



# How good is good enough? Data requirements for reliable crop yield simulations and yield-gap analysis



Patricio Grassini<sup>a,\*</sup>, Lenny G.J. van Bussel<sup>b</sup>, Justin Van Wart<sup>a</sup>, Joost Wolf<sup>b</sup>, Lieven Claessens<sup>c,d</sup>, Haishun Yang<sup>a</sup>, Hendrik Boogaard<sup>e</sup>, Hugo de Groot<sup>e</sup>, Martin K. van Ittersum<sup>b</sup>, Kenneth G. Cassman<sup>a</sup>

<sup>a</sup> University of Nebraska-Lincoln, PO Box 830915, Lincoln, NE 68583-0915, USA

<sup>b</sup> Plant Production Systems Group, Wageningen University, PO Box 430, 6700 AK Wageningen, The Netherlands

<sup>c</sup> International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), PO Box 39063, 00623 Nairobi, Kenya

<sup>d</sup> Soil Geography and Landscape Group, Wageningen University, PO Box 47, 6700 AA Wageningen, The Netherlands

<sup>e</sup> Alterra, Wageningen University and Research Centre, PO Box 47, 6700 AA Wageningen, The Netherlands

## ARTICLE INFO

### Article history:

Received 19 December 2014

Received in revised form 7 March 2015

Accepted 8 March 2015

Available online 30 March 2015

### Keywords:

Crop simulation  
Yield gap  
Yield potential  
Weather data  
Cropping system

## ABSTRACT

Numerous studies have been published during the past two decades that use simulation models to assess crop yield gaps (quantified as the difference between potential and actual farm yields), impact of climate change on future crop yields, and land-use change. However, there is a wide range in quality and spatial and temporal scale and resolution of climate and soil data underpinning these studies, as well as widely differing assumptions about cropping-system context and crop model calibration. Here we present an explicit rationale and methodology for selecting data sources for simulating crop yields and estimating yield gaps at specific locations that can be applied across widely different levels of data availability and quality. The method consists of a tiered approach that identifies the most scientifically robust requirements for data availability and quality, as well as other, less rigorous options when data are not available or are of poor quality. Examples are given using this approach to estimate maize yield gaps in the state of Nebraska (USA), and at a national scale for Argentina and Kenya. These examples were selected to represent contrasting scenarios of data availability and quality for the variables used to estimate yield gaps. The goal of the proposed methods is to provide transparent, reproducible, and scientifically robust guidelines for estimating yield gaps; guidelines which are also relevant for simulating the impact of climate change and land-use change at local to global spatial scales. Likewise, the improved understanding of data requirements and alternatives for simulating crop yields and estimating yield gaps as described here can help identify the most critical “data gaps” and focus global efforts to fill them. A related paper (Van Bussel et al., 2015) examines issues of site selection to minimize data requirements and up-scaling from location-specific estimates to regional and national spatial scales.

© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1. Introduction

Yield potential ( $Y_p$ ) is defined as the yield of an adapted crop cultivar as determined by solar radiation, temperature, carbon dioxide, and genetic traits that govern length of growing period, light interception by the crop canopy and its conversion to biomass, and partition of biomass to the harvestable organs (Evans, 1993; van Ittersum and Rabbinge, 1997). Water-limited yield potential ( $Y_w$ )

is determined by these previous factors and also by water supply amount and distribution during the crop growth period and field and soil properties that affect soil water availability such as slope, plant-available soil water holding capacity, and depth of the root zone (Lobell et al., 2009; van Ittersum and Rabbinge, 1997; Van Ittersum et al., 2013). For a specific location and year, the crop yield gap ( $Y_g$ ) is defined as the difference between  $Y_p$  (irrigated systems) or  $Y_w$  (rainfed) and average actual farm yield ( $Y_a$ ). The magnitude of  $Y_g$  provides a benchmark of current land productivity in relation to the biophysical yield ceiling, and an estimate of the additional crop production that could potentially be achieved, on existing cropland area, through improved management that alleviates all limiting factors other than weather factors. Estimates of  $Y_p$ ,

\* Corresponding author. Tel.: +1 402 472 5554; fax: +1 402 472 7904.

E-mail addresses: [pgrassini2@unl.edu](mailto:pgrassini2@unl.edu) (P. Grassini), [Lenny.vanBussel@wur.nl](mailto:Lenny.vanBussel@wur.nl) (L.G.J. van Bussel).

Yw, and Yg also provide the foundation for more detailed studies to identify underpinning causes of the observed Yg, and for *ex-ante* evaluation of impact from adoption of new technologies, changing climate, and land-use change.

Accuracy in Yg estimation depends on the error associated with estimates of Yp (or Yw) and Ya<sup>1</sup>. Amongst methods to estimate Yp or Yw, crop simulation models provide the most robust approach because they account for the interactive effects of genotype, weather, and management (GxExM) on yields across agro-ecological zones and years (Van Ittersum et al., 2013). To minimize errors in estimating Yp and Yw, crop simulation models require high-quality input-data on weather, soil, and crop management (Aggarwal, 1995; Rivington et al., 2005; Bert et al., 2007). These models need also to be rigorously evaluated for their ability to reproduce major GxExM interactions (Passioura, 1996; Kersebaum et al., 2007; Van Ittersum et al., 2013). Likewise, reliable simulation of Yp and Yw requires specification of the cropping system and water regime in which a crop is grown as determined by crop sequence, dates of sowing and physiological maturity for the most widely used cultivars, and whether the crop is fully irrigated, partially irrigated, or rainfed (Folberth et al., 2012; Van Wart et al., 2013c). Finally, the error associated with the estimate of average annual Ya will also determine the accuracy of the Yg estimate.

Crop yield simulation is an important component of yield-gap analysis, hence, the above-mentioned sources of uncertainty related with estimates of Yp (or Yw) also affect other kinds of studies that rely on crop yield simulations and the required data therein. For example, studies on climate change, and land use change involving crop simulation models applied at global or regional spatial scales are abundant in recent literature (e.g., Challinor et al., 2014a; Rosenzweig et al., 2014). However, several recent publications have identified a number of substantive concerns associated with data sources and methods used in such studies (Van Ittersum et al., 2013; Van Wart et al., 2013a). These concerns include: (i) poor quality of weather and soil data, (ii) unrealistic assumptions about the cropping-system context, (iii) poorly calibrated crop simulation models, and (iv) lack of transparency about underpinning assumptions and methods. For example, Nelson et al. (2010) used 50-y monthly average gridded (5' resolution) weather data and coarse assumptions about the cropping system (e.g., a single crop variety was simulated for the entire world) to produce a global assessment of climate change impact on crop yields and land-use change. A similar approach was followed by Bagley et al. (2012) to simulate changes in water availability and potential crop yields in the world's breadbaskets. In both studies, no information was provided about how models were calibrated to simulate yield potential. Similarly, Rosenzweig et al. (2014) used an ensemble of models to simulate crop yields based on gridded daily weather data, coarse assumptions about cropping systems, and crop model parameters that were forced to reproduce current regional or national Ya averages. Another pitfall of these three studies is failure to account for multiple-crop systems (i.e., fields planted with more than one crop in the same year, such as the rice-wheat system that is widely practiced in Asia) or cropping systems where irrigated and rainfed systems co-exist within the same geographic area.

In most cases, use of poor quality or coarse-scale weather, soil, and cropping-system data for yield-gap analysis, as well as for other studies on climate change, food security, and land-use change that rely on crop yield simulations, is due to the fact that high quality data at finer spatial resolution do not exist, so pragmatic short-cuts are required to achieve the full terrestrial coverage. These short-cuts, however, are rarely evaluated for their ability to reproduce

Yp, Yw and Yg values estimated using high-quality, measured data. Without such validation, Yp, Yw, and Yg estimates with coarse-scale data sources can seriously distort results, decreasing their usefulness to inform regional or national policies and effective prioritization of research and development investments for agriculture (Rivington et al., 2004; Van Wart et al., 2013a,c). In contrast, one can find studies on yield-gap analysis for specific locations with data that are only available for few and specific site-years, which are not representative of larger spatial areas and do not allow upscaling to regional or global levels (e.g., Fermont et al., 2009; Grassini et al., 2011). Surprisingly, despite wide use of crop simulation models for yield-gap analysis (263 results in the Web of Science by Nov 15th, 2014), there are no published guidelines about standard sources and quality of data input for weather, soil, actual yields, and cropping-system context, or requirements for calibration of crop models used in such studies.

In summary, a robust approach to simulate accurate crop yield potential and estimate Yg requires: (i) input data that meet minimum quality standards at the appropriate spatial scale, (ii) agronomic relevance with regard to cropping-system context, (iii) proper calibration of crop models used, and (iv) flexibility and transparency to account for different scenarios of data availability and quality. Here we address the current lack of guidelines on data and methods for yield gap analysis, by developing a systematic approach for selection of data inputs based on the lessons learned from establishing the Global Yield Gap Atlas ([www.yieldgap.org](http://www.yieldgap.org)). The paper focusses on yield-gap analysis at specific 'point' locations, and their surrounding inference zone, based on application of crop simulation models to estimate Yp or Yw (hereafter called 'targeted areas'). An inference zone is defined as an area with similar climate such that there is relatively little variation in crop management practices. This paper has implications not only for yield-gap analysis but also for other studies related with climate change, food security, and land-use change because these studies typically rely on crop yield simulations and the required data therein. A separate paper describes the methodology for site selection, spatial delimitation of the inference zone around a location, and up-scaling local estimates of Yg to regional and national scales (Van Bussel et al., 2015).

## 2. Data requirements for yield-gap analysis

### 2.1. Overview

Yield-gap analyses at large spatial scale require enormous amounts of input data, because simulated and actual crop yields are strongly determined by the spatial and temporal variation in environmental conditions and cropping system context. Based on the concept that it is better to use primary data for crop growth simulations than to use aggregated or interpolated average input data (De Wit and Van Keulen, 1987; Rabbinge and van Ittersum, 1994; Penning De Vries et al., 1997), the Global Yield Gap Atlas ([www.yieldgap.org](http://www.yieldgap.org)) utilizes a 'bottom-up' approach for yield-gap analysis. A limited number of locations are selected such that these account for the greatest proportion of total national production of the crop being evaluated. For these locations, 'point-based' estimates of Yp, Yw, Ya, and Yg are derived, which are subsequently up-scaled to climate zones and national spatial scales (Van Wart et al., 2013b; Van Bussel et al., 2015). This site selection and up-scaling process helps to limit the number of locations for which site-specific data on weather, soils, and cropping system are required, which in turn facilitates the focus on quality of the underpinning data and helps ensure local to global relevance of the analysis. Principles that underpin the data selection approach

<sup>1</sup> Accuracy is the closeness of a measurement (or simulation) to the true value.

**Table 1**  
Quality and availability of data required for yield gap analysis in three study regions.

Data input	Region		
	Nebraska, USA	Argentina	Kenya
<i>Weather</i>			
Source	HPRCC, NWS	INTA-SIGA, SMN	KMS
Availability of required variables	All	All, except solar radiation	All, except solar radiation
Available data-years	>20 yr	>20 yr	3–18 yr
Spatial distribution	High	Medium	Low
Data quality <sup>a</sup>	High	Medium	Low
Publicly accessible	Yes	Yes	No
<i>Soils</i>			
Source	USDA-NRCS	INTA-GeoINTA, INTA-Soil division	ISRIC-WISE
Spatial resolution	High	Intermediate	Coarse
Availability of required variables	All	All	All, except rootable depth
<i>Crop management<sup>b</sup></i>			
Source	USDA-RMA	None	None
Availability of required variables	Only sowing date	None	None
<i>Model calibration</i>			
Source	Research farms High-yield producer fields	Research farms	None
<i>Actual yield</i>			
Source	USDA-NASS	Ministry of Agriculture-SIIA	Ministry of Agriculture
Finest spatial resolution level <sup>c</sup>	County ( $\approx 2000 \text{ km}^2$ )	Department ( $\approx 4500 \text{ km}^2$ )	District ( $\approx 2500 \text{ km}^2$ )
Available data-years	All years	All years	Every other 2–3 yr
Data quality <sup>a</sup>	High	Intermediate	Poor
Publicly accessible	Yes	Yes	No

HPRCC: High Plains Regional Climate Center (<http://www.hprcc.unl.edu/>); NWS: National Weather Service (<http://www.weather.gov/>); INTA: Instituto Nacional de Tecnología Agropecuaria (<http://inta.gob.ar/>); SIGA: Sistemas de Información Clima y Agua (<http://climayagua.inta.gob.ar/>); SMN: Argentina National Meteorological Service (<http://www.smn.gov.ar/>); KMS: Kenya Meteorological Service (<http://www.meteo.go.ke/>); NRCS: National Resource Conservation Service (<http://www.nrcs.usda.gov/wps/portal/nrcs/site/soils/home/>); GeoINTA: (<http://geointa.inta.gov.ar/web/>); ISRIC-WISE: International World Soil Reference and Information Center; World inventory of soil emission potentials (<http://www.isric.org/projects/world-inventory-soil-emission-potentials-wise/>); USDA: United States Department of Agriculture (<http://www.usda.gov/wps/portal/usda/usdahome/>); NASS: National Agricultural Statistics Service (<http://www.nass.usda.gov/>); RMA: Risk Management Agency (<http://www.rma.usda.gov/>); SIIA: Sistema Integrado de Información Agropecuaria (<http://www.siiia.gob.ar/>).

