environmental impacts, such as the release of stored carbon (Fargione et al. 2008), degraded water quality, lowered water tables and reduced stream flow due to irrigation (Scanlon et al. 2007), loss of habitat, displaced wildlife and plant communities, and a substantial reduction in biodiversity (Cassman et al. 2003). If expansion of cropland and the associated negative environmental impacts are to be minimized, crop yield—the second primary factor of crop production—on existing cropland will need to be increased.

Increasing crop production can also be accomplished by improving crop yields such that they are closer to their yield potential. Yield potential can be estimated in a number of different ways including crop model simulation yields, certified yield contest-winning yields, surveys of historical highest-recorded agricultural research station yields (Duvick and Cassman 1999), and maximum farmer yields (Lobell et al. 2009). Variations of these yield potential estimates have been used in numerous studies of the gap between yield potential and actual yield (Abeledo et al. 2008; Aggarwal et al. 2008; Tittonell et al. 2008; Gerber et al. 2010; Grassini et al. 2011; Hochman et al. 2012; Laborte et al. 2012; Hall et al. 2013; Meng et al. 2013; Tack et al. 2015; van Ittersum et al. 2013; van Wart et al. 2013).

One commonly used method of estimating yield potential is to estimate the yield attained under absolute ideal conditions—ideal soil, ideal weather, ideal inputs, and ideal management—with every last bushel harvested. Such absolute ideal growing conditions may be impossible to attain (Evans 1993) and are unlikely to be achievable on a wide scale in the near term.

To estimate a more near-term, achievable value, yield potential can alternatively be estimated as the yield attained under realistic ideal growing conditions using best management practices. Best management practices are not optimized to achieve maximum yield at any cost; they are optimized to balance the cost of inputs with the potential for outputs. Farmers commonly strive for this type of yield potential given production is bound by limits in financial resources.

To help farmers compare performance among current crop varieties and understand the yields they might achieve through best management practices under realistic ideal conditions, universities in the United States plant trial plots of current crop varieties. Crop variety trials are conducted by the land-grant university (or, occasionally, universities) in each state. Land-grant universities are designated to receive federal support as a part of the land-grant system, which includes the Agricultural Experiment Station Program and the associated Cooperative Extension Service Program, created to publicize the experiment station data (What Is a Land-Grant College? 2009). Seed companies submit entries into the annual trials and typically pay a small fee for trial expenses. The results of these variety trials are used by seed companies as a source of information for marketing and decision-making and by farmers as a direct, objective, and reliable guide to which varieties to plant in the upcoming season to achieve the maximum yield and subsequent maximum profit.

Trials are most often performed in multiple regions within a state in an attempt to test the crop varieties across a range of different environments and representative soils. They are professionally conducted by university staff with state-of-the-art best management practices, tools, and techniques. Because of the use of best management practices and the realistic ideal growing conditions, trial yield (YT) is therefore an estimate of the yield potential more likely achievable on a wide scale in the near term.

In this study, YT was chosen as an estimate of yield potential in the interest of examining the gap between actual yield and a more realistic yield potential. It is different from modeled estimates of yield potential, calculated under absolute ideal growing conditions and management practices, in that it demonstrates the yields that can realistically be achieved on well-managed plots. It is the best a farmer could expect to do in a given area, employing best management practices and current genetics.

Actual yield is defined in this study as the farm yields (YF) reported by the United States Department of Agriculture National Agricultural Statistics Service (USDA-NASS) Agricultural Surveys. Surveys collecting row crop and hay production data are conducted annually each December. Data are collected from farm operators via mail, telephone, and interviews over the course of two weeks. Sample size targets range from 65,000 to 81,000, depending upon the crop being surveyed. Yield data are reported at the county level on an overall, rainfed, and/or irrigated basis. These USDA-NASS Agricultural Survey data are the measure of actual yield in this study as it represents the yields actually achieved by farmers.

Here we define the difference between YT and YF in a defined geographical range and time frame is called the yield gap (YG). Quantifying YG provides the information needed to describe how well farmers are doing and how much room is available for improvement. Analysis of YG data also supports the creation of accurate crop production forecast model parameters, improving the estimates of future food and biofuel production capabilities.

Spatial quantification of YG illustrates the areas that have the greatest potential for yield growth. This knowledge draws attention to which areas resources might be directed to potentially achieve the largest gains in yield and, thus, the largest gains in crop production. Prioritizing areas on which to focus agricultural resources and attention is essential if growing demand for crops for food and biofuel are to be met—and particularly so if it is to be done with minimal further conversion of land to cropland.

This study quantifies maize YG across the United States on a county level over the years 2006–2011 and characterizes the spatial distribution and variability of YG both in rainfed maize (YGr) and irrigated maize (YGi). Maize was the crop selected as the focus of this study as it is one of the leading crops used in both food and biofuel production. In 2011, The United States harvested just under 34 million hectares of maize, nearly one quarter of the total harvested hectares in the country, for a cash receipt of $63.9 billion (USDA-NASS n.d.), and the pressure for the production of more food and more biofuels is only increasing the demand for maize.

Materials and Methods

All yields are analyzed on a county level. This is the smallest scale that USDA-NASS reports the Agricultural Survey data for YF. Trial locations for calculation of YT are typically provided as the city or the county in which the trial was conducted, though sometimes latitude and longitude are provided.

Trials are only performed in a select number of counties in each state each year. USDA-NASS Agricultural Surveys report data for many more counties each year than trials but still do not report data for every county in every state each year. In addition, trials are either conducted and reported as rainfed or irrigated. USDA-NASS reports overall (non-irrigation-specific), rainfed, and irrigated yield values; a given county may have only an overall yield value reported, only a rainfed value reported, only an irrigated yield reported, or any combination of the three values reported. Non-irrigation-specific overall yield values were not used in this study so as to only make direct rainfed YT-to-rainfed YF comparisons and direct irrigated YT-to-irrigated YF comparisons.

This study analyzes yields over 2006–2011. For each year of data, only counties that had both a rainfed YT and a rainfed YF reported or both an irrigated YT and an irrigated YF were included. Therefore, each YG measured is on the same binary irrigation basis, in the same county, in the same year.

Yield Gap. In this study, YG is expressed as the percentage difference between the 90th percentile of the yields achieved by all hybrids entries tested in a land-grant university crop variety trial within a given county (YT) and the average yield achieved by farmers in the same county as reported by USDA-NASS Agricultural Surveys (YF).

The 90th percentile of entry yields in a given trial was chosen over the average of entry yields in a given trial as an estimate of YT because the average would not provide an estimate of the greatest potential of maize yield. The 90th or 95th percentile of a range of yields has been used in previous yield studies to estimate yield potential (Licker et al. 2010; Foley et al. 2011; Mueller et al. 2012). The 90th percentile of trial yields was chosen over the 95th percentile or the maximum of trial yields as an estimate of yield potential to produce a more conservative estimate.

YT data from counties within 32 states over the years 2006-2011 were compiled. Yields from 27 of the 32 states were able to be used to analyze YG due to the availability of the necessary irrigation- and county-year-matching YF data. The 32 states comprise approximately 99% of maize production and 99% of maize hectares harvested within the United States. The 27 states still comprise approximately 97% of maize production and 97% of maize hectares harvested within the United States, as found by calculating the percentage of total from state-level USDA-NASS data for each of the 50 states and summing over the 32 or 27 states, respectively.

The gap between YF and YT was calculated on a percentage basis:

$$YG = 100 - \frac{YF}{YT} \times 100$$

Hence, a positive YG percentage indicates that YF was greater than YT. Depending on the data available within a county, a YGr, a YGi, or both a YGr and a YGi were calculated for each county in this study.

Trial Data. As previously discussed, trials are conducted using best management practices under realistic ideal conditions. The best management practices used in each state's trials are described in Table 1. Irrigation in each state is described in more detail later.

