International Journal of Applied Earth Observation and Geoinformation 38 (2015) 321–334

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

journal homepage: www.elsevier.com/locate/jag

Global rain-fed, irrigated, and paddy croplands: A new high resolution map derived from remote sensing, crop inventories and climate data

J.Meghan Salmon[a,*], Mark A. Friedl[a], Steve Frolking[b], Dominik Wisser[c], Ellen M. Douglas[d]

[a] Department of Earth and Environment Boston University, 685 Commonwealth Ave, Boston, MA 02215, USA
[b] Institute for the Study of Earth, Oceans, and Space University of New Hampshire, 8 College Rd., Durham, NH 03824, USA
[c] Center for Development Research (ZEFc), University of Bonn, Walter-Flex-Straße 3, D-53113 Bonn, Germany
[d] School for the Environment, University of Massachusetts Boston, 100 Morrissey Blvd. , Boston, MA 02125, USA

ARTICLE INFO

Article history:
Received 24 January 2014
Received in revised form 23 January 2015
Accepted 27 January 2015
Available online 16 February 2015

Keywords:
Irrigation
MODIS
Remote sensing
Paddy
Cropland
Water management

ABSTRACT

Irrigation accounts for 70% of global water use by humans and 33–40% of global food production comes from irrigated croplands. Accurate and timely information related to global irrigation is therefore needed to manage increasingly scarce water resources and to improve food security in the face of yield gaps, climate change and extreme events such as droughts, floods, and heat waves. Unfortunately, this information is not available for many regions of the world. This study aims to improve characterization of global rain-fed, irrigated and paddy croplands by integrating information from national and sub-national surveys, remote sensing, and gridded climate data sets. To achieve this goal, we used supervised classification of remote sensing, climate, and agricultural inventory data to generate a global map of irrigated, rain-fed, and paddy croplands. We estimate that 314 million hectares (Mha) worldwide were irrigated circa 2005. This includes 66 Mha of irrigated paddy cropland and 249 Mha of irrigated non-paddy cropland. Additionally, we estimate that 1047 Mha of cropland are managed under rain-fed conditions, including 63 Mha of rain-fed paddy cropland and 985 Mha of rain-fed non-paddy cropland. More generally, our results show that global mapping of irrigated, rain-fed, and paddy croplands is possible by combining information from multiple data sources. However, regions with rapidly changing irrigation or complex mixtures of irrigated and non-irrigated crops present significant challenges and require more and better data to support high quality mapping of irrigation.

Introduction

In the coming decades, demand for food and competition for arable lands will intensify pressure on existing croplands to increase crop yields (Foley et al., 2011). This issue is especially pressing in the developing world, where population growth and higher per-capita consumption will increase demand for food, and where urban expansion and demand for biofuels will create competition for agricultural land (Godfray et al., 2010; Rounsevell et al., 2005; Searchinger et al., 2008). These demands will also compete with the need to preserve services provided by natural ecosystems (Green et al., 2005; Tilman et al., 2001). Given these constraints, management strategies that increase the productivity of existing agricultural land will be required.

Increased agricultural productivity is generally achieved by increasing inputs, augmenting the number of crops grown per year, or both. Worldwide, irrigation is the most important and widespread means of achieving this goal (Lobell et al., 2009). Irrigation reduces yield loss caused by drought and water stress, and increases flexibility in crop planting dates, crop types, and cultivars. Irrigation will therefore continue to be an important management strategy in the coming decades as farmers adapt to the Earth's changing climate (Evans, 1998; Rosenzweig and Parry, 1994).

Irrigation currently accounts for about 70% of global freshwater withdrawals, creating substantial changes to regional hydrology (FAO, 2011; Ozdogan et al., 2010a). In particular, overuse of irrigation has led to unsustainable groundwater depletion in many areas (Lemly et al., 2000; Rodell et al., 2009; Scanlon et al., 2012). Moving forward, it will be important to balance irrigation benefits to crop yields against negative impacts on long-term sustainability and regional hydrology.