<sup>a</sup> See Sections 2.2.2 and 2.5.2, respectively, on weather and actual yield data quality.

<sup>b</sup> Includes information on (or to estimate) dominant crop sequences and their relative proportion, sowing date, plant density, and cultivar maturity.

<sup>c</sup> Average size of administrative units located within the major maize production areas in each region.

implemented by the Global Yield Gap Atlas ([www.yieldgap.org](http://www.yieldgap.org)) include:

- (i) preference for using measured instead of estimated or interpolated data,
- (ii) transparency, reproducibility, and consistency in data selection,
- (iii) use of local expertise to corroborate data inputs (and collect them if necessary), and to ensure agronomic relevance, and
- (iv) strong preference for publicly accessible data.

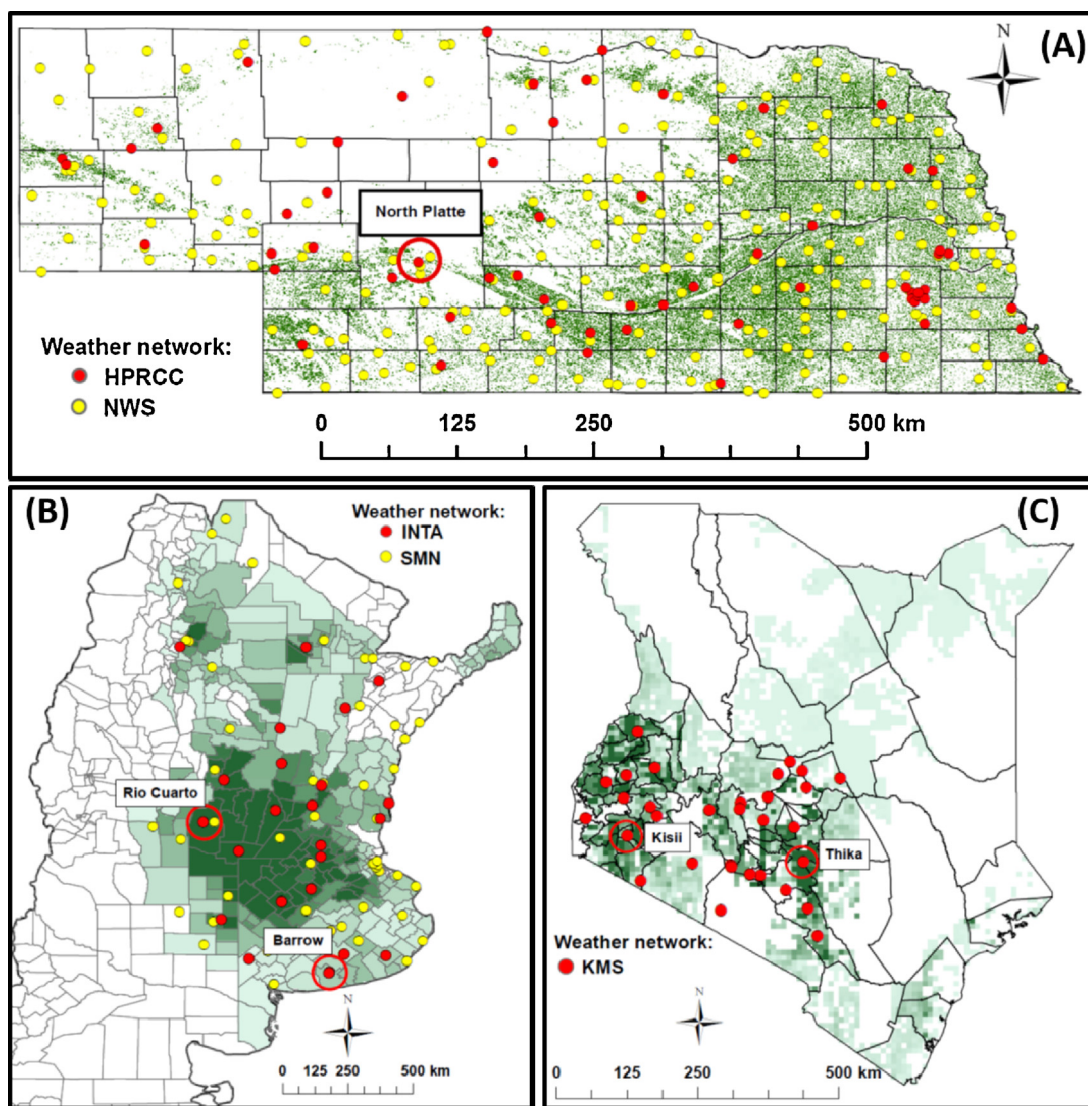
The methodology developed by the Global Yield Gap Atlas consists of a tiered approach, for each data-input type (*i.e.*, weather, cropping system, soil, Ya, and model calibration), which first defines the ‘ideal’ database for yield-gap analysis followed by “second- or third-choice” alternatives for cases in which the preferred data source does not exist or is not available. In fact, few countries or regions have good quality data at the fine degree of spatial resolution required for highly reliable yield gap analysis. Given this situation, we evaluate rainfed maize yield gaps in Nebraska (USA), Argentina, and Kenya to illustrate how to deal with a wide range of data quality and availability (Table 1, Fig. 1).

## 2.2. Weather data: The foundation for reliable crop simulation

### 2.2.1. How many years of weather data are needed?

Daily weather data of sufficient quantity and quality are required for robust simulation of  $Y_p$  and  $Y_w$  and their temporal variability (quantified by the coefficient of variation [CV]). A key question is how many years of weather data are needed to obtain a robust estimate of  $Y_p$ ,  $Y_w$ , and  $Y_g$  in order to account for year-to-year variation in weather. The answer depends on location and water regime. This is illustrated by looking at the range of possible

$Y_p$  and  $Y_w$  estimates, simulated based on different number of years of weather data, for rainfed and irrigated maize at North Platte (Nebraska, USA) and rainfed maize in Rio Cuarto and Barrow (favourable and harsh rainfed crop environments in Argentina, respectively) (Figs. 1 and 2, see details on model simulations in Appendix A). Simulations were performed using crop models that have been successfully validated on their ability to reproduce yields measured under optimal management conditions in each of the regions (see Section 2.6.4). Whereas rainfall is relatively low and highly variable at both North Platte and Barrow, the latter has soils in which a caliche layer limits the rootable soil depth. The sites can be categorized according to their average yield and inter-annual variation as follows: irrigated maize at North Platte and rainfed maize at Rio Cuarto (highest yield, lowest CV) and rainfed maize at North Platte and Barrow (lowest yield, highest CV). In favourable environments, 10 years of weather data are sufficient to estimate an average yield and CV that are within  $\pm 10\%$  of the estimates obtained with the entire 30-year database (*e.g.*, North Platte with irrigation and rainfed at Rio Cuarto) (Fig. 2). The number of required years increases to 15 to 20 years in less favourable environments (rainfed maize at North Platte and Barrow). Hence, depending upon water supply, 10 (irrigated or favourable rainfed environments) to 20 years of daily weather data (harsh rainfed environments) are needed for reliable estimates of  $Y_p$  (irrigated) or  $Y_w$  (rainfed) and their variability. These findings are consistent with Van Wart et al. (2013c), who showed that 6 to 15 years of weather data are required for reliable estimates of  $Y_w$  across an east-west transect in the U.S. Corn Belt where total rainfall, during the maize crop growing season, decreases from 900 mm (east) to 400 mm (west). Therefore, the number of available years of observed weather data shown in Table 1 seems sufficient for robust estimation of  $Y_p$  and  $Y_w$  in Nebraska and Argentina (>20 yr) but is probably insufficient for many locations



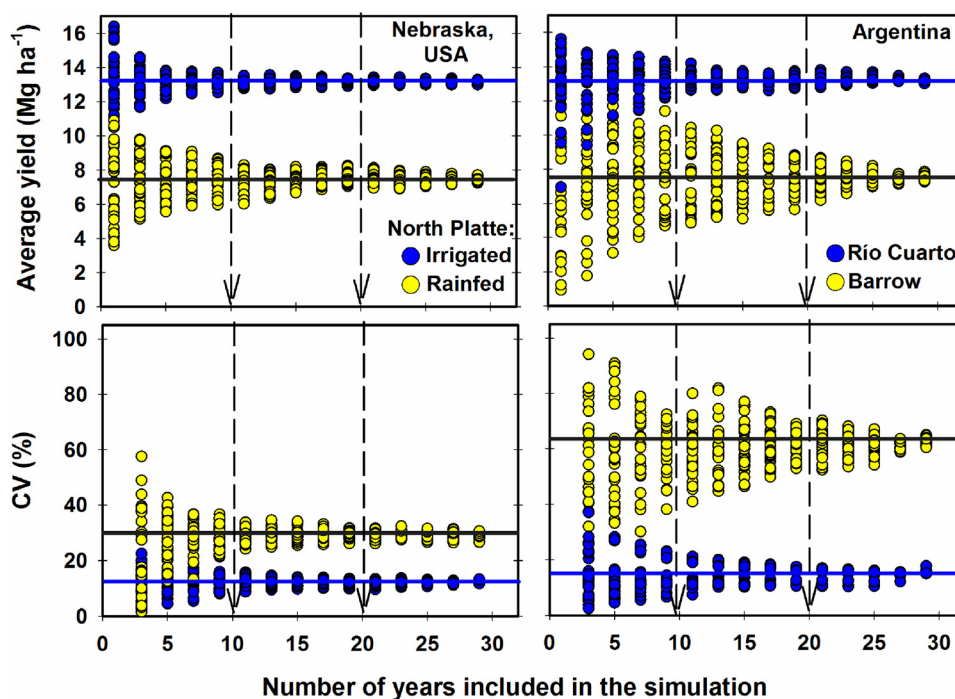
**Fig. 1.** Maps of Nebraska, USA (A), Argentina (B), and Kenya (C). Note scale differences among panels. Green intensity indicates maize harvested area density retrieved from USDA-NASS (Nebraska, USA), Ministry of Agriculture-SIIA (Argentina), and global SPAM maps (Kenya; You et al., 2014). Dots indicate locations of meteorological stations with  $\geq 3$  years of daily weather data situated within the major maize producing regions in each country. Lines indicate the boundaries of administrative units at which actual yield data are available: county (Nebraska, USA), department (Argentina), and district (Kenya). Meteorological stations used for specific analyses in the present article are circled and their names are shown. Meteorological weather networks are High Plains Regional Climate Center (HPRCC); US National Weather Service (NWS); Instituto Nacional de Tecnología Agropecuaria, Sistemas de Información Clima y Agua (INTA-SIGA); Argentina National Meteorological Service (SMN), Kenya Meteorological Service (KMS).

in Kenya (3 to 18 years depending upon location) where rainfall is low and highly variable. Use of insufficient number of years can bias estimates of  $Y_w$  due to inclusion of extreme weather years or short-term climate cycles in the weather data time series.