State	Plot Size	Fertilizer	Herbicide	Fungicide	Insecticide	Pesticide	Tillage	Planting Rate	Planting Method	Harvesting Method
Alabama	Two 30- to 36-in. rows, 20 to 30 ft. long	Yes, applied according to soil test reco-mmendations	Yes	--	--	--	Yes/No, differs by location	28,000 seeds/acre	--	--
Arkansas	Two rows wide, 20 to 25 ft. long	Yes, varying nutrients and amounts applied	Yes	--	Yes	--	--	Based on reco-mmen-dation of seed company	--	Plot combine
California	Four to six 30-in. rows, 476 to 1,250 ft. long	Yes, varying nutrients and amounts applied	Yes	--	Yes/No, differs by location	--	--	35,000 seeds/acre	Air planters	--
Colorado	Five ft. wide, 30 ft. long	Yes, varying nutrients and amounts applied	Yes	--	Yes	--	Yes/No, differs by location	--	--	--
Delaware	Four rows, 17.4 ft. long, center two rows harvested	Yes, varying nutrients and amounts applied	Yes	--	Yes	--	Yes/No, differs by location	--	Monosem air planter	--
Georgia	30-in. rows	Yes, varying nutrients and amounts applied	Yes	--	--	Yes	Yes	24,500 to 35,000 seeds/ acre, differs by location	--	--

State	Plot Size	Fertilizer	Herbicide	Fungicide	Insecticide	Pesticide	Tillage	Planting Rate	Planting Method	Harvesting Method
Illinois	Four rows wide, 23 ft. long, center two rows harvested	Yes, applied as needed	Yes	Yes/No, differs by location	Yes, applied in furrow at planting	--	Yes	--	Modern four-row planter modified for small plot work	Custom-built, self-propelled, corn plot combine
Indiana	Four 30-in. rows, 21.5 ft. long, center two rows harvested	Yes, applied based on location management program and desired plant population	Yes/No, differs by location	Yes	--	--	Yes	32,000 seeds/acre	Air plot planter	Self-propelled harvester without gleaning
Iowa	Four 30-in. rows, 20 ft. long (planted row length 17.4 ft. long), center two rows harvested	--	--	--	Yes	--	--	34,500 seeds/acre	--	Corn combine without gleaning
Kansas*	Two rows, 20 to 30 ft. long	Yes, varying nutrients and amounts applied	--	--	Yes/No, differs by hybrid	--	--	--	--	--
Kentucky	Two 30-in. rows, 20 ft. long	Yes, applied according to soil test reco-mmendations	Yes	--	Yes	--	No	28,000 plants/acre	John Deere Maxi-Emerge vacuum, two-row, no-till planter modified for small plot work	Massey Ferguson MF 8-XP two-row corn combine
Louisiana	38- to 40-in. rows	Yes, varying nutrients and amounts applied	--	--	--	Yes	Yes	32,000 seeds/acre	--	--
Maryland	Four 30-in. rows, 32-ft. harvest length, center two rows harvested	Yes, varying nutrients and amounts applied	Yes	--	Yes/No, differs by location	--	Yes/No, differs by location	29,500 seeds/acre	Modified, four-row John Deere 1750 planter equipped with coulters and trash-wheels for no-till planting	--
Michigan*	Four 30-in. rows, 22 ft. long, center two rows harvested	Yes, varying nutrients and amounts applied	--	--	--	--	--	--	--	Mech-anically
Minnesota	--	--	--	--	--	--	--	33,000 seeds/acre	--	--
Mississippi	Two 30-in. rows, 15 ft. long	Yes, applied according to soil test reco-mmendations	Yes, applied with strict adherence to label instructions	--	Yes	--	Yes/No, differs by location	Rates suggested by seed companies (28,000 to 36,000 seeds/ acre, differs by location)	Cone planter	--

State	Plot Size	Fertilizer	Herbicide	Fungicide	Insecticide	Pesticide	Tillage	Planting Rate	Planting Method	Harvesting Method
Missouri	Four 30-in. rows, 27 ft. long, center two rows harvested	Yes, applied at the discretion of the research manager	Yes	--	Yes, in-furrow	--	Yes/No, differs by location	30,000 seeds/acre for non-irrigated tests and 36,000 seeds/acre for irrigated tests	--	--
Montana	--	--	--	--	--	--	--	--	--	--
Nebraska	Two rows wide, 15 to 35 ft. long	Yes/No, differs by location	Yes/No, differs by location	One location only	--	--	Yes/No, differs by location	17,550 to 31,660 plants/acre	--	--
New Mexico	Two or four 30- to 40-in. rows, 10 or 20 ft. long (if four rows, center two rows harvested)	Yes, varying nutrients and amounts applied	Yes	--	Yes/No, differs by location	--	--	27,000 to 90,000 seeds/acre	--	--
New York	--	--	--	--	--	--	--	--	Machine	Combine
North Carolina	Two 30-, 36-, or 38-in. rows, 22 ft. long, center two rows harvested, 6-ft. alley width	Yes, varying nutrients and amounts applied	--	--	Yes	--	--	--	--	--
North Dakota	--	--	--	--	--	--	--	--	--	--
Ohio	Four 30-in. rows, approximately 25 ft. long	Yes, applied according to recommended cultural practices for obtaining optimum grain yields	Yes, applied according to reco-mmended cultural practices for obtaining optimum grain yields	--	Yes, applied according to reco-mmended cultural practices for obtaining optimum grain yields	--	No, or minimal	--	Commercial type planter adapted for plot planting	--
Oklahoma	Two rows, 25 ft. long, trimmed to 20 ft. long prior to harvest	Yes, varying nutrients and amounts applied	Yes	--	--	--	Yes	28,000 to 32,000 plants/acre	--	--
Penn-sylvania	Four rows, equal to 1/250 acre, center two rows harvested	--	--	--	Yes, at locations where previous crop was corn	--	--	30,000 plants/acre	Vacuum precision planter	Self-propelled combine
South Dakota	Four 30-in. rows, 20 ft. long, center two rows harvested	Yes	Yes, applied at recommen ded label rates	--	--	--	Yes/No, differs by location	29,621 seeds/acre	Monosem precision row crop planter	--
Tennessee	Two 30-in. rows, 28-30 ft. long	Yes, 150 lbs. N/acre applied at all locations	--	--	--	--	--	Uniform seeds/acre /location	Precision seeding planter	--
Texas	Two 30- to 40-in. rows, 21 to 26 ft. long	Yes, varying nutrients and amounts applied	Yes	--	Yes/No, differs by location	--	Yes	--	John Deere 7100 Max-Emerge plot planter equipped with cones	John Deere 3300 plot combine
Virginia	Two 30-in. rows, 25 ft. long (one location was one 30-in. row, 30 ft. long)	Yes, varying nutrients and amounts applied	--	--	--	Yes	--	--	Winter-steiger PlotKing 2600 (one location was hand-planted)	Massey Ferguson 8XP plot combine (half of one location was harvested by hand)

State	Plot Size	Fertilizer	Herbicide	Fungicide	Insecticide	Pesticide	Tillage	Planting Rate	Planting Method	Harvesting Method
Wisconsin	Two rows wide	Yes, applied as recommended by soil tests	Yes	--	Yes, at locations where previous crop was corn	--	Yes	--	Precision vacuum corn planter	Self-propelled corn combine
Wyoming	--	--	--	--	--	--	--	--	--	--

* Full 2011 report unavailable. 2012 report information given.
-- No information reported.

Table 1. Description of reported trial management practices by state in 2011.

A spatially explicit database of trial yields across the United States was essential to analysis of YT and YF. We determined which states perform maize variety trials through a review of the crop variety trials conducted in each state. We then compiled all annual maize variety trial reports available online through the website of the trial-conducting university in each state. Only the most recent year or few most recent years of trial data were typically available through a university's website.

In an effort to obtain older trial data and reduce the risk of human error in compiling and formatting data, we requested all available trial data in spreadsheet format from the maize variety trial personnel. Through the reports available online and through the correspondence with the university maize variety trial personnel, we amassed the necessary trial data.

There was a wide range as to which years of trial data each state had available and as to which metadata each state included in both its spreadsheet files and its reports. However, across all states included in this study, the following data were formatted into a layout easily transferred into the statistical analysis program JMP (SAS Institute Inc., Cary, NC): hybrid name, year of trial, trial location (city and/or county and state), maize yield, and binary irrigation data (rainfed trial or irrigated trial). The compiled data for each state were cleaned of any repetitions and missing data. If the county of the trial location was not provided, the county was determined based upon the provided city or the provided latitude and longitude coordinates and added into the database to serve as a basis of comparison with county-level USDA-NASS data.