* Corresponding author. Present address: Nelson Institute for Environmental Studies, University of Wisconsin – Madison, 1710 University Ave, Madison, WI 53726, USA. Tel.: +1 617 866 3522.
E-mail addresses: Salmon2@wisc.edu (J.Meghan Salmon), friedl@bu.edu (M.A. Friedl), Steve.frolking@unh.edu (S. Frolking), dwisser@uni-bonn.de (D. Wisser), Ellen.douglas@umb.edu (E.M. Douglas).

http://dx.doi.org/10.1016/j.jag.2015.01.014
0303-2434/

At global scale, assessment of the benefits and negative consequences of irrigation requires accurate information regarding the location and extent of irrigated croplands. However, current information regarding global patterns of irrigation includes significant uncertainty, and regional-to-global studies of hydrology, agricultural land use, and food security require improved information related to water use by croplands. Currently, the best available global irrigation data sets reflect irrigation infrastructure and not actual irrigation (Portmann et al., 2010), or are based on census data sets available at the scale of national or sub-national inventory units (e.g., FAOSTAT). Remote sensing, which provides repeated, high resolution information related to land cover and land use at global scales, provides an obvious data source for addressing this information gap. Indeed, several recent studies have used remote sensing with good success to map irrigation at regional to national scales (e.g., Biggs et al., 2006; Pervez and Brown, 2010; Ozdogan and Gutman, 2008). However, global scale remote sensing of irrigation faces many challenges including wide variations in climate conditions, natural vegetation regimes, agricultural management practices, and the quality of available ancillary data.

With these issues in mind, the research and data sets described in this paper have two main elements. First, we develop new methods to address the challenge of mapping global irrigated croplands based on a combination of remote sensing, climate, and inventory data sets. Second, we use these methods to compile a new map of global rain-fed, irrigated, and paddy croplands circa 2005, thereby providing an up-to-date representation of global agricultural water management at significantly higher spatial resolution than is currently available. Because irrigation status is tightly linked to both agricultural production and food security, the data sets developed through this work have significant potential to support studies of agricultural productivity, food security, and the impacts of agriculture on regional-to-global hydrology.

Background

Global irrigation datasets

The United Nations Food and Agriculture Organization (FAO) produces two data sets related to global irrigation. Since 2001, the Statistics Division of the FAO (FAOSTAT) has used annual questionnaires and expert analysis to generate estimates for the area of rain-fed croplands, irrigated croplands, and rice-growing (typically paddy) croplands for most countries worldwide. However, FAOSTAT data are only available at spatial resolutions corresponding to the inventory units for which they were collected. They therefore lack the spatial detail required for many applications. To overcome this, the FAO also maintains a gridded map of irrigation equipment, the Global Map of Irrigated Areas (FAO–GMIA), which is produced at 5 arc-minute spatial resolution (Siebert et al., 2013). This map provides a global representation of the area equipped for irrigation at 5 arc-minute spatial resolution, circa 2005. Note, however, that the area equipped for irrigation represents an upper limit to the irrigated area in each cell; drought, equipment failures, and above-average precipitation can all cause the actual area of irrigated croplands to differ from the area equipped for irrigation. Conversely, the presence of unrecorded irrigation infrastructure, such as illegal wells and boreholes, may cause the FAO–GMIA to under-report irrigation in other areas.

Recently, several projects have produced gridded data sets that represent downscaled versions of production statistics from FAOSTAT. Portmann et al. (2010) combined the FAO–GMIA map with crop type and crop calendar data to create a comprehensive database of global agricultural land use. The result, which they call the Monthly Irrigated and Rain-fed Crop Areas in 2000 (MIRCA2000) database, provides a characterization of global agricultural practices circa 2000 at 5 arc-minute spatial resolution. Although based on the FAO–GMIA data set, the methods used to generate MIRCA2000 account for irrigation infrastructure that was not in use circa 2000 and use data sets derived from agricultural censuses that include information on both irrigation infrastructure and irrigated areas that were harvested. Additionally, the Global Agro-Ecological Zones project produced by the FAO and the International Institute for Applied Systems Analysis offers 5 arc-min grids of irrigated area by crop type in 2000 (IIASA/FAO, 2012). Finally, Liangzhi et al. (2014) developed the Spatial Production Allocation Model to distribute agricultural statistics onto a regular grid.