### 2.2.2. Required weather variables for crop modelling and data quality

Daily incident solar radiation and temperature (maximum [ $T_{max}$ ] and minimum [ $T_{min}$ ]) are required for estimating  $Y_p$ , whereas estimation of  $Y_w$  also requires precipitation. Depending on the method used to estimate reference evapotranspiration ( $ET_0$ ) in the simulation model, vapour pressure and wind speed may also be needed. Although measured data are always preferable to propagated or derived weather data, daily data for the other variables required for crop modelling besides  $T_{max}$ ,  $T_{min}$ , and precipitation (*i.e.*, solar radiation, vapour pressure) can be estimated, in absence of measured data, with a reasonable degree of accuracy using temperature data or retrieved from other data sources. An exception is wind speed, which cannot readily be

estimated from other variables, hence, a default world average value of  $2 \text{ m s}^{-1}$  is typically used to estimate  $ET_0$  when measured wind speed data are not available (Allen et al., 1998). In contrast, solar radiation can be estimated using equations that rely on sunshine hours (*e.g.*, Angstrom formula) or temperature (*e.g.*, Hargreaves formula) (Allen et al., 1998). Likewise, in regions with relatively level topography and little air pollution, gridded solar radiation reported by The Prediction of Worldwide Energy Resource (POWER) dataset from the National Aeronautics and Space Administration (<http://power.larc.nasa.gov/>), hereafter called NASA-POWER, can be used with confidence for crop simulation (Bai et al., 2010; White et al., 2011; Van Wart et al., 2013a,c). Vapour pressure is typically derived from relative humidity or dew point temperature measurements. In absence of measured data, vapour pressure can be estimated from the measured  $T_{min}$  assuming that dew point temperature is near the daily  $T_{min}$  (Allen et al., 1998). In all cases, it is desirable to locally validate these approaches using good quality observed data from a representative subset of years and locations in the region of interest.



**Fig. 2.** Average simulated maize yield potential and its temporal variability (estimated by the coefficient of variation [CV]) as a function of the number of years of weather data used in the simulations. Simulations were performed for favourable (blue symbols) and unfavourable environments (yellow symbols) for maize production in Nebraska (USA) and Argentina. Water inputs from irrigation and rainfall decrease in this order: irrigated maize at North Platte > rainfed maize at Rio Cuarto > rainfed maize at North Platte  $\approx$  rainfed maize at Barrow. Soils were deep at North Platte and Rio Cuarto ( $\geq 1.5$  m) but shallower at Barrow (0.8–1.2 m). Simulations based on Hybrid-Maize (Nebraska) and CERES-Maize (Argentina) models using observed weather data and dominant soils and management in each location and water regime (see Table S1). The data points, for a given  $n_y$ , represent the average yield potential and CV values as calculated based on 30 subsets of  $n_y$  re-sampled from the 30-yr weather database. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Besides data availability, robustness of simulated  $Y_p$  and  $Y_w$  depends on the quality of measured data. Weather data quality can be evaluated by prevalence of suspicious and missing values. Quality control screening methods have been developed to identify suspicious values in weather datasets (e.g., Allen et al., 1998; Hubbard et al., 2005). As a general guideline, we define a year of weather data as suitable for direct use in crop models, when  $\geq 80\%$  of all data for  $T_{\max}$ ,  $T_{\min}$ , and precipitation are recorded and  $< 20$  consecutive days are missing or suspicious for  $T_{\max}$  and  $T_{\min}$ , and  $< 10$  consecutive days for precipitation. For countries and regions where the weather station network is relatively dense (e.g., in Nebraska, on average, there is one HPRCC and NWS meteorological station per 3180 and 860 km<sup>2</sup>, respectively), and each station has long-term daily weather records, a robust approach to quality control with regard to identification and replacement of suspicious values and filling of missing data, is by evaluating correlations among adjacent weather stations (e.g., Hubbard et al., 2005; You et al., 2008). Unfortunately, in many regions of the world weather station networks have coarser spatial and temporal coverages. In these cases, identification of suspicious values is more problematic. Linear interpolation can also be employed, to a certain extent, to fill-in missing or erroneous  $T_{\max}$  and  $T_{\min}$  data, while gridded precipitation data from NASA-POWER or the Tropical Rainfall Measuring Mission (TRMM, <http://trmm.gsfc.nasa.gov/>) can be used to fill-in missing days (although TRMM data are only available over the latitude band 50° N–S). An alternative for filling missing  $T_{\max}$  and  $T_{\min}$  is to use relationships between observed and gridded weather data based on a limited number of data-years to perform a location-specific correction of the latter and use these to fill in values for missing days (e.g., Chaney et al., 2014; Van Wart et al., 2015).

Two other factors influence quality of weather data for agricultural assessments. The first is the degree to which the location of a selected weather station is representative of the surrounding

agricultural land on which the simulated crop is grown. Solar radiation,  $T_{\max}$ , and  $T_{\min}$  can be biased by topography, water bodies, surrounding vegetation, and urban areas. For agricultural applications, weather data should ideally be measured at meteorological stations situated in a rural setting surrounded by agricultural land (e.g., HPRCC and INTA weather networks in Nebraska and Argentina). Still, observed weather data from stations located in cities or airports are preferable to gridded weather data (see Van Wart et al., 2013a). Second, crop modelling to represent weather, soil, current crop management and cropping systems should use weather data from recent decades (preferably last 2–3 decades) because data from previous decades may not be representative of current climate where there have been significant changes in weather due to climate change (e.g., Kassie et al., 2014; Rurinda, 2014).

### 2.2.3. Selection of weather data sources

Selection of sources of weather data is based on the goal of using as much observed weather data as possible while reaching the minimum number of years required for robust estimates of  $Y_p$  or  $Y_w$  and their variability (Fig. 2). In many parts of the world, weather data availability and quality are far from optimal for some or all required weather variables. Hence, our protocol follows a tiered approach (Fig. 3) in which the focus shifts from the ideal scenario towards acquisition of the minimally required weather variables for the simulation (i.e.,  $T_{\max}$ ,  $T_{\min}$ , and precipitation) as data quality and availability become limiting. To this end, three levels of weather data availability are defined:

- Level 1: suitable weather data available for  $> 10$  years, preferably from recent decades to avoid misleading effects of climate change. While we recognize that 10–15 years of data may still be insufficient for a robust estimate of  $Y_w$  and its variability in semi-arid

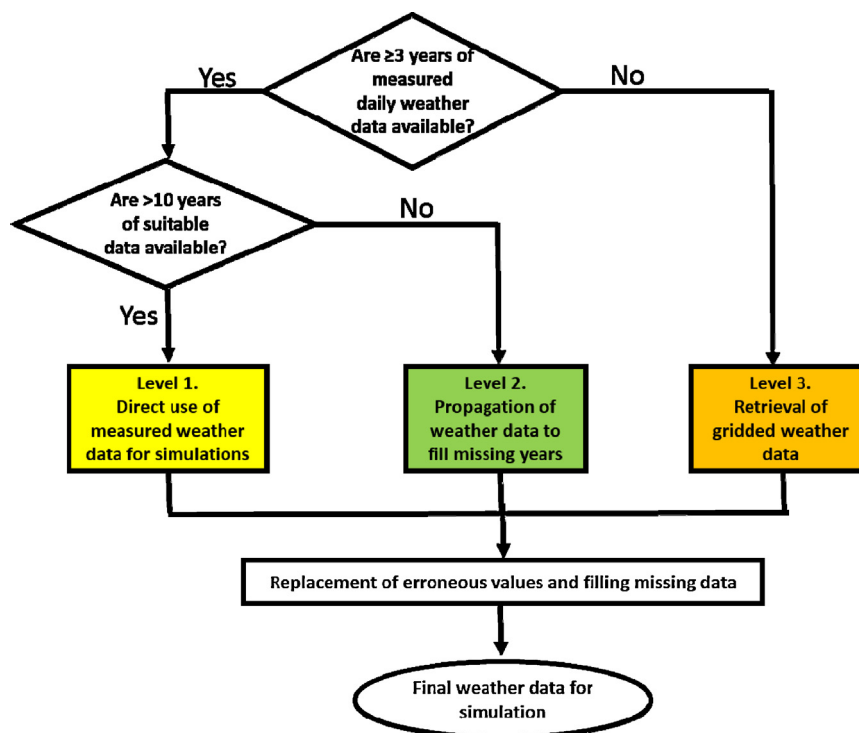


Fig. 3. Flow chart for selection of weather data for crop simulation modelling as used in the Global Yield Gap Atlas ([www.yieldgap.org](http://www.yieldgap.org)).

environments, this is still superior to use of propagated weather data or gridded weather databases (Van Wart et al., 2013a, 2015).

- Level 2: suitable weather data available for  $\leq 10$  years. In these cases, the best option is to use the existing weather data and to generate the missing years of data following the methodology described by Van Wart et al. (2015), to obtain a minimum of 15–20 years of weather data. Briefly, this method consists of (a) correcting long-term, daily NASA-POWER  $T_{\max}$  and  $T_{\min}$  values on the basis of, at least, 3 years of observed  $T_{\max}$  and  $T_{\min}$  data and (b) retrieving precipitation data from TRMM or NASA-POWER databases.
- Level 3: suitable weather data are available for  $< 3$  years or do not exist at all. In this case the only option is to use gridded or generated weather databases; however, resulting simulations need to be flagged as less reliable than Yw or Yp estimates based on observed weather data and updated, when observed weather data become available for the targeted area. It is difficult, however, to recommend the best gridded weather database to use, because, without site-specific correction, all of them appear to have substantial biases when compared against measured weather data, and the biases are not consistent in sign and magnitude across locations (Mearns et al., 2001; Baron et al., 2005; van Bussel et al., 2011; Van Wart et al., 2013a, 2015).

#### 2.2.4. Selection of weather data for the three case study areas

The three countries shown in Table 1 illustrate how the protocol can be applied across the spectrum of data availability. Nebraska approaches the 'ideal' scenario, where all required weather variables are measured and available from HPRCC meteorological stations located on agricultural land, with a sufficient number of locations and years, and data are subjected to robust measures of quality control. Argentina deviated from the ideal condition because (i) solar radiation data are not available, (ii) some meteorological stations are located in airports or cities (those belonging to the SMN network), and data quality is an issue for

some locations or time periods. Solar radiation was retrieved from the NASA-POWER database, which was evaluated against measured solar radiation for a subset of location-years (total of 18,375 daily observations), showing remarkably good agreement (root mean square error:  $3.5 \text{ MJ m}^{-2} \text{ d}^{-1}$ ,  $r^2 = 0.84$ ). To comply with quality standards, all daily observations for each variable were screened by looking at correlations between the selected weather station and the two adjacent stations following the method described by Van Wart et al. (2013c). In contrast, almost all meteorological stations in Kenya were located at airports or in cities and did not have suitable data for a sufficient number of years ( $< 10$  years). For those targeted areas where  $\geq 3$  years were available (but less than 10), the propagation technique developed by Van Wart et al. (2015) was applied to produce long-term weather data ( $\geq 10$  years), keeping all observed data within the dataset and only using propagated data for missing time periods. NASA-POWER was used as source of solar radiation data and also to estimate humidity from dew point temperature. For those targeted areas that have  $< 3$  years of data or no data at all, NASA-POWER weather data for all variables were used without correction, but results were flagged as highly suspicious given the uncertainty in weather data quality for the site in question.

### 2.3. Cropping-system context

#### 2.3.1. What is the cropping-system context?

Specification of dominant water regimes (i.e., rainfed, fully-irrigated, or partially-irrigated), crop sequence(s), and their proportion of total harvested crop area, are essential for accurate estimation of Yp, Yw and Yg at local to national scales. Explicit quantitative accounting of this cropping system context is especially important where rainfed and irrigated crops co-exist within the same geographic area and where the climate allows 2–3 crop cycles per year on the same field. Likewise, the same crop can be grown in very different crop sequences so that Yp (or Yw) differs depending on sequence. Each water regime and cropping system is defined

by average sowing date<sup>2</sup>, cultivar maturity (growing degree days or, when not available, time duration from sowing to physiological maturity), and plant density (number of plants per ha). Stored soil water at sowing in the root zone also needs to be specified for rainfed or partially-irrigated cropping systems (see Section 2.6.4). For each water regime, separate Yp (or Yw) are simulated for each crop cycle and, if there is more than one cycle, a weighted average is estimated based on the relative proportion of total harvested crop area of each cycle. A similar approach is followed when the same crop is grown in different crop sequences. This aggregation is needed because Ya data are typically reported on a per-harvested hectare basis, without disaggregation by crop cycle (see Section 2.5.1).