Following the compilation of the years of data for each state, trial data from the 32 available trial-performing states across the years 2006–2011 were compiled into a single JMP (SAS Institute Inc., Cary, NC) file (Figure 1).

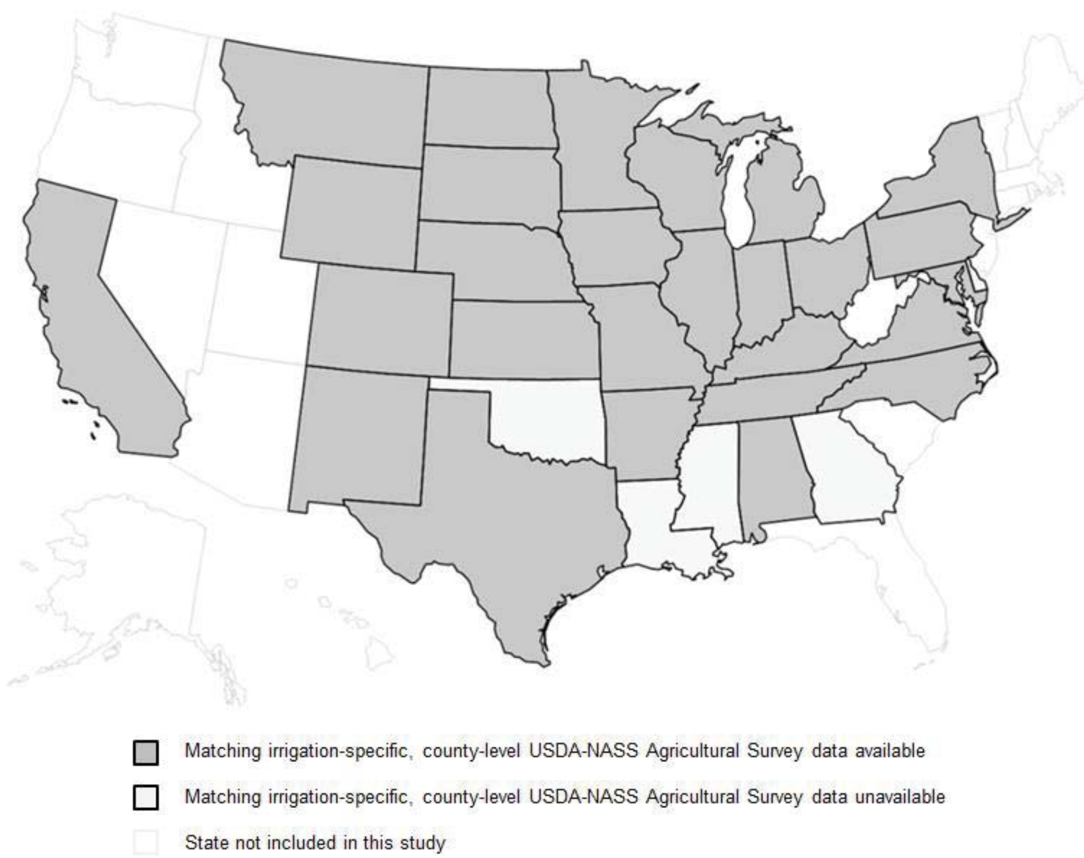


Figure 1. The 32 maize trial-performing states included in this study and the 27 states that have matching irrigation-specific, county-level YF data available.

The resultant 32-state database contains 129,499 hybrid entries. The years 2006–2011 were selected because a majority of the 32 states in the database had maize trial yield data extending back to 2006, though all of those years of data were not obtainable for all states due to lack of trials or lack of available data. Prior to 2006, data availability among states was much more sparse.

Each county tested between six and 354 unique entries per year, with a mean of 93 ± 62, calculated as the mean of all county-years across all states ± standard deviation (Table 2).

State Name	Years with No Trial Data Available	Total No. Counties Represented in 2006-2011 Trials	No. Counties Represented per Year in 2006-2011 Trials	Total No. County-Years 2006-2011	Mean No. Hybrids Tested per County-Year ± S.D.
Alabama		7	5-7	40	49 ± 32
Arkansas		6	4-6	29	88 ± 7
California		3	3	18	18 ± 3
Colorado		11	3-9	35	25 ± 12
Delaware	2006-2008	3	2	6	92 ± 45
Georgia		6	6	36	56 ± 29
Illinois		12	11-12	68	144 ± 40
Indiana		14	9-11	60	129 ± 53
Iowa		22	15-17	98	149 ± 67
Kansas		20	13-17	93	71 ± 33
Kentucky		12	5-7	37	151 ± 15
Louisiana	2008	4	2-4	17	83 ± 45
Maryland		4	4	24	115 ± 43
Michigan		18	10-13	75	103 ± 48
Minnesota		11	8-10	51	118 ± 38
Mississippi		9	6-8	39	82 ± 24
Missouri	2006	18	14-16	74	102 ± 38
Montana	2010	1	1	5	20 ± 8
Nebraska	2010	29	16-23	90	60 ± 28
New Mexico		1	1	6	19 ± 8
New York	2006	11	8-9	43	26 ± 18

State Name	Years with No Trial Data Available	Total No. Counties Represented in 2006-2011 Trials	No. Counties Represented per Year in 2006-2011 Trials	Total No. County-Years 2006-2011	Mean No. Hybrids Tested per County-Year ± S.D.
North Carolina		9	5-8	38	67 ± 18
North Dakota		14	3-10	40	64 ± 44
Ohio		11	8-10	54	131 ± 36
Oklahoma		3	1-2	8	26 ± 13
Pennsylvani a		20	10-16	81	28 ± 13
South Dakota		6	6	36	231 ± 65
Tennessee	2006, 2007	7	4-6	20	156 ± 63
Texas		13	8-12	60	34 ± 10
Virginia		7	5-7	35	86 ± 56
Wisconsin		14	11-13	74	139 ± 48
Wyoming		2	1-2	9	11 ± 3

Table 2. Description of the trial data by state. A county-year is a unique county and year combination. Mean number of hybrid entries tested per county-year calculated as mean of all county-years grouped by state, plus or minus standard deviation (S.D.).

For a yield-to-yield comparison to USDA-NASS data, whose yields are simply reported as one yield value per county (unless an irrigated or rainfed yield is also reported), the 90th percentile of hybrid yields in each trial was calculated. In this study, the 90th percentile of hybrid entry yields in each trial is YT, the estimate of yield potential. YTr indicates a rainfed YT, and YTi indicates an irrigated YT.

Farm Data. Rainfed, irrigated, and overall county-level maize yield data were downloaded from the USDA-NASS Quick Stats 2.0 (USDA-NASS n.d.). These data were downloaded only for those county-years from the period of 2006–2011 represented in the maize trial database. YFr indicates a rainfed YF, and YFi indicates an irrigated YF.

All states represented in this study had YF data; however, only 10 of the 32 states reported YF on a rainfed and/or irrigated basis for any number of years from 2006–2011. As irrigation is a critical factor impacting maize yield, an overall (i.e., non-irrigation-discriminate) YF would not provide a direct comparison to either YTr or YTi. The USDA-NASS Census of Agriculture, conducted every five years, is more comprehensive than the annual Agricultural Survey, and the 2007 Census of Agriculture was therefore used to create a more complete characterization of irrigation in the 32 states (Figure 2).