An alternative to using inventory data, which we exploit here, is to use remotely sensed data sources. For example, Thenkabail (2006) produced a global map of irrigation circa 2000 using data from the Advanced Very High Resolution Radiometer (AVHRR), the Système pour l'Observation de la Terre Vegetation, and several ancillary layers. In addition to mapping the presence of irrigation, the resulting Global Irrigated Area Map which was produced by the International Water Management Institute (hereafter referred to as IWMI–GIAM), characterizes irrigation water source type and cropping intensity. Unfortunately, the area of global irrigation circa 2000 from IWMI–GIAM and FAOSTAT differ by more than 30%. Hence, improved information related to the global distribution of irrigated croplands is urgently needed.

It is important to note that it is relatively rare for irrigation mapping efforts to rely either solely on remote sensing data or not at all. Rather, most global irrigation data sets incorporate a variety of data sources, some inventory based and others derived from remote sensing observations. The most common approach uses remote sensing data sets to spatially distribute inventory statistics (e.g., MIRCA2000). In this study, we similarly blend data sources, but we do not constrain irrigated areas to those estimated in inventory statistics.

Remote sensing of agriculture

The use of remote sensing in agriculture began nearly four decades ago through a series of large scale field experiments: the Corn Blight Watch Experiment (MacDonald et al., 1971), the Large Area Crop Inventory Experiment (MacDonald et al., 1975), and the Agriculture and Resources Inventory Surveys Through Aerospace Remote Sensing experiment (AgRISTARS, 1981). These studies laid the groundwork for modern quantitative remote sensing, and remote sensing is now commonly used in a wide range of agricultural applications including real time monitoring of crop health, famine early warning, and yield forecasts at regional scale (FEWS NET, 2012; USDA FAS, 2012).

Over the last decade several global land cover products have been produced using moderate resolution remote sensing (300-m–1-km) that include one or more classes of agricultural land use. The most prominent examples include GLC2000 (Bartholomé and Belward, 2005), GlobCover (Arino et al., 2007), and NASA's Moderate Resolution Imaging Spectroradiometer (MODIS) Land Cover Type product (Friedl et al., 2010). However, these data sets are intended to capture a variety of land cover attributes, and are not optimized for capturing agricultural land management. Additionally, their moderate resolutions can obscure subpixel variability in land management common in many agricultural regions, particularly those in Africa and Asia.

Building off these efforts, some studies have combined inventory statistics with products derived from remote sensing to compile lower resolution (e.g., 5 arc-minute) maps of global agricultural intensity and extent. Specifically, Ramankutty et al. (2008) used FAO inventory statistics in combination with maps of agricultural land use derived from the MODIS land cover product to create

a global map of agricultural intensity at 5 arc-minute spatial resolution. In a similar study, Goldewijk, et al. (2007) combined data from the IGBP DISCover data set (Loveland et al., 2000), GLC2000, and FAO inventory statistics to create a global map of agricultural extent, also at 5 arc-minute spatial resolution. The datasets created through both these efforts have been widely used, reinforcing the importance and utility of spatially explicit global cropland databases. However, they represent the state of the art in global cropland mapping, and are not perfect representations of agricultural practices on the ground. As they utilize global land cover maps, they incorporate the previously mentioned issues present in those. Additionally, they are based upon the FAO inventory statistics, which are not highly accurate in countries with unstable government or poor data collection networks.