### 2.3.2. Sources of error associated with cropping system data

In many cropping systems, availability of machinery and labour constrain timely crop sowing, and plant density is sometimes sub-optimal due to high seed cost or manual sowing. In cases in which there is a clear indication that sowing date or plant density are sub-optimal, it is useful to distinguish between simulations based on actual management *versus* those using 'optimal' management and provide a justification for the latter. In all cases, the 'optimal' management scenario must be constrained within the boundaries imposed by the crop sequence under the assumption that, in general, farmers are efficient in allocation of land, labour, and time within the limitations imposed by existing economic and biophysical environments (Herdt and Mandac, 1981; Hopper, 1965; Sheriff, 2005).

Because breeding efforts for most crops have improved yields and yield stability over the past 30 years (Connor et al., 2011; Fischer et al., 2014), simulations of Yp and Yw should be based on recently released high-yielding crop cultivars, grown in pure stands, widely used by farmers in the region. Ideally, it is desirable to have cultivar maturity reported in growing-degree days (GDD) from sowing to maturity, preferably also the GDD from sowing-to-flowering, and, for those cultivars in which development is also modulated by photoperiod and vernalisation requirements, to have all the cultivar-specific parameters that account for the developmental responses to these two factors. In developed countries, this information is sometimes available through seed catalogues published or provided on websites by seed companies or from public-sector cultivar testing programs. In most developing countries, however, the only indicator of cultivar maturity is average crop cycle duration, that is, the number of days 'typically' required for a crop at a specific location to reach physiological maturity. A backwards procedure can be followed in these cases to derive cultivar GDD by running long-term simulations and adjusting phenology-related coefficients until simulations reproduce the reported average date of physiological maturity. When this approach is used, estimated Yp or Yw can still be biased due to uncertainties in the simulated flowering date, or when crop cycle duration is based on the date of harvest instead of physiological maturity (e.g., Bagley et al., 2012). For example, in large-scale, mechanized commercial farming, harvest takes place when grain moisture content reaches a level at which mechanical harvest is possible and drying costs are minimized. Hence, in some cases, harvest can take place up to 4 weeks after the crop has reached physiological maturity. By contrast, in small scale, non-mechanized farming in tropical and semi-tropical regions, reported harvest date is typically much closer to physiological maturity due to the value of crop residues for livestock feeding, risk of yield losses due to insects, diseases, birds, and rodents, and opportunities to plant subsequent crops in the same year. Using

<sup>2</sup> Average sowing date is defined as the approximate calendar date at which 50% of the final sown hectare is complete.

**Table 2**

Cropping-system context in the three cases of study presented in this study.

Region	Cropping system feature	
	Water regime	Crop intensity (maize crops yr <sup>-1</sup> )
Nebraska (USA)	Irrigated & rainfed	One
Argentina	Rainfed	One
Kenya	Rainfed	One (east Kenya) or two (west Kenya)

maturities longer than those used by producers typically leads to unrealistically high Yp in irrigated systems or Yw in favourable rainfed environments while Yw can be unrealistically low and variable at locations with severe terminal water deficit.

### 2.3.3. Cropping system data used for the three case studies

Differences in cropping systems are illustrated for the three case studies (Table 2). In Nebraska, irrigated and rainfed maize co-exist (with 60 and 40% of total harvested area, respectively) and a separate set of management practices, in particular plant population density, is required for each water regime. In contrast, maize area under irrigation in Argentina and Kenya is negligible (<3% of total maize harvested area). Whereas only one annual maize crop is grown in Nebraska and Argentina, typically in a 2-y rotation with soybean, two maize crops are grown in the same field each year at many locations in west Kenya where a bi-modal annual rainfall pattern occurs. Hence, separate specification of management practices for each maize crop cycle was needed for these locations in Kenya for an accurate simulation of Yp or Yw. Resulting Yp and Yw needs to be averaged, weighted by their relative area, as explained in Section 2.3.1.

In all three case studies, the required cropping-system information was not readily available, except for sowing date data in Nebraska. Data on sowing date progress are annually collected for major U.S. crops, on a county basis, by the Risk Management Agency (<http://www.rma.usda.gov/>). While this information is collected for insurance purposes, it also provides an objective way to define the range of sowing dates at an adequate spatial resolution. In contrast, data on dominant cultivar and plant density, for each water regime, are not publicly available and simulations rely on expert opinion from local agronomists and information provided by seed dealers and seed companies. Once the dominant cultivar is determined, the GDD from emergence to flowering and from flowering to physiological maturity can be retrieved from private seed company catalogues and information available on their websites. In Argentina, accurate information on dominant cultivar, sowing dates, and recommended plant population densities were obtained from local agronomists working in each of the targeted areas. GDD of dominant cultivars was available through seed companies and confirmed with detailed phenological observations in research station experiments (Monzon et al., 2012). All management data in Kenya were collected from local collaborators but, in contrast to Nebraska (USA) and Argentina, wide ranges were reported (e.g., a 2-month window for sowing date), reflecting important variation in management practices across years and farms due to variation in timing of rainfall at the beginning of the rainy season.

## 2.4. Soil data

### 2.4.1. Selection of dominant soil types

The present paper does not attempt to provide a review of the available data sources or different approaches to obtain soil input data required by each crop model. Readers are referred to papers that consider different approaches for obtaining adequate soil data for crop yield simulations (e.g., Ritchie et al., 1990; Gijssman et al.,

2007; Batjes, 2012; Romero et al., 2012). Instead, our aim is to develop scientifically justifiable and efficient protocols for selecting the most widely used soils for production of a given crop at a specific location, and then specifying the soil properties for those soils that are required for crop modelling (hereafter called 'functional' soil properties). Soil mapping units and soil series were used as the basis for deriving required soil properties. A soil map unit is a collection of areas grouped according to landscape position, profile characteristics, relationships between these two, suitability for various uses, and need for particular types of management such as soil erosion control practices. Each soil map unit may be composed of one or more soil series. It is important to define the dominant soil series that are most widely used for production of the targeted crop in the area of interest (Van Bussel et al., 2015). To avoid biases due to inclusion of soil units not relevant for crop production, soils with negligible crop area (*i.e.*, <10% coverage of crop harvested area in the area of interest) or those where sustainable long-term annual crop production is not likely such as shallow soils (rootable depth <0.5 m), sandy soils (PASW <7 cm cm<sup>-3</sup> or sand content >75%), and soil series with very steep terrain (slope >10%) are excluded. Moreover, all else being equal, farmers have a preference for growing certain crops on the best soils, as it is the case of maize in Argentina and the USA.

#### 2.4.2. Required soil variables for crop modelling

While soil input data required by different crop simulation models to simulate Yw may differ to some extent, all such models require rootable soil depth and volumetric plant-available soil water holding capacity (PASW; in cm<sup>3</sup> cm<sup>-3</sup>). Hence, soil profile data should include these 'functional' soil properties (*e.g.*, soil water retention limits) or, at least, data from which these can be derived (*e.g.*, soil texture class). Other soil and terrain attributes such as slope and drainage class are also needed to determine the amount of surface runoff. An accurate simulation of surface runoff requires a level of model precision and data detail that current data availability does not allow in most countries, hence, semi-empirical approaches for runoff estimation are acceptable (*e.g.*, Soil Conservation System (SCS), 1972; Campbell and Diaz, 1988).

Besides soil water holding capacity, rootable soil depth is the most important soil property influencing Yw and its year-to-year variability (*e.g.*, Sadras and Calvino, 2001). The rootable soil depth is defined as the soil depth that can be effectively explored by the crop root system to absorb water and nutrients without severe physical or chemical constraints to root growth or functionality. Root growth restrictions include bedrock, caliche layer, abrupt textural change, alkalinity, sodicity, acidity, etc. (USDA-NRCS National Soil Survey Handbook). Even in absence of these constraints, there is a limit to the rootable soil depth defined by crop genotype and length of the crop season. For most grain crop species in rainfed systems, a value of ≈1.5 m can be assumed for soils without physical or chemical limitations (*e.g.*, Dardanelli et al., 1997). Although data needed to define the rootable soil depth can be retrieved from soil series descriptions, in many cases soil data are limited to the topsoil and it is not clear if the sampling depth can be taken as a proxy for the rootable soil depth. In absence of this information, determination of rootable depth must rely on local experts though, based on our own experience in the Global Yield Gap Atlas, knowledge about subsoil properties is generally poor in many countries and should be used with caution.

The other mandatory variable for simulating Yw is plant available soil water (PASW) as determined by upper and lower soil limits for water retention (*i.e.*, field capacity and permanent wilting point, respectively, which correspond roughly to a suction of -33 and -1500 kPa). Actual measurements of soil water retention limits are rarely available, hence, these are typically estimated using pedo-transfer functions (PTF) based on soil texture. Many PTFs

are available to derive soil moisture limits as discussed by Tietje and Tapkenhinrichs (1993), Rawls et al. (1991), and Gijsman et al. (2002). An important, though often overlooked consideration when using a PTF is that the range of soil texture and clay mineralogy of the targeted areas should be within the range of validity of the PTF. In particular, PTFs developed for temperate soils (*e.g.*, Saxton and Rawls, 2006) should not be used for estimating water retention limits in strongly weathered tropical soils (Tomasella et al., 2000; Hodnett and Tomasella, 2002).

The potential degree of error due to incorrect specification of PASW and rootable depth is illustrated for two locations in Kenya, which represent favourable (second-season crop at Kisii) and harsh (single-season crop at Thika) rainfed crop environments, and for North Platte, USA (Figs. 1 and 4). Maize Yw was simulated using (i) generic soil water retention limits reported for each textural class by Driessen and Konijn (1992) for temperate soils *versus* values estimated from a PTF developed for tropical soils (Hodnett and Tomasella, 2002) and (ii) rootable depth of 1 m *versus* 1.5 m (Fig. 4). Average Yw and its CV vary greatly among combinations of PTF × soil rootable depth. For example, average Yw ranged from 8.7 to 10.8 Mg ha<sup>-1</sup> at Kisii, with CV ranging from 24% to 42%. These ranges were relatively much wider at North Platte, where Yw ranged from 3.4 to 6.3 Mg ha<sup>-1</sup>, with CV ranging from 18 to 91%.

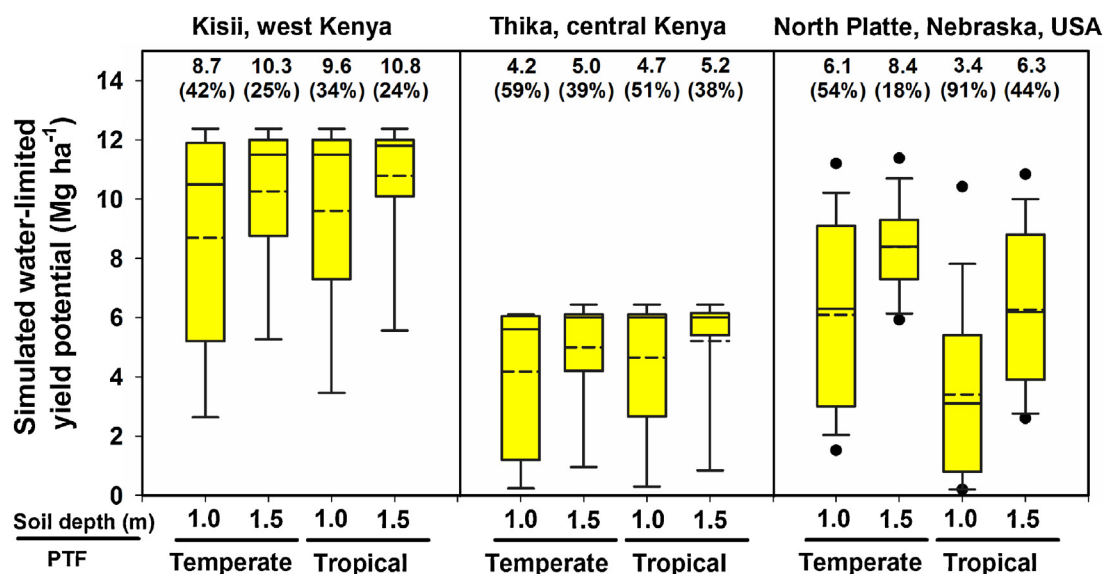
#### 2.4.3. Soil data retrieval for the three case studies

Soil data sources for the three case studies include: detailed national soil maps and profile databases in Nebraska and Argentina and the ISRIC-WISE (Batjes, 2012) global soil database for Kenya (Table 1). For Nebraska and Argentina, relevant soil types for crop production in the targeted areas can be easily identified and information to determine the rootable soil depth and PASW is available. For Kenya, the ISRIC-WISE global soil database was selected because this database provides the required information on soil properties for crop modelling. Sources of uncertainty when using ISRIC-WISE (and other global soil databases) include: (i) difficulty to determine which soil units are relevant for crop production, (ii) little available data on rootable soil depth, and (iii) uncertainty about selection of an appropriate, well-calibrated PTF for tropical soils. Relevant soil units for crop production in the targeted areas of Kenya were selected following the rules for soil type selection described in Section 2.4.1, together with information on crop harvested area distribution from SPAM (You et al., 2009, 2014). PTFs derived for tropical soils by Hodnett and Tomasella (2002) were used to estimate soil water retention limits based on the reported soil texture. Due to lack of data on rootable soil depth, a standard 1-m depth was used for all Yw simulations based on observations of Nye and Greenland (1960) about savannah soils in Sub-Saharan Africa.