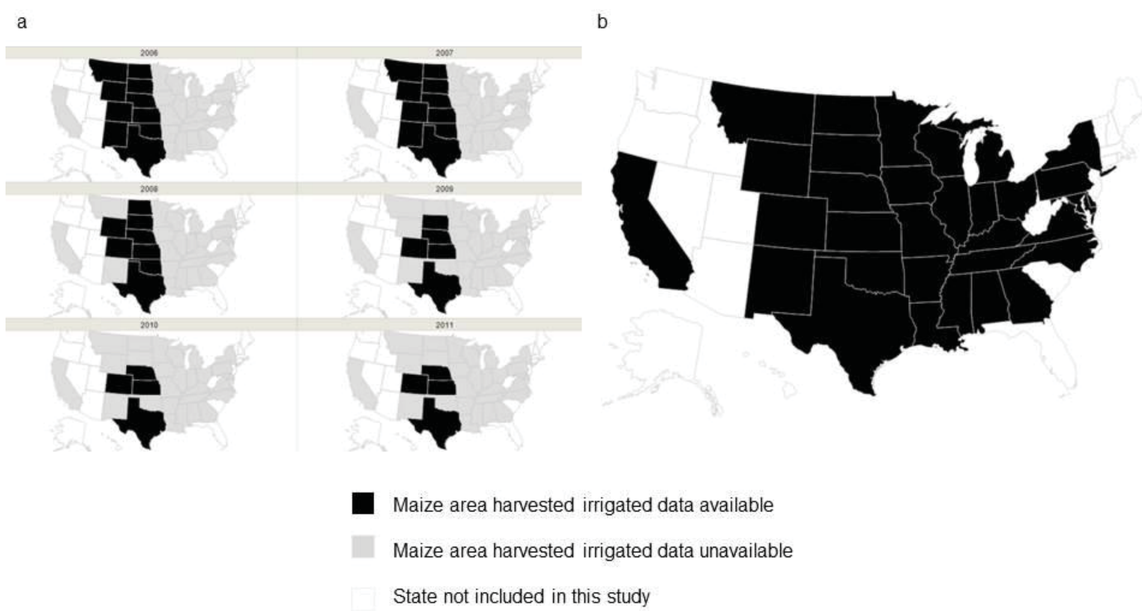


Figure 2. Availability of maize area harvested irrigated data in each trial state as taken from the 2006-2011 USDA-NASS Agricultural Surveys (a) and the 2007 USDA-NASS Census of Agriculture (b).

The percentage of area harvested that is irrigated was calculated (by dividing irrigated area harvested by overall area harvested and multiplying by 100) from both state-level survey and state-level census data. Only Colorado, Kansas, Nebraska, and Texas had complete irrigation data available for each year during 2006-2011 through the agricultural surveys. For the 28 other states, only six states had between two and four years of survey irrigation data; the 22 remaining states only had 2007 census irrigation data available. Of these 28 states with incomplete irrigation data available, the percentage of area harvested that was irrigated was below 25% for 18 states, and the percentage of area harvested that was irrigated was above 75% for five states. Irrigation assumptions were made in these 23 states with missing irrigation data to provide as many YFr-to-YTr and YFi-to-YTi comparisons as possible from the available data. If the percentage of area harvested irrigated was below 25%, the yields in the state were assumed to be rainfed yields; if the percentage of area harvested was above 75%, the yields in the state were assumed to be irrigated (Figure 3). Five states had percentages of area harvested irrigated between 25 and 75%. Irrigation was unable to be conclusively assumed for these states, so these five states (Delaware, Georgia, Louisiana, Mississippi, and Oklahoma) were eliminated from the yield gap analyses to avoid unreliable YF-to-YT comparisons.

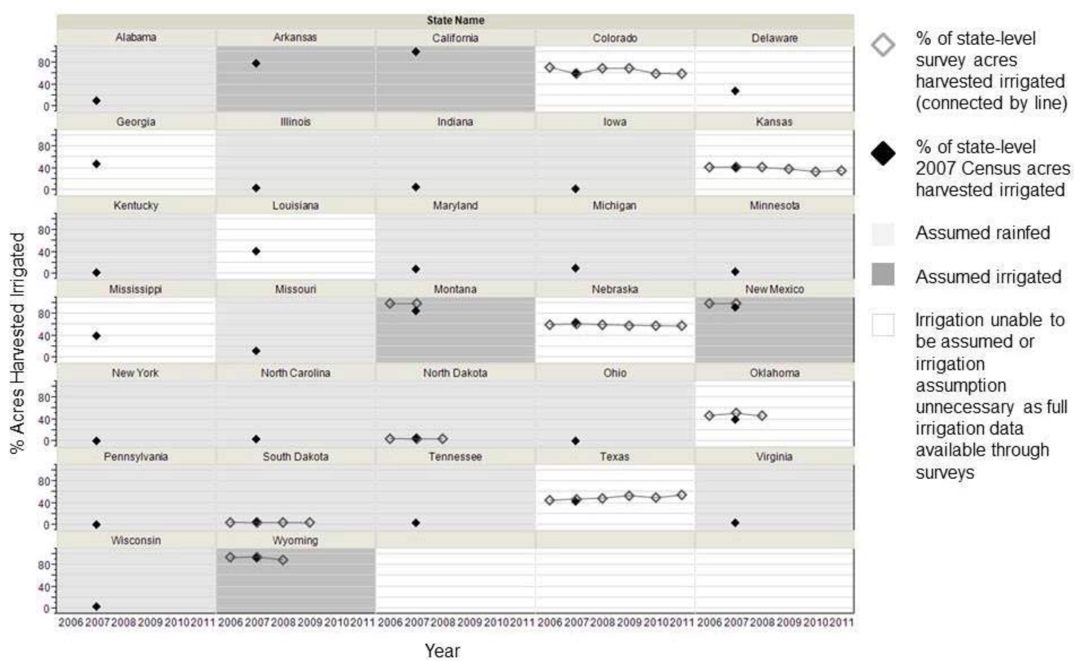


Figure 3. Percentage of corn grain acres harvested irrigated in each trial state as taken from the 2006-2011 USDA-NASS Agricultural Surveys and the 2007 USDA-NASS Census of Agriculture.

In summary, we began with all available 2006-2011 yield data from all of the counties in the United States performing maize variety trials that we were able to collect. All YT values were identified as either rainfed or irrigated. Next, USDA-NASS yield data for only those maize variety trial-performing counties, and only those years for each of those counties for which we had YT data, were downloaded. Rainfed or irrigated conditions were identified for YF data through thorough USDA-NASS Agricultural Survey data for four states and USDA-NASS 2007 Census of Agriculture assumptions for 23 states. YT data for which matching county-year YF data were not available or for which binary irrigation designations could not be determined for the matching county-year YF data were eliminated from the database. YF data for which binary irrigation designations could not be determined were eliminated from the database. These data eliminations were made to ensure only the most relevant, direct comparisons between YF and YT. This produced one comprehensive database of 2006-2011 YT and YF data with only irrigation-specific, YF-YT-matching county-years included (Table 3).

State Name	Years with Matching Data Unavailable	Total No. Matching Counties Represented in 2006-2011	No. Matching Counties Represented per Year in 2006-2011	Total No. Matching County-Years 2006-2011	No. Matching County-Years without Irrigation Data Reported or Assumed	No. Remaining Matching County-Years 2006-2011 for Direct Irrigation Comparison
Alabama		7	2-6	30		30
Arkansas		6	3-5	26		26
California		3	3	18		18
Colorado		11	3-9	30	9	21
Delaware	2006-2008	3	2	6	6 (All)	0
Georgia		5	2-5	20	20 (All)	0
Illinois		12	11-12	68		68
Indiana		14	9-11	59		59
Iowa		22	15-17	98		98

State Name	Years with Matching Data Unavailable	Total No. Matching Counties Represented in 2006-2011	No. Matching Counties Represented per Year in 2006-2011	Total No. Matching County-Years 2006-2011	No. Matching County-Years without Irrigation Data Reported or Assumed	No. Remaining Matching County-Years 2006-2011 for Direct Irrigation Comparison
Kansas		19	9-16	75	37	38
Kentucky		11	5-6	34		34
Louisiana	2008	4	2-4	16	16 (All)	0
Maryland		4	3-4	22		22
Michigan		17	10-13	74		74
Minnesota		11	7-10	50		50
Mississippi		8	5-7	37	37 (All)	0
Missouri	2006	18	14-15	72		72
Montana	2010	1	1	5		5
Nebraska	2010	29	16-23	90	4	86
New Mexico	2010-2011	1	1	4		4
New York	2006	11	3-9	34		34
North Carolina		9	5-8	38		38
North Dakota		13	3-7	33		33
Ohio		11	8-10	53		53
Oklahoma	2009, 2010	2	1-2	5	5 (All)	0
Pennsylvania		20	9-14	70		70
South Dakota		6	6	36		36
Tennessee	2006, 2007	5	3-4	14		14
Texas		13	7-11	58	50	8
Virginia		7	4-7	31		31
Wisconsin		13	10-13	72		72
Wyoming		2	1-2	8		8

Table 3. Description of the final database by state. This final database contains only matching, irrigation-specific county-years of maize variety trial and USDA-NASS data.