At regional scales, remote sensing data have been used to refine information related to agriculture by supplementing agricultural extent data sets with information related to crop management. For example, Biradar and Xiao (2011) used data from MODIS to map multiple cropping systems in India. Recent work has also demonstrated the feasibility of using remote sensing in concert with crop inventory statistics and gridded climate data to map irrigation at regional-to-national scales (Biggs et al., 2006; Ozdogan and Gutman, 2008; Ozdogan et al., 2010b; Pervez and Brown, 2010). Moreover, several studies have shown the utility of remote sensing data for mapping paddy rice management in Asia (Gumma, 2011; Xiao et al., 2005; Shao et al., 2001).

The goal of the research described in this paper is to build upon these previous efforts by developing a new database depicting Global Rain-fed, Irrigated and Paddy Croplands circa 2005, which we refer to hereafter as GRIPC. Specifically, we implement an approach that uses the agricultural extent information available from the MODIS Land Cover Type Product, then stratifies global cropland areas into the three classes of agricultural water management. The classification utilizes spectral data from remote sensing, climate data, and previous knowledge from the MIRCA2000 data set.

Data and methods

Classification scheme and training data

We map three classes of agricultural land use in GRIPC: rain-fed croplands, irrigated croplands, and paddy croplands (Table 1). Paddy croplands are mapped separately, regardless of irrigation status, due to their distinct hydrological footprint and unique signature in remote sensing data sets (Xiao et al., 2005). Uncropped areas, including both unvegetated and natural vegetation, were excluded from the database using a mask based on the MODIS Land Cover Type product. This mask was created by Damien Sulla-Menashe at Boston University (personal communication, February 14, 2011). It is based primarily on the land cover classification assigned during 2004–2006, but also considered the classifications from surrounding years in the decade. Multiple years of data were considered to reduce the influence of noise in the classified maps as well as short-term changes in land management, such as fallow seasons. The MODIS Land Cover Type Product contains two cropland classes: class 12 (Agriculture) and class 14 (Agriculture/Natural Vegetation Mosaic). In general, areas were retained as cropland when they were classified as a cropland class in a majority of years.

To create GRIPC, we used a supervised classification approach, which relies on training data to estimate a classification model. The training database we used contains 352 sites that were selected to capture the worldwide diversity of global croplands, where each site in the database was composed of a polygon averaging about 600 ha, or roughly 25 MODIS pixels. To build this database, we first assessed 328 agricultural sites that were previously compiled to create the MODIS Global Land Cover Type Product. To support GRIPC, each of these sites was screened and characterized with regard to its irrigation status using a combination of data sources including high-resolution imagery available through Google Earth, the FAO–GMIA map, and in some cases, time series of modeled evaporative demand based on the approach described by Thornthwaite (1948). Google Earth imagery provided information regarding irrigation status when it displayed out-of-season greenness or the presence of irrigation infrastructure. Sites where there was no clear convergence of evidence regarding how the agriculture was managed with respect to irrigation or paddy flooding were eliminated.

To ensure that the training site database successfully captured geographic variability in global croplands, we used the Global Agro-Ecological Zones described by Fischer et al. (2001) (Table 2) and assessed how well each agro-ecozone in each continent was sampled in the site database. Based on this assessment, we added sites to capture rain-fed, irrigated, and paddy croplands in under-represented agro-ecozones and regions. The final site database consisted of 149 sites retained from the MODIS Land Cover Type product database, supplemented with 203 additional sites created using high-resolution satellite data available in Google Earth, yielding a total of 352 sites worldwide (Fig. 1).

Decision tree classification

Classification of cropland types in GRIPC was performed using C4.5, a univariate decision tree classification algorithm that is widely used for land cover mapping and other supervised classification applications (Quinlan, 1993). In generating decision trees, we used a machine learning technique known as "boosting" (Quinlan, 1996), which has been shown to improve the accuracy of classification results. Following the general approach used by the MODIS Land Cover Type Product, boosting allows us to estimate likelihoods for each class at each pixel (McIver and Friedl, 2002). In addition to MODIS data, inputs to the classification included gridded data related to climate and agro-ecozones, which are both predictive of irrigation and therefore improve classification results (Table 3).