#### 2.5. Actual yield: Often a bottleneck for estimating yield gaps

Actual yield is defined as the average annual yield obtained by farmers in a geographic area for a given crop with a given water regime. There are four key aspects related to Ya data: (i) level of disaggregation by crop and water regime, (ii) number of available data-years, (iii) spatial resolution, and (iv) data quality. Another important, though often overlooked aspect, is the dry matter concentration at which Ya values are reported so that the Ya and Yw (or Yp) data used for calculation of Yg are at equivalent moisture content. For example, the most widely used database for retrieving national Ya averages (FAOSTAT, FAO, 2014; <http://faostat.fao.org/>) does not explicitly define the moisture content at which crop yields are reported. In contrast, grain yields reported by government agencies in the USA and Argentina are provided at standard moisture content (*e.g.*, ca. 15% for maize grain).





**Fig. 4.** Box plots of simulated maize water-limited yield potential at Kisii and Thika (Kenya) and North Platte, Nebraska (USA) using Hybrid-Maize model based on 1 m and 1.5 m rootable depth and soil water limits retrieved from [Driessen and Konijn \(1992\)](#) for temperate soils versus values estimated using pedo-transfer functions (PTF) for tropical soils ([Hodnett and Tomasella, 2002](#)). Lower and upper boundaries for each box are the 25th and 75th percentiles. The solid and dashed lines inside each box indicate the median and mean, respectively. Whiskers (error bars) above and below the box indicate the 90th and 10th percentiles. Dots above and below the whiskers indicate the 95th and 5th percentiles. The means over years and the inter-annual coefficient of variation (in %) are also presented above the bars. Simulations were performed using local weather and management data. At Kisii, simulations were performed only for second-season maize sown on Sept 9. Clay and silt loam soils were used for the sites in Kenya and Nebraska, respectively.

#### 2.5.1. Level of disaggregation and number of available data years

Actual yields need to be disaggregated by water regime wherever irrigated and rainfed crop systems coexist within the same geographic area. Likewise, in multiple cropping systems where 2 or more cycles of the same crop are grown on the same field each year or the same crop can be grown in very different crop sequences so that Yw differs depending on sequence, it is preferred to have separate Ya estimates for each crop cycle and sequence, which allows estimating separate Yg values. With few exceptions, however, Ya data are reported on an aggregated harvested-area basis, without disaggregating Ya by crop cycle or sequence. Hence, mean Yg is estimated as the difference between the long-term weighted averages of Yp (or Yw) and Ya, both expressed on a per-harvested hectare basis (see Section 2.3.1).

The number of years of Ya data to calculate average Ya should be determined on a case-by-case basis, following the principle of including as many recent years of Ya data as possible, to account for weather variability but not climate change, while avoiding the bias due to a technological time-trend ([Van Ittersum et al., 2013](#)). Likewise, the years of Ya data should be within the range of years for which Yw (or Yp) was simulated. As a general guideline for data-rich countries that show a steep yield trend (or trend break), we recommend using the Ya reported for the 5 most recent years for the calculation of average yield; if there is no trend, the Ya reported for the most recent 10 years can be used. However, this approach cannot be followed in data-poor countries where long-term yield statistics are not available. For these cases, we recommend a minimum of 5 recent years of Ya data (3–4 years are acceptable if more years are not available), recognizing that this may not be sufficient to account for year-to-year variability in Ya due to weather, especially in harsh rainfed environments.

#### 2.5.2. Actual yield source, spatial resolution, and data quality

Ideally, Ya should be based on yield statistics available for sub-national administrative units such as municipalities, counties, departments, sub-districts, districts, or provinces. Ultimately, the location and extent of the administrative unit should be

(reasonably) congruent with the location and spatial extent of the targeted area for yield gap analysis. If two or more administrative units (or parts of them) are located within the targeted area, a weighted average yield can be estimated based on their relative area-basis coverage. Ya can also be estimated from values reported for larger administrative units such as regions, provinces, and states but resulting Ya estimates need to be flagged (and eventually replaced by more spatially granular estimates) because yield reported at a coarse level of spatial resolution may not be representative of the Ya of the targeted area, when the latter is smaller than the area reporting Ya.

In many cases, Ya data can be accessed directly through national statistics bureaus websites ([Table 1](#)), FAO/IFPRI/SAGRE agromaps ([FAO/IFPRI/SAGRE, 2006](#); <http://kids.fao.org/agromaps/>), CountrySTAT (<http://www.countrystat.org/>), Eurostat (<http://epp.eurostat.ec.europa.eu/portal/page/portal/agriculture/data/database>), or retrieved by agronomists from their local statistical bureaus or institutions. A viable alternative, when national statistics at an appropriate level of spatial resolution do not exist or are unreliable, is to estimate Ya from existing data collected through farm surveys and by local agronomists administered by national agricultural research institutions, universities, CGIAR centers, [World Bank](#) (LSMS), private sector, or other on-going projects such as TAPRA survey panel (<http://www.tegemo.org/index.php/component/k2/item/258-tapra-ii-household-panel-survey-coverage>). Spatial coverage of the survey should be consistent with the targeted area and include five years of data to account for weather variability (again, 3–4 years are acceptable if no more years are available). Another source of yield data is from on-farm experiments that include a treatment that follows local 'farmer practices' over several years (e.g., [Tittonell et al., 2008](#); [Fermont et al., 2009](#); [Wairegi et al., 2010](#)) or producer self-reported data (e.g., [Grassini et al., 2014](#)). These sources of data can be useful to determine Ya as long as the farms where the studies were conducted are representative of the population of farms within the area of study. If no yield data are available at any sub-national level or through survey or field trial data, Ya can be based on local

knowledge (local agronomists, agricultural input or seed dealers, or others engaged in businesses that deal directly with producers). The aim would be to estimate average  $Y_a$  in the most recent past 3-year period (preferably longer) with the goal of replacing these estimates with official statistics when these become available.

Determining the degree of uncertainty related to the accuracy of  $Y_a$  data is an important component of yield gap assessment. Whereas it is not feasible to survey most farms within a region or year in a cost-effective way, comparison of  $Y_a$  using several independent data sources, for the same region-year, can be used to assess the  $Y_a$  data uncertainty. This comparison does not determine which data source is more accurate, but a substantial difference in estimates of  $Y_a$  among data sources provides insight about the uncertainty in  $Y_a$  and  $Y_g$ . Unfortunately, there are only very few examples of verification of  $Y_a$  estimates using truly independent datasets (Sadras et al., 2014 and references cited therein). These previous studies have shown that estimates of  $Y_a$  from different data sources are similar or markedly different, depending upon the crop/country in question, with the magnitude of the  $Y_a$  mismatch also varying across years and regions within the same country. In other published studies aiming at assessing quality of gridded  $Y_a$  data, the comparison is not valid because the databases compared were derived from the same underpinning  $Y_a$  data, resulting in a misleading assessment about the quality of  $Y_a$  (e.g., Iizumi et al., 2013).

### 2.5.3. Actual yield sources and quality-control for the three study cases

Availability and quality of  $Y_a$  markedly differed among the three case studies (Table 1, Figs. 1 and 5). In Nebraska, long-term (>30 years) annual  $Y_a$  data were available through USDA-NASS ([www.nass.usda.gov/](http://www.nass.usda.gov/)) for each water regime and county (roughly 2000 km<sup>2</sup> or a circle with radius of ca. 25 km). Comparison of  $Y_a$  data reported by USDA-NASS against  $Y_a$  data independently collected through the Nebraska Natural Resources Districts (<http://www.nrdnet.org/>) indicated an overall difference of 0.6 Mg ha<sup>-1</sup>, which represented only 6% of average yield calculated using both data sources, so, there is confidence in the reported  $Y_a$  data. Data availability was similar in Argentina though at a coarser spatial aggregation (roughly 4500 km<sup>2</sup>, i.e., a circle around a location with radius of ca. 38 km) and average mismatch between independent  $Y_a$  data sources represented 14% of the yield mean (though relatively large differences >15% were found for 33% of region-years). Finally, though the spatial resolution of the  $Y_a$  data in Kenya was acceptable, only a limited number of years of  $Y_a$  data were available (Table 1) and time periods were not consistent across locations. Also notable was a large discrepancy between two sources (45% of  $Y_a$  mean), though discrepancy was small in absolute values due to very low average farm yield levels (Fig. 5).

### 2.6. Model calibration and long-term simulations of yield potential

#### 2.6.1. Selection of crop simulation model

Desirable attributes of crop simulation models were summarized by Van Ittersum et al. (2013) and are not addressed in this paper. Like Van Ittersum et al. (2013), we argue against using a single generic model globally because it is more important that the model used has been calibrated and evaluated for the conditions to be simulated. Thus, models may differ for the same crop in different regions or countries, and for different crops, as long as the models used have been calibrated under those conditions of the targeted areas. Preferably, the same model should be used for the same crop to simulate  $Y_w$  and  $Y_p$  at locations that are then aggregated to give estimates at larger spatial scales (Van Bussel et al., 2015). We also argue that it is preferable to use one (or few)

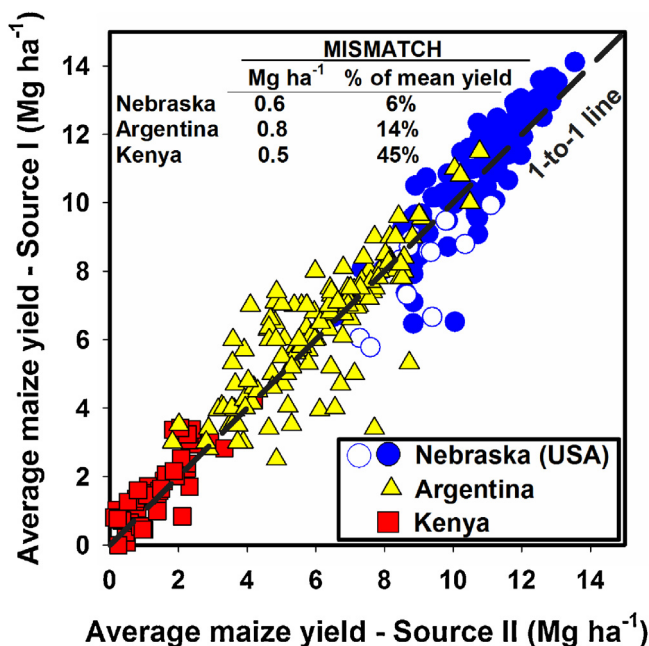


Fig. 5. Comparison between two independent sources of actual grain yield for maize in Nebraska, USA (blue circles), Argentina (yellow triangles), and Kenya (red squares). Data from Nebraska include both rainfed and irrigated crops (open and solid circles, respectively). Yield sources I and II are, in this order, Natural Resources Districts ([www.nrdnet.org/](http://www.nrdnet.org/)) versus National Agricultural Statistics Service of the United States Department of Agriculture ([www.nass.usda.gov/](http://www.nass.usda.gov/)) in Nebraska (USA), Bolsa de Cereales de Buenos Aires ([www.bolcereales.com.ar/pas](http://www.bolcereales.com.ar/pas)) versus Ministerio de Agricultura, Ganadería y Pesca ([www.minagri.gov.ar/site/index.php](http://www.minagri.gov.ar/site/index.php)) in Argentina, and Tegemeo Institute ([www.tegemeo.org/](http://www.tegemeo.org/)) versus Ministry of Agriculture ([www.kilimo.go.ke](http://www.kilimo.go.ke)) in Kenya. Data are from 12 counties and at least 6 cropping seasons per county (Nebraska, USA), 13 regions and 9–10 cropping seasons per region (Argentina), and 47 districts and one season (2011) for Kenya. Average mismatch between data sources is shown (absolute value and as percentage of the mean of the two actual yield data sources) for each region. Data from Nebraska and Argentina have been adapted from Sadras et al. (2014). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

well-calibrated simulation models to estimate  $Y_p$  and  $Y_w$  than using ensembles of numerous, in many cases poorly-calibrated, models as proposed by others (e.g., Asseng et al., 2013; Rosenzweig et al., 2014; Challinor et al., 2014b). In fact, careful examination of this approach (i.e., ensembles) in recent publications shows that it can perform very poorly at specific locations (e.g., Martre et al., 2014). Indeed, a strong justification for using an ensemble of models, each developed for different purposes and few validated for the environmental conditions in question, has yet to be articulated. Likewise, most crop-modelling papers do not report data about model calibration within the targeted agro-ecological zones under study, and how the models perform in terms of reproducing  $Y_w$  and  $Y_p$  measured in well-managed experiments.