Of the 1,286 matching county-years, we were unable to determine YF irrigation in 184 county-years due to the states' irrigation data not being included in the USDA-NASS annual Agricultural Survey or the states' irrigation not being above 75% or below 25% by the 2007 Census of Agriculture and therefore unable to be safely assumed as irrigated or rainfed. Thus, the final database contains 1,102 county-years of matching, irrigation-specific data.

As 2006–2011 is not a long enough time period to examine meaningful temporal trends, the mean of those years of data was calculated for each county, resulting in one YF value and one YT value reported for each county instead of one YF and one YT value reported for each year of each county (each county-year).

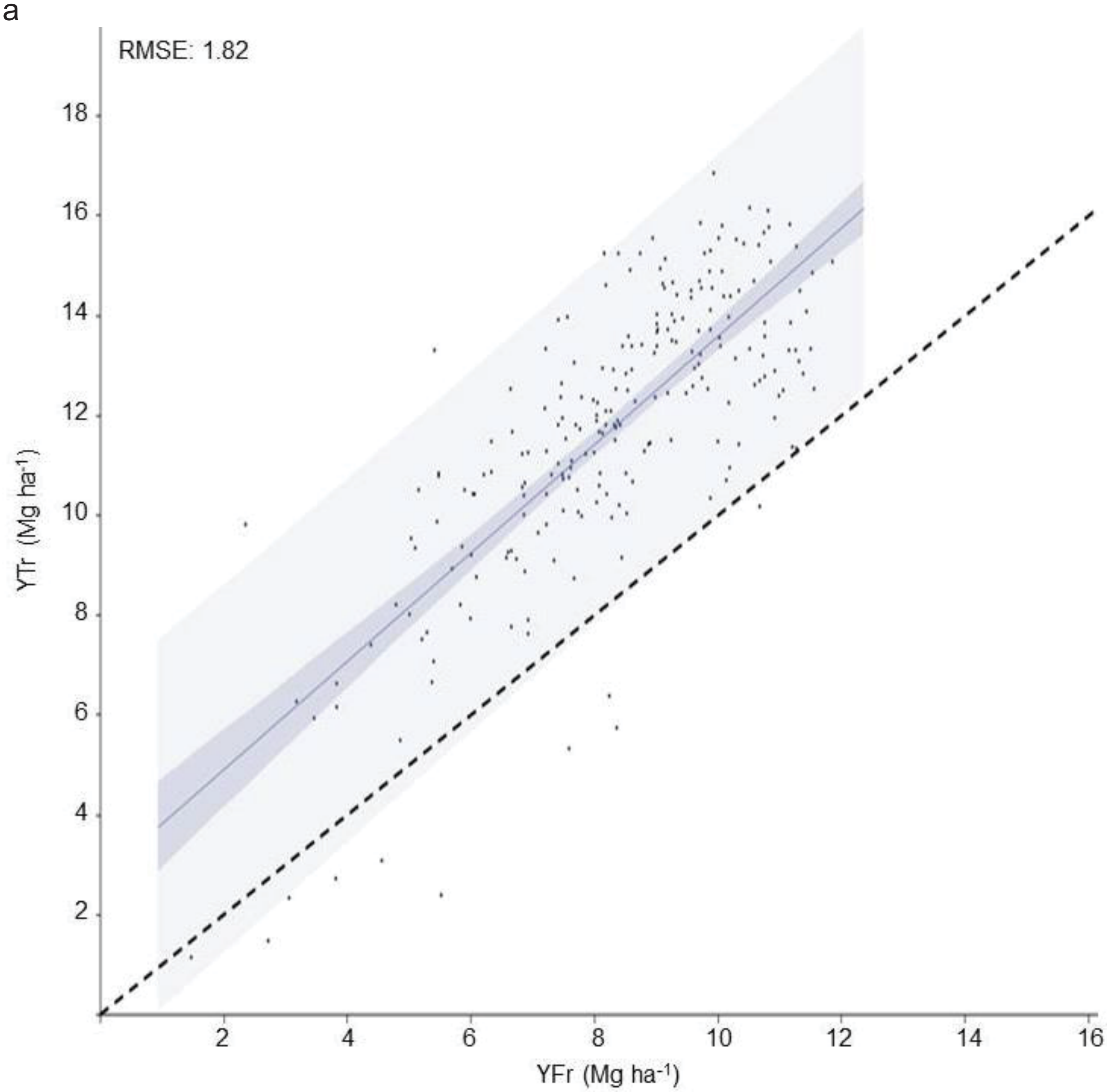
Results

County-Level. A scatter plot of county-level rainfed (a) and irrigated (b) YT versus YF with linear fit displays the larger overall depiction of all YG calculated in this study (Figure 4). The dashed black line represents what the line of best fit would be if the YG did not exist (i.e., if YF and YT were equal). The same trend is apparent in both rainfed and irrigated data. Though both the data-fitted and the 1:1 lines have positive slopes, the data-fitted lines of best fit and the corresponding 95% confidence intervals lie above the 1:1 theoretical zero YG lines. The 95% rainfed and irrigated prediction intervals lie almost completely above the 1:1 line. The rate of change of the conditional mean of YT with respect to YF was about 1.1 for rainfed maize and about 0.7 for irrigated maize, as seen in the equations of linear fit:

$$YTr = 2.8 + 1.1 * YFr$$

$$YTi = 6.4 + 0.7 * YFi$$

a



b

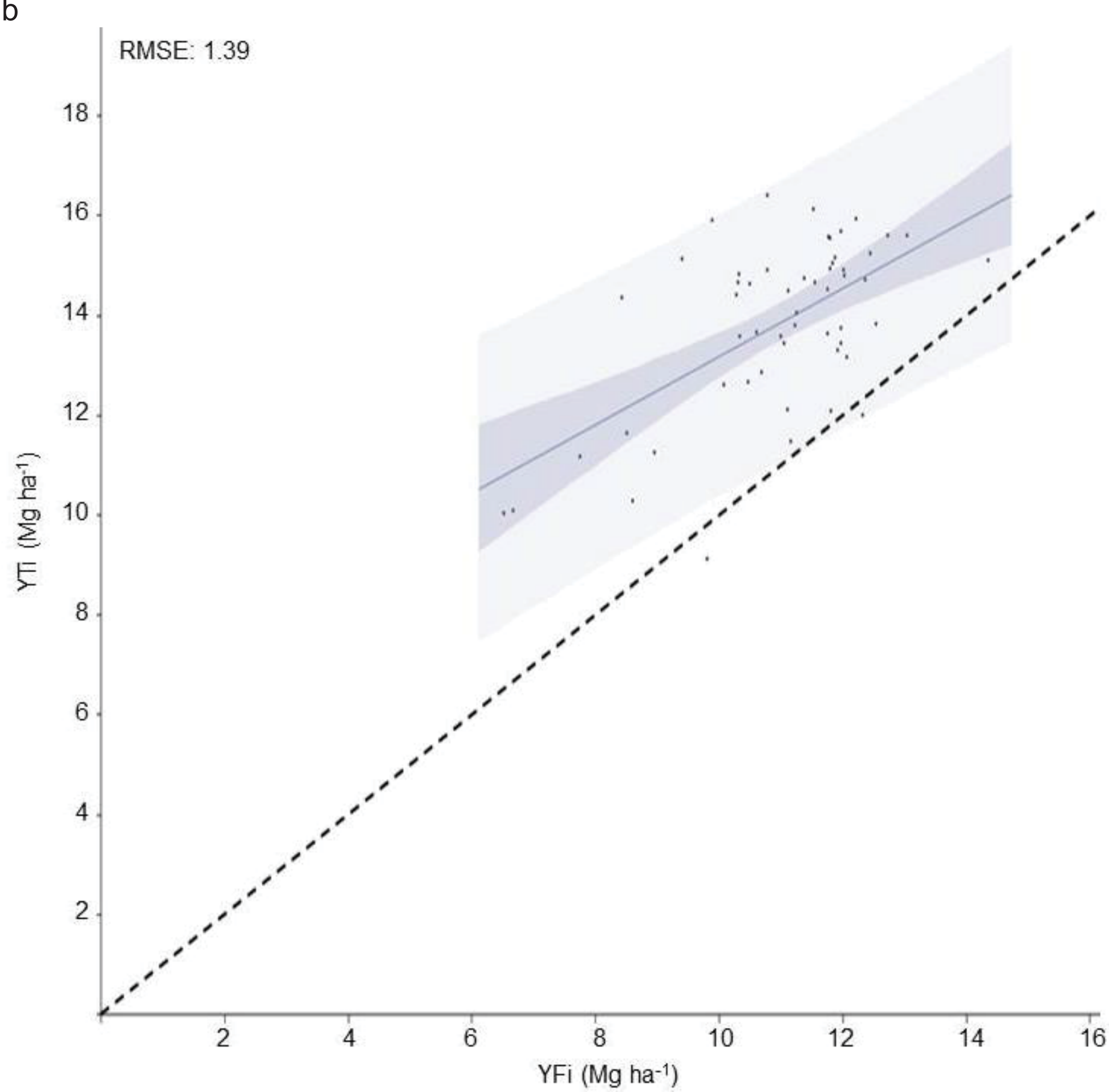


Figure 4. Linear regression of county-level YF on county-level YT (points) with line of fit (solid line), 95% confidence interval (darkly shaded areas), 95% prediction interval (lightly shaded areas), and root mean square error values for (a) rainfed and (b) irrigated data. The black dashed lines represent what the lines of fit would look like without the existence of a YG.

Grouped by State. To provide a YG value for each state, the medians of county-level YFs, YTs, and YGs were calculated for each state. Median state-grouped YTr ranged from 4.2 Mg ha^{-1} in Colorado to 15.0 Mg ha^{-1} in Indiana. Median state-grouped YFr ranged from 2.1 Mg ha^{-1} in Wyoming to 11.0 Mg ha^{-1} in Illinois. Median state-grouped

YTi ranged from 9.0 Mg ha^{-1} in Virginia to 16.2 Mg ha^{-1} in New Mexico. Median state-grouped YFi ranged from 7.6 Mg ha^{-1} in Texas to 12.0 Mg ha^{-1} in Colorado (Figure 5).

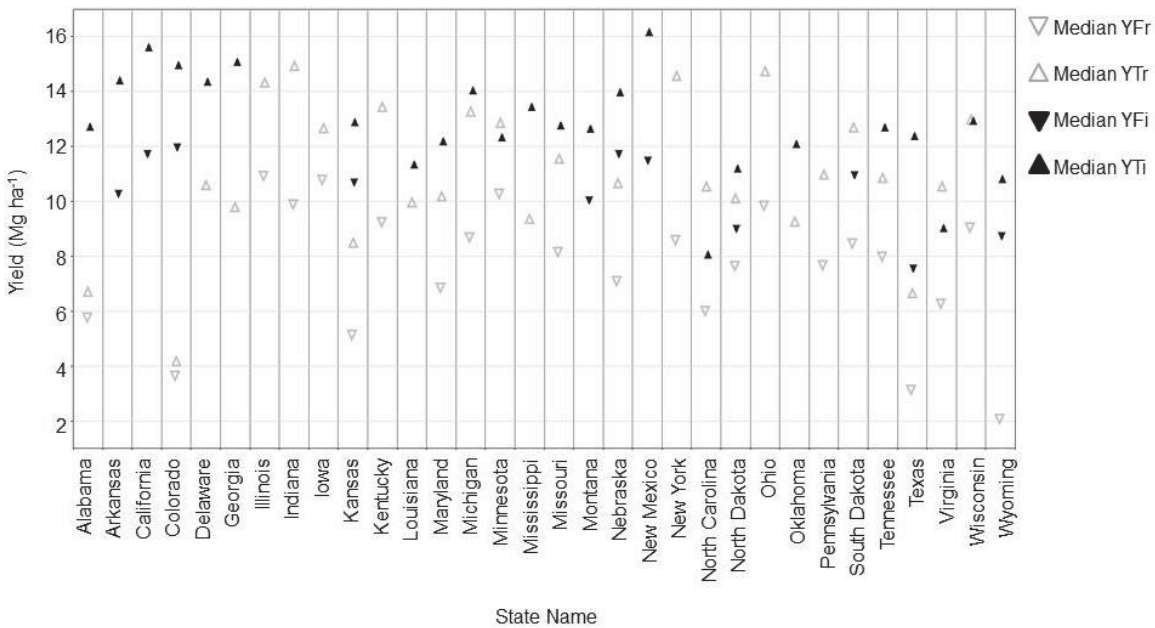


Figure 5. Median YFr, YFi, YTr, and YTi points by state, in Mg ha^{-1}.

The largest YGr of 53%, by a margin of 10%, was seen in Texas. The next largest median YGr of 43% was seen in North Carolina, closely followed by the median YGr of 41% in New York. The smallest YGr of 13% was seen in Colorado, closely followed by a median YGr of 14% in Alabama and 15% in Iowa. By a margin of 10%, the largest YGi of 39% was seen in Texas. The next largest median YGi of 29% was seen in both Arkansas and New Mexico. The smallest YGi of 16% was seen in Nebraska, closely followed by a median YGi of 17% in Kansas. This equates to a 40% range of median YGr values and a 23% range of median YGi values (Figure 6).

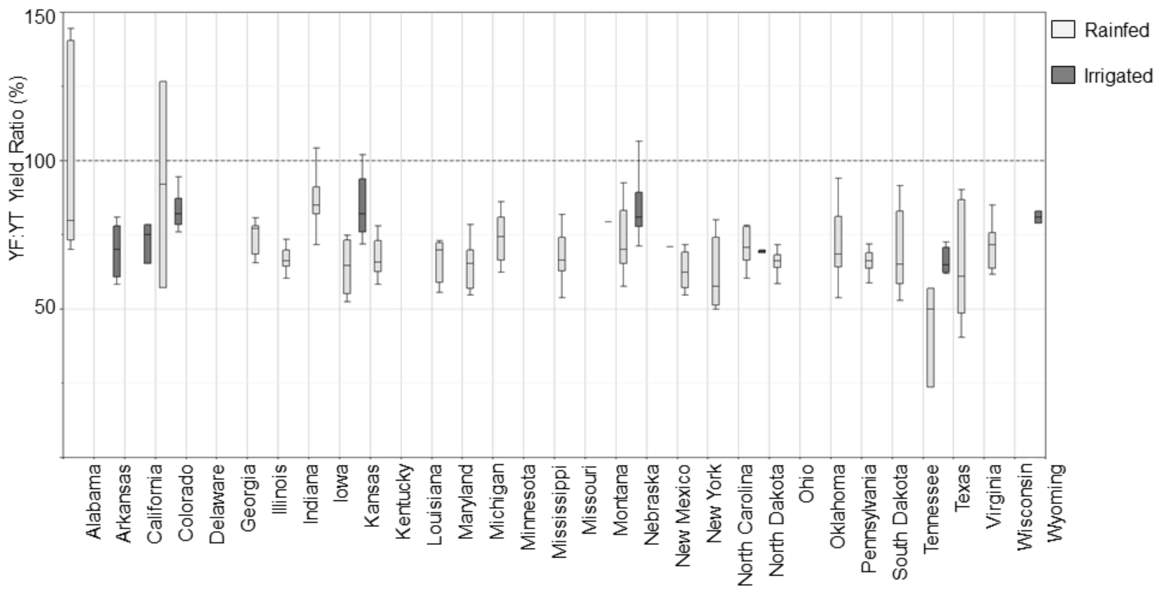


Figure 6. Box plots of the YF:YT yield ratio (%) by state. Boxes represent the interquartile ranges. Whiskers represent 1.5 times the interquartile ranges. Zero YG (YF equal to YT) is indicated by the horizontal dashed, black line. YG above this line indicate a YF larger than YT. Outliers are omitted; see Table 4 for summary of outlier data

Outlier YG Points *(Not Shown)*							
State Name	County Name	YTr (Mg ha^{-1})	YFr (Mg ha^{-1})	YTi (Mg ha^{-1})	YFi (Mg ha^{-1})	YGr (%)	YGi (%)
Indiana	Jackson	15.3	8.1			47	
Kansas	Greeley	2.8	3.8	16.0	12.2	-37	24
Nebraska	Hayes	1.5	2.7		12.4	-77	
Nebraska	Keith	1.2	1.4		11.5	-22	
North Dakota	Sargent	6.4	8.2			-28	
North Dakota	Williams	2.4	5.5	11.2	7.7	-125	31
Ohio	Wood	11.5	10.0			14	
Pennsylvania	Venango	5.8	8.3			-44	

Table 4. Summarized Figure 6 outlier data.