MODIS input data were mostly derived from Nadir-BRDF-Adjusted Surface Reflectances (NBAR; Schaaf et al., 2002), which

Table 1
Description of classes of agricultural water management mapped in GRIPC.

Class name	Class description	Number of training sites
Rain-fed croplands	Rain-fed croplands, also called dryland farming, include all cropland where no water from any storage or delivery mechanism is utilized, but crops are not flooded. Harvest must occur at least once per year.	144
Irrigated croplands	Irrigated croplands include cropland where water from available sources is delivered to crops, but crops are not flooded. Harvest must occur at least once per year.	161
Paddy croplands	Paddy croplands, typically used for growing rice, include croplands where fields are flooded, leading to inundation. Flooding must persist longer than two weeks. Water sources for flooding may be rain or irrigation. Harvest must occur at least once per year.	47

Table 2
Description of agroecozones based on Fischer et al. (2001).

	Description	Temperature growing period (TGP)[a]	Length of growing period (LGP)[b]	Number of rain-fed sites	Number of irrigated sites	Number of paddy sites
1	No thermal season	TGP < 90 days	LGP < 90 days	0	0	0
2	One thermal season, moisture limited	90 days < = TGP < 210 days	LGP < TGP	11	12	0
3	One thermal season, no moisture limits	90 days < = TGP < 210 days	LGP > = TGP	10	3	0
4	Two thermal seasons, both moisture limited	210 days < = TGP < 365 days	LGP < 90 days	3	19	0
5	Two thermal seasons, one moisture limited	210 days < = TGP < 365 days	90 days < = LGP <= 210 days	11	13	1
6	Two thermal seasons, no moisture limits	210 days < = TGP < 365 days	LGP > = 210 days	19	15	3
7	No freeze, year-round moisture limits	TGP = 365 days	LGP < 90 days	10	44	1
8	No freeze, one season without moisture limits	TGP = 365 days	90 days < = LGP <= 210 days	34	32	7
9	No freeze, two or more seasons without moisture limits	TGP = 365 days	LGP > = 210 days	39	25	24

[a] TGP is computed by Fischer, et al. (2001) as the number of days when mean daily temperature exceeds 5° C, using the Climate Research Unit (CRU) database climatology (years 1961–1990) (New et al., 1999).
[b] LGP is computed as above, with additional requirements on computed values of actual evapotranspiration (ET_a).