#### 2.6.2. Data for model calibration

Different crop cultivars are planted across locations, hence, it is necessary to calibrate crop models to account for differences in crop phenology and growth-related factors (Jones et al., 2003). A robust calibration requires estimates of  $Y_p$  or  $Y_w$  from high-yielding field experiments in which crops are grown without nutrient limitations or yield loss from biotic adversities (e.g., insects, disease, weeds), and where all required weather, soil, and management data are available to run the field-year specific simulations (see Appendix B in Supplemental information). Variety trials (if of proper plot size and with near-optimal management) are a good source of yield and phenology data as well. If such experiments are not available for a specific country or

region within a country, an alternative is to use crop growth data from experiments in which crops are grown with optimal management for analogous regions in terms of climate and soils. Ultimately, the goal should be to evaluate the ability of the model to reproduce major  $G \times E \times M$  interactions across a relevant range of potential yields.

If robust calibration is not possible due to lack of field studies in which crops were grown with near-optimal management, the methodology proposed by Van Wart et al. (2013c) can be used to calibrate the simulated crop phenology. Briefly, the model coefficients related to phenology can be adjusted until the simulated physiological maturity matches the typical date of physiological maturity reported within the targeted area (see Section 2.3) while growth-related coefficients can be based on generic model parameters reported in the literature or derived from previous modelling studies (e.g., van Heemst, 1988) or adjusted within limits as detailed in Appendix B.

### 2.6.3. Simulation of long-term yield potential and its variability

Simplification of the cropping-system features by averaging weather, soil or cropping-system data, typically results in biased results and a substantial reduction in agronomic relevance of Yp estimates (De Wit and Van Keulen, 1987). Therefore, the basic unit for a crop simulation is given by a combination of crop cycle (within a cropping system)  $\times$  soil type  $\times$  water regime  $\times$  year. These “simulation units” can then be aggregated to higher spatial scales and longer time periods by weighting for harvested crop area under each unit as previously described in Sections 2.3 and 2.4.

Once Yp and Yw are simulated for a given simulation unit, estimated values can be screened for inconsistencies or errors. The following quality-control measures can be applied to screen simulated yields:

- (i) years with Yp or Yw  $\leq Y_A$ ,
- (ii) Yw  $\approx$  Yp and Yw has a small CVs (<5%) in water-limited environments,
- (iii) Yp or Yw or harvest index estimates far beyond reported record yields,
- (iv) years with Yp or Yw  $\approx 0 \text{ Mg ha}^{-1}$ , and
- (v) simulated yields for particular locations/years that look ‘suspiciously’ lower or higher than in the rest of the sites/years.

Other approaches to derive Yp or Yw, such as boundary functions relating crop yield to water availability, can also be used to check suspicious values (e.g., French and Schultz, 1984). If any of the above cases are detected, underpinning weather, soil, management, and model parameters should be re-checked for the suspicious values as well as the value of Ya itself.

### 2.6.4. Model calibration and long-term simulations for the three case studies

The three examples presented in the paper portray well the range of conditions in data availability for model calibration and evaluation. Simulations of maize Yp and Yw in Nebraska and Kenya were performed using the Hybrid-Maize model (Yang et al., 2004). Model calibration was performed using high-quality data from experiments and high-yield producer fields in the U.S. Corn Belt where crops had been grown under near-optimal conditions (Yang et al., 2004). Model performance at reproducing yields in well-managed crops has been exhaustively evaluated across a wide range of environments in the U.S. Corn Belt, with measured yields ranging from 0.5 to 18  $\text{Mg ha}^{-1}$  along a wide range of water supplies (Yang et al., 2004; Grassini et al., 2009). In contrast, lack of high-quality experimental data in Kenya did not allow an independent evaluation of the Hybrid-Maize model and only phenology-related coefficients were modified to better represent

the crop cycle duration reported by local collaborators for the targeted areas. In Argentina, CERES-Maize (Jones and Kiniry, 1986), embedded in DSSAT v 4.5 (Jones et al., 2003), was used to estimate maize Yp and Yw. Model calibration was performed with detailed measurements from a number of well-managed rainfed and irrigated maize experiments (Monzon et al., 2007, 2012).

Available soil water content at sowing within the rootable soil depth can have a large impact on Yw, especially in harsh rainfed environments. Ideally, crop simulation models can be used to simulate the soil water balance during the entire crop rotation, including the non-growing season and this approach was followed for simulating the maize-soybean rotation in Argentina. However, it was not possible to follow this approach in Nebraska (USA) and Kenya because the Hybrid-Maize model does not simulate crop rotations. For these cases, the soil water balance was initialized over a period of time before the sowing date, beginning around physiological maturity of the previous crop in the rotation, assuming a typical low initial soil water content at end of the growing season of the previous crop of 50% of available soil water (or as estimated by expert opinion).

## 3. Discussion

### 3.1. Key principles

Robust protocols to support crop modelling and yield-gap analysis at a specific location are presented based on the lessons learned from establishing the Global Yield Gap Atlas ([www.yieldgap.org](http://www.yieldgap.org)). These methods were developed to be flexible enough to account for a wide range of data availability and quality, while ensuring minimum standards of data quality, agronomic relevance, and transparency in selection and documentation of data sources as summarized in Table 3. Application of the methodology was illustrated for maize production in three countries representing a wide range of data availability and quality. While the methodology does not overcome challenges due to lack of data, either because the required data do not exist or are not publicly available, it provides the most appropriate alternatives consistent with a transparent framework and rationale that can be used for all countries and crops. There are two guiding principles at the core of the methodology. First, that the simulation unit to estimate Yp and Yw has relevant agronomic context (combining location  $\times$  water regime  $\times$  crop cycle  $\times$  soil type) and can be aggregated to larger spatial scales through an upscaling protocol based on weighted crop area within each simulation unit (Van Bussel et al., 2015). Second, that all underpinning data should rely as much as possible on observed data, and these data should be publicly available to the extent possible. For data that are of poor quality or currently do not exist or are unavailable (e.g., weather data in many countries in Sub-Saharan Africa), the global agricultural research community should strive to achieve open public access to these weather data because of the importance of estimating yield gaps and food production capacity to support strategic evaluation of local to global food security scenarios (e.g., Global Open Data for Agriculture and Nutrition initiative; [www.godan.info](http://www.godan.info)).

### 3.2. Global databases and their lack of local precision

Given the proliferation of global databases on weather, soil, crop systems and actual yield data that provide required data for crop modelling at global scale, we caution that these ‘new’ databases are, in most cases, recycled existing data of highly varying quality and spatial resolution. For example, many recent databases report data on Ya at a high degree of spatial resolution in gridded global databases (Monfreda et al., 2008; Ray et al., 2012; Iizumi et al., 2013;

**Table 3**  
Summary of data source selection depending upon data availability.

Data availability and quality	Weather	Cropping system	Soil	Data for model calibration	Actual yield
High	Measured data with good quality, >10 years	National databases	National maps linked with high resolution soil profile databases with functional soil properties	High-quality site-year experiments	Most recent annual values reported at a fine spatial level
Intermediate	Propagated ( <i>i.e.</i> , few years of measured data used to create long-term weather)	Expert opinion	Global soil databases	Default parameters retrieved from the literature for similar regions	Annual yields reported at coarser spatial levels or from census, trials, etc.
Low	Best available source of gridded data	Global crop calendars	Local experts information	Default parameters retrieved from the literature for other regions	Yield retrieved from local experts or national-level average

You et al., 2014). Yet this fine resolution is achieved by using data reported at much coarser spatial scales and thus can give a false sense of confidence about data quality. This is especially true for many developing countries where reporting of actual yields is not well developed and weather data and soil data are of poor quality. Moreover, methods used to create these databases are tortuous, not very transparent, and have undergone little independent validation because of the time and effort required. Likewise, data on cropping systems and agronomic practices at a fine spatial scale are scarce. And while recent global databases can help to identify the dominant crop sequence and management (FAO Crop Calendar, 2010; Sacks et al., 2010; Waha et al., 2012; HarvestChoice, 2013), in general, they are too spatially coarse for simulating Yp, Yw, and Yg at specific locations or in small geographic regions. Hence, the most pressing bottleneck for locally relevant crop modelling and yield-gap analysis is not computing power or sophistication of geo-statistical methods running many thousands of simulations and mapping the results, but rather the availability of high-quality, relevant agronomic data on weather, soil, cropping systems, actual yields, and experimental data for model calibration. Indeed, the improved understanding of data requirements and alternatives for yield gap analysis at local to global scales as described here can help identify the most critical “data gaps” and focus global efforts to fill them. Our paper provides a first step in this direction by establishing minimum requirements and quality standards for each data type (weather, crop system, soil, Ya, and model calibration) but further research should be directed to quantitatively determine the relative importance of each data type, relative to the others, for accurate Yg determination.

### 3.3. Public availability of weather data

An uncomfortable truth about weather data is that records taken by government meteorological agencies are often not made publicly available, or they are only available for a price. Table 4 summarizes the weather data sources and confidentiality in countries where yield-gap analysis was performed or is being undertaken by the Global Yield Gap Atlas ([www.yieldgap.org](http://www.yieldgap.org)). Of all locations where yield gap assessments were performed ( $n=365$ ), a respective 65%, 20%, and 15% relied on observed, propagated, and gridded weather data. Weather data could not be made publicly available for 68% of the locations for which observed data were available ( $n=237$ ). In such situations, a viable alternative is to use synthetic weather data created for an adequate time interval using the propagation technique described by Van Wart et al. (2015). This option has the advantage of providing weather data that are similar, though not identical, to the observed weather data, while preserving data confidentiality.

### 3.4. Minimum standards to guide improvement

Whereas the protocol described here sets minimum standards for data selection and quality for yield gap analysis, the current guidelines can be further improved as more and better weather, soil, and cropping system data become available. For example, sowing date is relatively stable across years in temperate climates of Nebraska and Argentina, but highly variable in Kenya where a tropical or sub-tropical climate gives a much wider sowing window (which can be as wide as two months) due to large year-to-year

**Table 4**  
Weather data used in the Global Yield Gap Atlas (GYGA), and public availability of measured weather data.

Country	Number of simulated sites	Proportion (%) of sites with each type of weather data			Proportion (%) of sites for which measured weather data can be made publicly available
		Measured	Propagated	Gridded	
Argentina	16	100	0	0	100
Australia	22	100	0	0	100
Brazil	39	100	0	0	0
Sub-Saharan Africa <sup>a</sup>	183	30	39	31	0 <sup>c</sup>
Bangladesh	11	100	0	0	100
Europe <sup>b</sup>	94	100	0	0	29
Overall	365	65	20	15	32

<sup>a</sup> Includes Burkina Faso, Ghana, Mali, Nigeria, Niger, Ethiopia, Kenya, Uganda, Tanzania, and Zambia.

<sup>b</sup> Includes Denmark, Germany, Poland, Spain, and The Netherlands.

<sup>c</sup> Some of these datasets are available for purchase from the national meteorological organization and were made available to the Global Yield Gap Atlas, but they cannot be provided for open access on the Atlas website ([www.yieldgap.org](http://www.yieldgap.org)).

variation in onset of the rainy season. In Kenya, a dynamic simulation of the sowing date, based on decision rules considering the amount of rainfall or soil water storage, is, perhaps, a more robust approach to better mimic farmer behaviour. Implementing realistic rules to simulate sowing date requires local information about the time window when sowing is likely to occur (given the crop sequence and labour and/or machinery constraints), the specific weather conditions that trigger sowing, and expected management changes that occur when sowing is delayed (e.g., decisions to grow shorter cultivar maturities or to use the crop for forage).