	AL	AR	CA	CO	IL	IN	IA	KS	KY	MD	MI	MN	MO	MT	NE	NM	NY	NC	ND	OH	PA	SD	TN	TX	VA	WI	WY
Rainfed County N	7	0	0	2	12	14	22	8	11	4	13	10	13	0	14	0	11	9	13	11	20	6	5	2	7	13	0
Irrigated County N	0	6	3	8	0	0	0	9	0	0	0	0	0	1	18	1	0	0	1	0	0	0	0	4	0	0	2

Table 5. Numbers of both rainfed and irrigated counties represented in each state in Figure 6.

Discussion

It was hypothesized that the trials would out-perform the average farmer, as this is what has been shown by previous yield gap studies. In this study, median state-grouped YGr range from 13 to 53%, and median state-grouped YGi range from 16 to 39%. The magnitude of these differences between YF and YT indicates that maize yields in the United States, particularly rainfed, have considerable room for improvement.

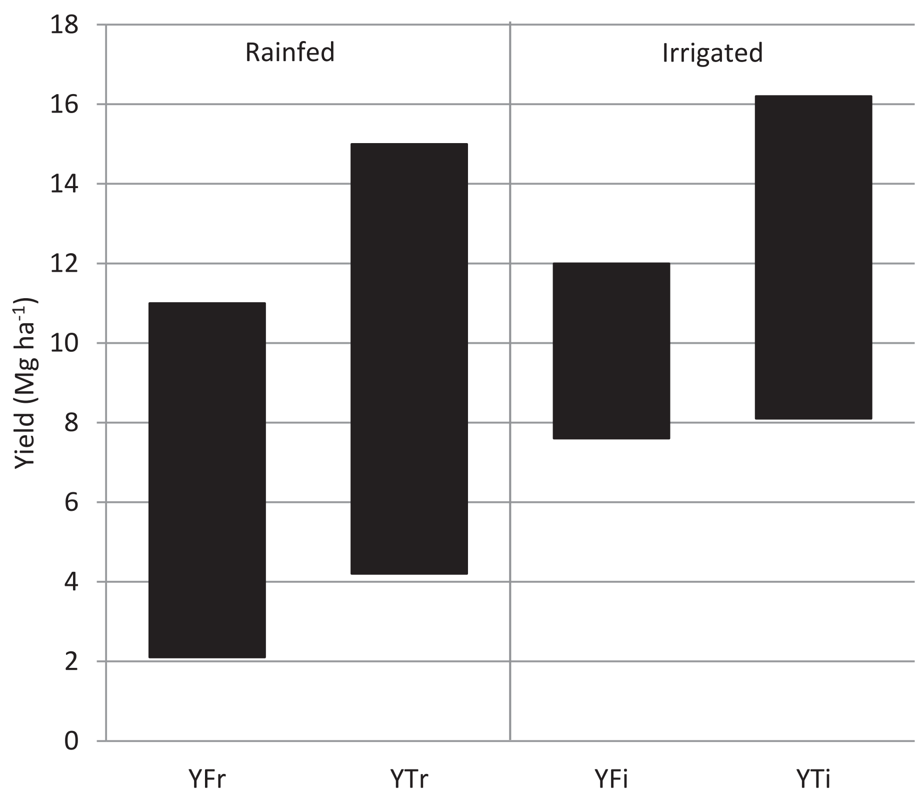


Figure 7. Summary of median state-grouped yield minimums and maximums.

Irrigated yields and YGi show narrower ranges than rainfed yields and YGr. YFi and YTi ranges have greater minimums and greater maximums than YFr and YTr, though YGr has a greater maximum than YGi. The 40% range of median state-grouped YGr values and the 23% range of median state-grouped YGi values suggest that across the United States, the YG varies greatly between states. As irrigation is a large contributing factor to yield, rainfed maize is more greatly affected by precipitation patterns and susceptible to drought stress if it is a growing season with little rain. Without irrigation to provide more consistent yields, it is not surprising that YGr had a greater range than YGi. Despite the differences between rainfed and irrigated yield data, both rainfed and irrigated maize variety trials are producing substantially higher yields than those yields produced on the average farm field.

The rate of change of the conditional mean of YT with respect to YF was about 1.1 for rainfed maize and 0.7 for irrigated maize. This means that for every 1 Mg ha^{-1} increase in YFr, YTr increased 1.1 Mg ha^{-1}, and for every 1 Mg ha^{-1} increase in YFi, YTi increased 0.7 Mg ha^{-1}. Both rainfed and irrigated fitted line slopes are close to 1, which is the slope of the theoretical zero YG, or 1:1, line, but the intercepts of 2.8 Mg ha^{-1} for rainfed maize and 6.4 Mg ha^{-1} for irrigated maize show that the fitted lines lie well above the zero YG line.

While there are no directly comparable results to which we might compare ours, there are others that estimate yield potential through other methods (Table 7 and Figure 7). Duvick and Cassman (1999) found a YGi of 44% and a YGr of 60% for maize in Nebraska, United States, using contest-winning yields as an estimate of yield potential. Lobell et al. (2009) states that YG is often greater than 80% for tropical maize in Africa. In addition, based upon their literature review, global YG across multiple crops (including maize) were found to range from 5% to 84%, however, nearly all major global cropping systems have YG in the 20% to 80% range. The average YGr for United States maize was found to be 35%, and the average YGi for United States maize was found to be 25% by Lobell et al. (2009), with yield potential estimated by a previous study's Hybrid-Maize model simulations (Yang et al. 2004). Van Wart et al. (2013) reported a 27% YGr and a 23% YGi for maize in the United States. Meng et al. (2013) examined the maize YG in China in three different ways. The authors reported a 44% to 52% YG when yield potential was estimated by the Hybrid-Maize Model, a YG of 49% when yield was estimated by the highest recorded yield, and a YG of 36% when yield potential was estimated by farm field experiments.

It has been hypothesized that, in practicality, the average national YG plateaus at approximately 15% to 25% (Cassman et al. 2003; Lobell, Cassman, and Field 2009). The YG studies of United States maize mentioned above by Duvick and Cassman (1999), Lobell et al. (2009), and van Wart et al. (2013) report YG ranging from 23% to 60%.

Study	YGi (%)	YGr (%)
Current	16–39	13–53
Van Wart et al. (2013)	23	27
Lobell et al. (2009)	25	35
Duvick and Cassman (1999)	44	60

Table 6. Summary of this and other cited studies that examined YG as YGi and YGr.

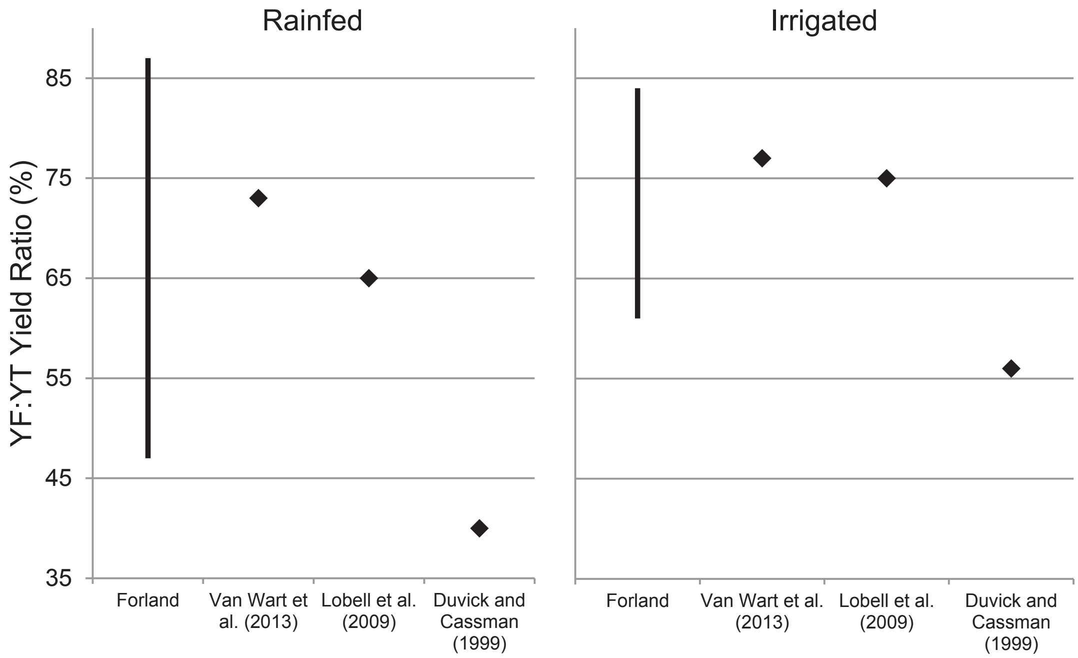


Figure 8. Depiction of the study results summarized in Table 7 as yield ratios.