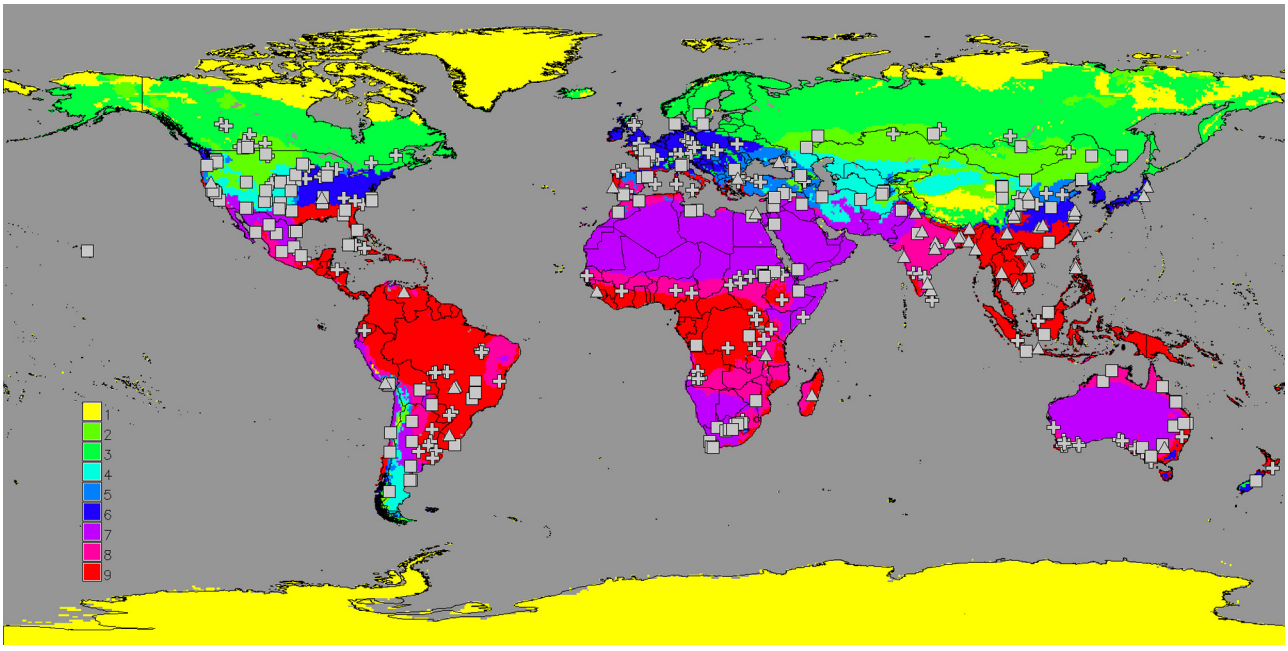

Fig. 1. Distribution of rain-fed (cross), irrigated (square) and paddy (triangle) sites overlaid on the agroecozone stratification (See Table 2 for descriptions).

span the visible, near- and mid-infrared spectrum and are available at 8-day temporal resolution. We also included three indices based on combinations of NBAR data: the two band Enhanced Vegetation Index (EVI2; Jiang et al., 2008), the Normalized Difference Wetness Index (NDWI; Gao and Goetzt, 1995), and the Normalized Difference Infrared Index (NDII; Jang et al., 2006). These indices were included to help capture two key surface properties that are affected by irrigation and paddy management: the amount of photosynthesizing vegetation and the surface wetness. We also included MODIS daytime Land Surface Temperature (LST) and the MODIS diurnal surface temperature range, which provide information related to vegetation surface moisture status (Ozdogan et al., 2010; Wan, 2008). To reduce temporal correlation, we aggregated data in each band to 40-day periods using the mean value of all high quality observations during each 40-day period based on NBAR quality assurance data.

Table 3
Data sets used in decision tree classification of water management classes for agricultural areas.

Name	Description	Spatial resolution	Temporal resolution
MODIS NBAR	8-day observations in 7 bands averaged for 9 periods/year	500 m	~6 weeks
MODIS EVI2, NDWI, and NDII	Calculated from 8-day MODIS data, then averaged for 9 periods/year	500 m	~6 weeks
MODIS LST	Daytime and nighttime 8-day observations averaged for 9 periods/year	1 km	~6 weeks
MODIS annual metrics	Minimum, maximum, mean, and standard deviation of all the above data sets	500 m–1 km	Annual
WORLDCLIM	19 bioclimatic variables	~1 km	Static
Average Climate Moisture Index	Computed using WBM-Plus (Douglas pers comm)	0.5 Degree	Static
Average Annual Moisture Index	Based on the approach described by Ramankutty et al. (2002), computed by Sibley (pers comm.)	5 Min	Static
Agroecozones, generated by BU	Classification based on IIASA Agroecological Zones project plate 07 (Length of Growing Period) and plate 04 (Temperature Growing Period)	0.5 Degree	Static
Hemispheric code	A discrete index (0,1,2) distinguishing among Northern hemisphere (above 23°N), the tropics, and Southern hemisphere (below −23°S)	500 m	Static