Estimating crop yield gaps within re-designed cropping systems (including different crops, crop sequences within a year, or crop rotations across years) is beyond the scope of the protocol described here because the number of possible permutations is enormous. Although some studies have attempted such re-design, they can only evaluate a limited number of options and selection of these options requires substantial working knowledge and subjective judgement about feasibility given the economic environment and infrastructure (e.g., Davis et al., 2012; Speelman et al., 2014). Likewise, estimating yield gaps for mixed crops stands, where diverse crop species are grown as inter-crops at the same time on the same piece of land, or for local landrace varieties, is made difficult by lack of robust crop models for such complex systems, with lack of uniform sowing patterns and spatial arrangement, and lack of uniformity in genotype-specific attributes governing Yp or Yw in land race seed populations. Due to this complexity, effective yield-gap protocols for such systems have not been developed. It is notable, however, that the global trend of crop agriculture for the past 50 years is towards adoption of modern, improved cultivars grown in pure stands because of higher yields, greater responsiveness to fertilizer, reduced labour, and easier management (e.g., weed control, sowing and harvesting) once farmers have access to inputs and markets (Loomis, 1984; Connor et al., 2011).

## Acknowledgments

We are grateful to the many country agronomists that collaborated with the Global Yield Gap Atlas during the first three years of the project. Their work was reported under the CGIAR research program on Climate Change, Agriculture and Food Security (CCAFS). We especially thank Dr Ochieng Adimo (Jomo Kenyatta University of Agriculture and Technology at Nairobi, Kenya) and Drs Juan Pablo Monzon and Fernando Aramburu Merlos (Instituto Nacional de Tecnología Agropecuaria, University of Mar del Plata, and CONICET, Argentina) for the data provided for this paper. We also thank Nicolas Guilpart (University of Nebraska-Lincoln) for helping preparing Fig. 4. Funding sources include the Bill & Melinda Gates Foundation, Robert B. Daugherty Water for Food Institute at University of Nebraska-Lincoln, USAID, and Wageningen University.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.fcr.2015.03.004>.

## References

- Aggarwal, P.K., 1995. Uncertainties in crop, soil and weather inputs used in growth models—implications for simulated outputs and their applications. *Agric. Syst.* 48, 36–384.
- Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. *Crop evapotranspiration: guidelines for computing crop water requirements*. In: *FAO Irrigation and Drainage Paper 56*. FAO, Rome, Italy.
- Asseng, S., Ewert, F., Rosenzweig, C., Jones, J.W., Hatfield, J.L., Ruane, A.C., Boote, K.J., Thorburn, P.J., Rötter, R.P., Cammarano, D., Brisson, N., Basso, B., Martre, P., Aggarwal, P.K., Angulo, C., Bertuzzi, P., Biernath, C., Challinor, A.J., Doltra, J., Gayler, S., Goldberg, R., Grant, R., Heng, L., Hooker, J., Hunt, L.A., Ingwersen, J., Izaurralde, R.C., Kersebaum, K.C., Müller, C., Naresh Kumar, S., Nendel, C., O'Leary, G., Olesen, J.E., Osborne, T.M., Palosuo, T., Priesack, E., Ripoche, D., Semenov, M.A., Shcherbak, I., Steduto, P., Stöckle, C., Stratonovitch, P., Streck, T., Supit, I., Tao, F., Travasso, M., Waha, K., Wallach, D., White, J.W., Williams, J.R., Wolf, J., 2013. Uncertainty in simulating wheat yields under climate change. *Nat. Clim. Change* 3, 827–832.
- Bagley, J.E., Desai, A.R., Dirmeyer, P.A., Foley, J.A., 2012. Effects of land cover change on moisture availability and potential crop yield in the world's breadbaskets. *Environ. Res. Lett.* 7, 1–9.
- Bai, J., Chen, X., Dobermann, A., Yang, H.S., Cassman, K.G., Zhang, F., 2010. Evaluation of NASA satellite- and model-derived weather data for simulation of maize yield potential in China. *Agron. J.* 102, 9–16.
- Baron, C.B.S., Balme, M., Sarr, B., Teare, S., Lebel, T., Janicot, S., Dingkuh, M., 2005. From GCM grid cell to agricultural plot: scale issues affecting modelling of climate impact. *Philos. Trans. R. Soc. London, Ser. B: Biol. Sci.* 360, 2095–2108.
- Batjes, N.H., 2012. ISRIC WISE derived soil properties on a 5 by 5 arc-minutes global grid (ver. 1. 2). In: *Report 2012/01. ISRIC—World Soil Information*, Wageningen, The Netherlands, pp. 57pp.
- Bert, F.E., Lacianna, C.E., Podesta, G.P., Satorre, E.H., Menendez, A.N., 2007. Sensitivity of CERES-Maize simulated yields to uncertainty in soil properties and daily solar radiation. *Agric. Syst.* 94, 141–150.
- Campbell, G.S., Diaz, R., 1988. Simplified soil–water balance models to predict crop transpiration. In: Bidinger, F.R., et al. (Eds.), *Drought Research Priorities for the Dryland Tropics*. ICRISAT, Patancheru, India, pp. 15–26.
- Challinor, A.J., Martre, P., Asseng, S., Thornton, P., Ewert, F., 2014b. Making the most of climate impacts ensembles. *Nat. Clim. Change* 4, 77–78.
- Challinor, A.J., Watson, J., Lobell, D.B., Howden, S.M., Smith, D.R., Chhetri, N., 2014a. A meta-analysis of crop yield under climate change and adaptation. *Nat. Clim. Change* 4, 287–329.
- Chaney, N.W., Sheffield, J., Villarini, G., Wood, E.F., 2014. Development of a high resolution gridded daily meteorological data set over sub-Saharan Africa: spatial analysis of trends in climate extremes. *J. Clim.* 27, 5815–5835.
- Connor, D.J., Loomis, R.S., Cassman, K.G., 2011. *Crop Ecology. Productivity and Management in Agricultural Systems*. Cambridge Univ. Press, Cambridge, UK.
- Dardanelli, J.L., Bachmeier, O.A., Sereno, R., Gil, R., 1997. Rooting depth and soil water extraction patterns of different crops in a silty loam Haplustoll. *Field Crops Res.* 54, 29–38.
- Davis, A.S., Hill, J.D., Chase, C.A., Johanns, A.M., Liebman, M., 2012. Increasing cropping system diversity balances productivity, profitability and environmental health. *PLoS ONE* 7, e47149.
- De Wit, C.T., Van Keulen, H., 1987. Modelling production of field crops and its requirements. *Geoderma* 40, 253–265.
- Driessen, P.M., Konijn, N.T., 1992. *Land-use Systems Analysis*. Wageningen Agricultural University, Wageningen.
- Evans, L.T., 1993. *Crop Evolution*. In: *Adaptation and Yield*. Cambridge University Press, Cambridge.
- FAO Crop Calendar, 2010. *FAO/AGPS. FAO Crop Calendar*, Available at: (<http://www.fao.org/agriculture/seed/cropcalendar/welcome.do>).
- FAO, 2014. *FAO Statistical Databases*. FAO, Available at: (<http://faostat.fao.org/>).
- FAO, IFPRI, SAGE, 2006. *Agro-Maps. A Global Spatial Database of Agricultural Land-use Statistics Aggregated by Sub-National Administrative Districts*, Available at: (<http://kids.fao.org/agromaps/>).
- Fermont, A.M., van Asten, P.J.A., Tittonell, P., van Wijk, M.T., Giller, K.E., 2009. Closing the cassava yield gap: an analysis from smallholder farms in East Africa. *Field Crops Res.* 112, 24–36.
- Fischer, R.A., Byerlee, D., Edmeades, G.O., 2014. Crop yields and global food security: will yield increase continue to feed the world? In: *ACIAR Monograph No. 158. Australian Centre for International Agricultural Research*, Canberra (xxii + 634 pp).
- Folberth, C., Yang, H., Wang, X., Abbaspour, K.C., 2012. Impact of input data resolution and extent of harvested areas on crop yield estimates in large-scale agricultural modeling for maize in the USA. *Ecol. Model.* 235–236, 8–18.
- French, R.J., Schultz, J.E., 1984. Water use efficiency of wheat in a Mediterranean-type environment. I. The relation between yield, water use and climate. *Aust. J. Agric. Res.* 35, 743–764.
- Gijsman, A.J., Jagtap, S.S., Jones, J.W., 2002. Wadding through a swamp of complete confusion: how to choose a method for estimating soil water retention parameters for crop models. *Euro. J. Agron.* 18, 77–106.
- Gijsman, A.J., Thornton, P.K., Hoogenboom, G., 2007. Using the WISE database to parameterize soil inputs for crop simulation models. *Comp. Elect. Agric.* 56, 85–100.
- Grassini, P., Thorburn, J., Burr, C., Cassman, K.G., 2011. High-yield irrigated maize systems in the Western U.S. Corn-Belt. I. On-farm yield, yield-potential, and impact of agronomic practices. *Field Crops Res.* 120, 142–150.
- Grassini, P., Torrión, J.A., Cassman, K.G., Yang, H.S., Specht, J.E., 2014. Drivers of spatial and temporal variation in soybean yield and irrigation requirements. *Field Crops Res.* 163, 32–46.
- Grassini, P., Yang, H.S., Cassman, K.G., 2009. Limits to maize productivity in the Western Corn-Belt: a simulation analysis for fully irrigated and rainfed conditions. *Agric. For. Meteorol.* 149, 1254–1265.
- HarvestChoice, 2013. *A Time to Sow: How HarvestChoice Is Mapping the Life Cycle of Crops*. International Food Policy Research Institute and University of Minnesota, Washington, DC and St. Paul, MN, Available online at (<http://harvestchoice.org/node/8862>).