Land-grant university crop variety trial data are vast and valuable in analyzing yield. These data have always been available but to our knowledge have not yet been extensively compiled and analyzed with this method. Although variety trial data further the understanding of yield potential and YG, it would be unrealistic to say that trial yields are currently attainable by all farmers or that the YG could ever be zero (Fischer, Byerlee, and Edmeades 2009). Farmers are limited by economics and the land quality of their fields, among other things. Management of trial plots is overseen by agronomic experts in university trials, whereas farmers, either by necessity or lack of expert knowledge, may not be managing their fields with best management practices. In addition to this, farmer's fields are many times larger than the small-plot university trials and therefore difficult to manage with the same attention to detail. These items need to be taken into account when considering how to apply the results of this study.

The bold biofuel production targets of the United States' current renewable energy policies are placing strain on crop production already pushed by a forecasted approximate doubling of global food demand by 2050. The pressures to intensify crop production and to determine realistic crop production potential are immense, meaning there is a great need for research to improve the understanding and quantification of yield potential and YG and to provide estimates of the spatial variability of these two parameters on a focused spatial scale.

This study met these research needs by compiling a database of 2006–2011 maize yield data for nearly all of the maize variety trial-performing states for comparison of YF to YT and the calculation of YG. Further research that could prove beneficial includes examination of temporal trends in land grant university crop trial best management practices, comparison of trial yields to other measures of yield potential, and application of these methods to other crops.

References

Abeledo, L.G., R. Savin, and G.A. Slafer. 2008. Wheat Productivity in the Mediterranean Ebro Valley: Analyzing the Gap between Attainable and Potential Yield with a Simulation Model. European Journal of Agronomy 28: 541–550.

Aggarwal, P.K., K.B. Hebbar, M.V. Venugopalam, et al. 2008. Quantification of Yield Gaps in Rain-Fed Rice, Wheat, Cotton and Mustard in India. 43. Global Theme on Agroecosystems. Patancheru, India: International Crops Research Institute for the Semi-Arid Tropics.

Cassman, K.G., A.R. Dobermann, D.T. Walters, and H. Yang. 2003. Meeting Cereal Demand While Protecting Natural Resources and Improving Environmental Quality. Annual Review of Environment and Resources 28: 315–358.

Duvick, D.N., and K.G. Cassman. 1999. Post-Green Revolution Trends in Yield Potential of Temperate Maize in the North-Central United States. Crop Science 39: 1622–1630.

Evans, L.T. 1993. Crop Evolution, Adaptation and Yield. New York: Cambridge University Press.

Fargione, J., J. Hill, and D. Tilman. 2008. Land Clearing and the Biofuel Carbon Debt. Science 319: 1235–1238.

Fischer, R.A., D. Byerlee, and G.O. Edmeades. 2009. Can Technology Deliver on the Yield Challenge to 2050? *In* . Rome. ftp://ftp.fao.org/docrep/fao/012/ak977e/ak977e00.pdf, accessed October 24, 2013.

Foley, J.A., N. Ramankutty, K.A. Brauman, et al. 2011. Solutions for a Cultivated Planet. Nature 478(7369): 337–342.

Gerber, J., N. Mueller, and J. Foley. 2010. Empirical Estimation of Global Yield Gaps. *In* . Montpellier, France.

Grassini, P., J. Thornton, C. Burr, and K.G. Cassman. 2011.High-Yield Irrigated Maize in the Western U.S. Corn Belt. I. On-Farm Yield, Yield Potential, and Impact of Agronomic Practices. Field Crops Research 120: 142–150.

Hall, A.J., C. Feoli, J. Ingaramo, and M. Balzarini. 2013. Gaps between Farmer and Attainable Yields across Rainfed Sunflower Growing Regions of Argentina. Field Crops Research 143: 119–129.

Hochman, Z., D. Gobbett, D. Holzworth, et al. 2012. Quantifying Yield Gaps in Rainfed Cropping Systems: A Case Study of Wheat in Australia. Field Crops Research 136: 85–96.

Johnston, M., J.A. Foley, T. Holloway, C. Kucharik, and C. Monfreda. 2009. Resetting Global Expectations from Agricultural Biofuels. Environmental Research Letters 4(014004): 9pp.

Laborte, A.G., C.A.J.M. de Bie, E.M.A. Smaling, et al. 2012. Rice Yields and Yield Gaps in South East Asia: Past Trends and Future Outlook. European Journal of Agronomy 36: 9–20.

Licker, R., M. Johnston, J.A. Foley, et al. 2010. Mind the Gap: How Do Climate and Agricultural Management Explain the “yield Gap” of Croplands around the World? Global Ecology and Biogeography 19(6): 769–782.

Lobell, D.B., K.G. Cassman, and C.B. Field. 2009. Crop Yield Gaps: Their Importance, Magnitudes, and Causes. Annual Review of Environment and Resources 34: 179–204.

Meng, Q., P. Hou, L. Wu, X. Chen, and F. Zhang. 2013. Understanding Production Potential and Yield Gaps in Intensive Maize Production in China. Field Crops Research 143: 91–97.

Mueller, N. D., J.S. Gerber, M. Johnston, et al. 2012. Closing Yield Gaps through Nutrient and Water Management. Nature 490(7419): 254–257.

Scanlon, B., I. Jolly, M. Sophocleous, and L. Zhang. 2007. Global Impacts of Conversions from Natural to Agricultural Ecosystems on Water Resources: Quantity versus Quality. Water Resources Research 43(3): W03437.

Tack, J., A. Barkley, and L.L. Nalley. 2015. Estimating Yield Gaps with Limited Data: An Application to United States Wheat. American Journal of Agricultural Economics Advance Access.

Tilman, D., C. Balzer, J. Hill, and B. Befort. 2011. Global Food Demand and the Sustainable Intensification of Agriculture. Proceedings of the National Academy of Sciences USA 108: 20260–20264.

Tittonell, P., B. Vanlauwe, M. Corbeels, and K.E. Giller. 2008. Yield Gaps, Nutrient Use Efficiencies and Response to Fertilisers by Maize across Heterogeneous Smallholder Farms of Western Kenya. Plant Soil 313: 19–37.

United States Department of Agriculture-- National Agricultural Statistics Service. N.d. http://www.nass.usda.gov/, accessed September 20, 2013.

Van Ittersum, M.K., K.G. Cassman, P. Grassini, J. Wolf, P. Tittonell, and Z. Hochman. 2013. Yield Gap Analysis with Local to Global Relevance--A Review. Field Crops Research 143: 4-17.

Van Wart, J.K. C. Kersebaum, S. Peng, M. Milner, and K.G. Cassman. 2013. Estimating Crop Yield Potential at Regional to National Scales. Field Crops Research 143: 34–43.

What Is a Land-Grant College? 2009. Washington State University Extension. http://ext.wsu.edu/documents/landgrant.pdf, accessed October 24, 2013.

Yang, H.S., A. Dobermann, J.L. Lindquist, et al. 2004. Hybrid Maize—a Maize Simulation Model That Combines Two Crop Modeling Approaches. Field Crops Research 87(2-3): 131–154.