- Herdt, R.W., Mandac, A.M., 1981. Modern technology and economic-efficiency of Philippine rice farmers. *Econ. Dev. Cult. Change* 29, 375–399.
- Hodnett, M.G., Tomasella, J., 2002. Marked differences between van Genuchten soil moisture-retention parameters for temperate and tropical soils: a new water retention pedo-transfer function developed for tropical soils. *Geoderma* 108, 155–180.
- Hopper, W.D., 1965. Allocation efficiency in a traditional Indian agriculture. *J. Farm. Econ.* 47, 611–624.
- Hubbard, K.G., Goddard, S., Sorensen, W.D., Wells, N., Osugi, T.T., 2005. Performance of quality assurance procedures for an applied climate information system. *J. Atmos. Oceanic Technol.* 22, 105–112.
- Iizumi, T., Yokozawa, M., Sakurai, G., Travasso, M.A., Romanerukov, V., oetli, P., Newby, T., Ishigooka, Y., Furua, J., 2013. Historical changes in global yields: major cereal and legume crops from 1982 to 2006. *Global Ecol. Biogeogr.* 23, 346–357.
- Jones, C.A., Kiriya, J.R., 1986. CERES-Maize: A Simulation Model of Maize Growth and Development. Texas A&M University Press, College Station, TX.
- Jones, J.W., Hoogenboom, G., Porter, C.H., Boote, K.J., Batchelor, W.D., Hunt, L.A., Wilkens, P.W., Singh, U., Gijsman, A.J., Ritchie, J.T., 2003. The DSSAT cropping system model. *Eur. J. Agron.* 18, 235–265.
- Kassie, B.T., Van Ittersum, M.K., Hengsdijk, H., Asseng, S., Wolf, J., Rötter, R.P., 2014. Climate-induced yield variability and yield gaps of maize (*Zea mays* L.) in the Central Rift Valley of Ethiopia. *Field Crops Res.* 160, 41–53.
- Kersebaum, K.C., Hecker, J.M., Mirschel, W., Wegehenkel, M., 2007. Modelling water and nutrient dynamics in soil–crop systems: a comparison of simulation models applied on common data sets. In: Kersebaum, K.C., Hecker, J.M., Mirschel, W., Wegehenkel, M. (Eds.), *Modelling Water and Nutrient Dynamics in Soil Crop Systems*. Springer, Dordrecht, pp. 1–17.
- Lobell, D.B., Cassman, K.G., Field, C.B., 2009. Crop yield gaps: their importance, magnitude, and causes. *Annu. Rev. Environ. Resour.* 34, 1–26.
- Loomis, R.S., 1984. Traditional agriculture in America. *Ann. Rev. Ecol. Syst.* 15, 449–478.
- Martre, P., Wallach, D., Asseng, S., Ewert, F., Jones, J.W., Rötter, R.P., Boote, K.J., Ruane, A.C., Thorburn, P.J., Cammarano, D., Hatfield, J.L., Rosenzweig, C., Aggarwal, P.K., Angulo, C., Basso, B., Bertuzzi, P., Biernath, C., Brisson, N., Challinor, A.J., Doltra, J., Gayler, S., Goldberg, R., Grant, R.F., Heng, L., Hooker, J., Hunt, L.A., Ingwersen, J., Izaurralde, R.C., Kersebaum, K.C., Müller, C., Kumar, S.N., Nendel, C., O'leary, G., Olesen, J.E., Osborne, T.M., Palosuo, T., Priesack, E., Ripoche, D., Semenov, M.A., Shcherbak, I., Steduto, P., Stöckle, C.O., Stratonovitch, P., Streck, T., Supit, I., Tao, F., Travasso, M., Waha, K., White, J.W., Wolf, J., 2014. Multimodel ensembles of wheat growth: many models are better than one. *Glob. Change Biol.* 10.1111/gcb.12768.
- Mearns, L.O., Easterling, W., Hays, C., Marx, D., 2001. Comparison of agricultural impacts of climate change calculated from high and low resolution climate change scenarios: Part I. The uncertainty due to spatial scale. *Clim. Change* 51, 131–172.
- Monfreda, C., Ramankutty, N., Foley, J.A., 2008. Farming the planet: 2, Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. *Global Biogeochem. Cycles* 22, 1–19.
- Monzon, J.P., Sadras, V.O., Abbate, P.A., Caviglia, O.P., 2007. Modelling management strategies for wheat-soybean double crops in the south-eastern Pampas. *Field Crops Res.* 101, 44–52.
- Monzon, J.P., Sadras, V.O., Andrade, F.H., 2012. Modelled yield and water use efficiency of maize in response to crop management and southern oscillation index in a soil-climate transect in Argentina. *Field Crops Res.* 130, 8–18.
- Nelson, G.C., Rosegrant, M.W., Palazzo, A., Gray, I., Ingersoll, C., Robertson, R., Tokgoz, S., Zhu, T., 2010. Food Security, Farming, and Climate Change to 2050. IFPRI, Washington, DC.
- Nye, P.H., Greenland, D.J., 1960. The Soil under Shifting Cultivation. Commonwealth Agricultural Bureau, Farnham Royal Bucks, UK.
- Passioura, J.B., 1996. Simulation models: science, snake oil, education, or engineering? *Agron. J.* 88, 690–716.
- Penning De Vries, F.W.T., Rabbinge, R., Groot, J.J.R., 1997. Potential and attainable food production and food security in different regions. *Philos. Trans. R. Soc. B: Biol. Sci.* 352, 917–928.
- Rabbinge, R., van Ittersum, M.K., 1994. Tension between aggregation levels. In: Fresco, L.O., Stroosnijder, L., Bouma, J., van Keulen, H. (Eds.), *The Future of the Land: Mobilising and Integrating Knowledge for Land Use Options*. John Wiley & Sons, New York, NY, pp. 31–40.
- Rawls, W.J., Gish, T.J., Brakensiek, D.L., 1991. Estimating soil water retention from soil physical properties and characteristics. *Adv. Soils Sci.* 16, 213–234.
- Ray, D.K., Ramankutty, N., Mueller, N.D., West, P.C., Foley, J.A., 2012. Recent patterns of crop yield growth and stagnation. *Nat. Commun.* 3, 1293.
- Ritchie, J.T., Godwin, D.C., Singh, U., 1990. Soil and weather inputs for the IBSNAT crop models. The International Benchmark Sites Network for Agrotechnology Transfer (IBSNAT) Project. In: Proceedings of the IBSNAT Symposium: Decision Support System for Agrotechnology Transfer. Part I. Symposium Proceedings, Department of Agronomy and Soil Science, College of Tropical Agriculture and Human Resources, October 16–18, 1989, University of Hawaii, Honolulu, Hawaii, Las Vegas, NV.
- Rivington, M., Bellocchi, G., Matthews, K.B., Buchan, K., Donatelli, M., 2004. An integrated modelling approach to conduct multi-factorial analyses on the impacts of climate change on whole-farm systems. In: Pahl, C., Schmidt, S., Jakeman, T. (Eds.), *Proceedings of the iEMSs 2004 International Congress: Complexity and Integrated Resources Management*. International Environmental Modelling and Software Society, June 2004, Osnabrueck, Germany.
- Rivington, M., Matthews, K.B., Bellocchi, G., Buchan, K., 2005. Evaluating uncertainty introduced to process-based simulation models estimated by alternative sources of meteorological data. *Agric. For. Meteorol.* 88, 451–471.
- Romero, C.R., Hoogenboom, G., Baigorria, C.A., Koo, J., Gijsman, A.J., Wood, S., 2012. Reanalysis of a global soil database for crop and environmental modelling. *Environ. Modell. Soft.* 35, 163–170.
- Rosenzweig, C., Elliott, J., Deryng, D., Ruane, A.C., Müller, Ch., Arneth, A., Boote, K.J., Folberth, C., Glotter, M., Khabarov, N., Neumann, K., Piontek, F., Pugh, T.A.M., Schmid, E., Stehfest, E., Yang, H., Jones, J.W., 2014. Assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison. *Proc. Natl. Acad. Sci. U.S.A.* 111, 3268–3273.
- Rurinda, J., 2014. Vulnerability and Adaptation to Climate Variability and Change in Smallholder Farming Systems in Zimbabwe. Wageningen University, Plant Production Systems (PhD thesis).
- Sacks, W.J., Deryng, D., Foley, J.A., Ramankutty, N., 2010. Crop planting dates: an analysis of global patterns. *Glob. Ecol. Biogeogr.* 19, 607–620.
- Sadras, V.O., Calvino, P.A., 2001. Quantification of grain yield response to soil depth in soybean, maize, sunflower, and wheat. *Agron. J.* 93, 577–583.
- Sadras, V.O., Grassini, P., Costa, R., Cohan, L., Hall, A.J., 2014. How reliable are crop production data? Case studies in USA and Argentina. *Food Secur.* 6, 447–459.
- Saxton, K.E., Rawls, R.J., 2006. Soil water characteristic estimates by texture and organic matter for hydrologic solutions. *Soil Sci. Soc. Am. J.* 70, 1569–1578. Available at: (<http://hydrolab.arsusda.gov/soilwater/Index.htm>).
- Sheriff, G., 2005. Efficient waste? Why farmers over-apply nutrients and the implications for policy design. *Rev. Agric. Econ.* 27, 542–557.
- Soil Conservation Service (SCS), 1972. National Engineering Handbook, Hydrology Section 4. Soil Conservation Service (SCS) (Chapters 4–10).
- Speelman, E.N., Garcia-Barrios, L.E., Groot, J.C.J., Tittonell, P., 2014. Gaming for smallholder participation on the design of more sustainable agricultural landscapes. *Agric. Syst.* 126, 62–75.
- Tietje, O., Tapkenhinchs, M., 1993. Evaluation of pedo-transfer functions. *Soil Sci. Soc. Am. J.* 57, 1088–1095.
- Tittonell, P., Corbeels, M., van Wijk, M.T., Vanlauwe, B., Giller, K.E., 2008. Combining organic and mineral fertilizers for integrated soil fertility management in smallholder farming systems of Kenya—explorations using the crop–soil model FIELD. *Agron. J.* 100, 1511–1526.
- Tomasella, J., Hodnett, M.G., Rossato, L., 2000. Pedotransfer functions for the estimation of soil water retention in Brazilian soils. *Soil Sci. Soc. Am. J.* 69, 649–652.
- U.S. Department of Agriculture, Natural Resources Conservation Service. National Soil Survey Handbook, Title 430–VI. Available at: (<http://www.nrcs.usda.gov/wps/portal/nrcs/detail/soils/ref/?cid=nrsc142p2.054223#52>).
- van Bussel, L.G.J., Muller, C., van Keulen, H., Ewert, F., Leffelaar, P.A., 2011. The effect of temporal aggregation of weather input data on crop growth models' results. *Agric. For. Meteorol.* 151, 607–619.
- Van Bussel, L.G.J., Grassini, P., Van Wart, J., Wolf, J., Claessens, L., Yang, H., Boogaard, H., de Groot, H., Saito, K., Cassman, K.G., Van Ittersum, M.K., 2015. From field to atlas: upscaling of location-specific yield gap estimates. *Field Crops Res.* <http://dx.doi.org/10.1016/j.fcr.2015.03.005>, In Press.
- van Heemst, H.D.J., 1988. Plant data values required for simple crop growth simulation models: review and bibliography. In: *Simulation Report CABO-TT Nr. 17*. CABO and Department of Theoretical Production Ecology, Agricultural University, Wageningen University.
- Van Ittersum, M.K., Cassman, K.G., Grassini, P., Wolf, J., Tittonell, P., Hochman, Z., 2013. Field gap analysis with local to global relevance—a review. *Field Crops Res.* 143, 4–17.
- van Ittersum, M.K., Rabbinge, R., 1997. Concepts in production ecology for analysis and quantification of agricultural input–output combinations. *Field Crops Res.* 52, 197–208.
- Van Wart, J., Grassini, P., Cassman, K.G., 2013a. Impact of derived global weather data on simulated crop yields. *Glob. Change Biol.* 19, 3822–3834.
- Van Wart, J., Kersebaum, C., Peng, S., Milner, M., Cassman, K.G., 2013c. Estimating crop yield potential at regional to national scales. *Field Crops Res.* 143, 34–43.
- Van Wart, J., Van Bussel, L.G.J., Wolf, J., Licker, R., Grassini, P., Nelson, A., Boogaard, H., Gerber, J., Mueller, N.D., Claessens, L., Cassman, K.G., Van Ittersum, M.K., 2013b. Reviewing the use of agro-climatic zones to upscale simulated crop yield potential. *Field Crops Res.* 143, 44–45.
- Van Wart, J., Grassini, P., Yang, H.S., Claessens, L., Jarvis, A., Cassman, K.G., 2015. Creating long-term weather data from the thin air for crop simulation modelling. *Agric. For. Meteorol.* <http://dx.doi.org/10.1016/j.agrformet.2015.02.020> (In Press).
- Waha, K., van Bussel, L.G.J., Müller, C., Bondeau, A., 2012. Climate-driven simulation of global crop sowing dates. *Glob. Ecol. Biogeogr.* 21, 247–259.
- Waigri, L.W.I., van Asten, P.J.A., Tenywa, M.M., Bekuna, M.A., 2010. Abiotic constraints override biotic constraints in East African highland banana systems. *Field Crops Res.* 117, 146–153.
- White, J.W., Hoogenboom, G., Wilkens, P.W., Stackhouse Jr., P.W., Hoell, J.M., 2011. Evaluation of satellite-based, modeled-derived daily solar radiation data for the continental United States. *Agron. J.* 103, 1242–1251.
- World Bank—Living Standards Measurement Study—Integrated Study on Agriculture (LSMS-ISA). URL: (<http://econ.worldbank.org/WBSITE/EXTERNAL/EXTDEC/EXTRESEARCH/EXTLSMS/0,contentMDK:21610833~pagePK:64168427~piPK:64168435~theSitePK:3358997,00.html>). [www.fao.org/docrep/X0490e/x0490e00.htm](http://www.fao.org/docrep/X0490e/x0490e00.htm)

- Yang, H.S., Dobermann, A., Lindquist, J.L., Walters, D.T., Arkebauer, T.J., Cassman, K.G., 2004. Hybrid-Maize: a maize simulation model that combines two crop modeling approaches. *Field Crops Res.* 87, 131–154.
- You, J., Hubbard, K.G., Goddard, S., 2008. Comparison of methods for spatially estimating station temperatures in a quality control system. *Int. J. Climatol.* 28, 777–787.
- You, L., Wood, S., Wood-Sichra, U., 2009. Generating plausible crop distribution maps for Sub-Saharan Africa using a spatially disaggregated data fusion and optimization approach. *Agric. Syst.* 99, 126–140.
- You, L., Wood, S., Wood-Sichra, U., Wu, W., 2014. Generating global crop distribution maps: from census to grid. *Agric. Syst.* 127, 53–60.