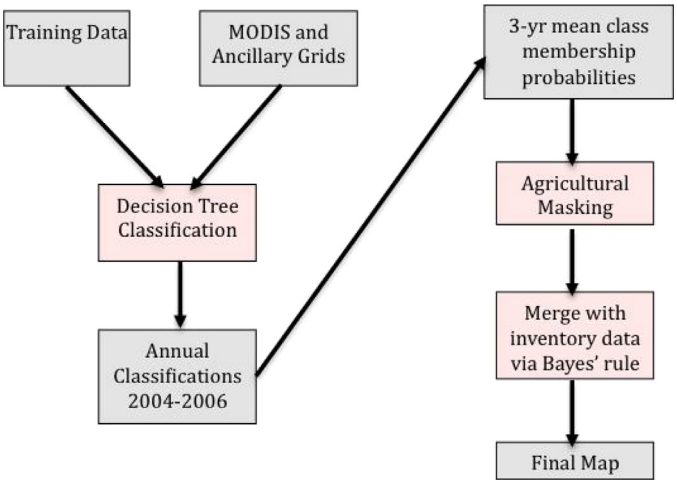

Fig. 2. Flow chart of production method for GRIPC.

To complement these remote sensing inputs, we also included monthly bioclimatic indices from the WORLDCLIM database (Hijmans et al., 2005 http://www.worldclim.org/bioclim), along with two additional climate-based indices that characterize the hydrologic regime at each pixel (Ramankutty et al., 2002; Willmott and Feddema, 1992). These two indices both measure aridity, but differ in the method of estimating moisture availability. One index was developed by Ramankutty, et al. (2002) for assessing agricultural suitability and updated by Adam Sibley at Boston University (personal communication, June 3, 2010). It is based on actual evapotranspiration and potential evapotranspiration. We also used the Climate Moisture Index developed by Willmott and Feddema (1992), which is based upon precipitation and potential evapotranspiration. A visual inspection of both aridity estimates determined that they include unique information, and are both useful in this study.

Fusing classification results with crop inventory data

Crop inventory data provide a wealth of information at national and sub-national scales related to the global distribution of croplands. Remote sensing data, on the other hand, provide indirect but much finer spatial resolution information related to global land cover and land use. To exploit the strengths of each data source we fused results from the remote sensing-based classification described in the previous section with existing inventory-based data sets (Fig. 2). To do this, we used Bayes' rule to combine class-specific likelihoods generated from the decision tree classification at each pixel with spatially explicit prior probabilities for the presence of irrigated, rain-fed, and paddy croplands. To prescribe the prior probability of each class at each cell, we used two layers that are included in the MIRCA2000 data set at 5 arc-minute spatial resolution: (1) global crop extent, and (2) global maximum monthly growing areas. The first layer provides the area for all crops grown in each cell, and the second layer provides the maximum area under cultivation in each month for 26 crop types under rain-fed and irrigated conditions. The area of paddy croplands in each cell was prescribed by adding the areas of rice cultivation under both rain-fed and irrigated conditions from MIRCA2000. The areas corresponding to non-paddy rain-fed and non-paddy irrigated conditions were then estimated by subtracting the area of rain-fed and irrigated rice crops from the total area of rain-fed and irrigated crops. The resulting proportions of rain-fed, irrigated, and paddy croplands in each 5 arc-minute cell were then used to prescribe the prior probability for each class in each cell.

To account for the fact that 500 m MODIS pixels include mixtures of land cover types, we treated pixels labeled as "Agriculture" in the MODIS Land Cover Product as being composed of 80 percent agricultural land use by area, and pixels labeled as agricultural mosaic as being composed of 50 percent agricultural land use. To estimate total irrigated area (including irrigated paddies) at global, continental, and national levels, we used estimates for the proportion of rice paddies managed by irrigation from Huke and Huke (1997). For all Asian countries, we used the country-level values reported by Huke and Huke (1997). For all other countries, we used the average proportion of rice paddies managed by irrigation (60%).

Assessment of irrigation maps

Assessment is a key component of map production, especially when remote sensing data are utilized. Thorough validation exercises are the preferred method for validation, but require high-quality data sets collected at appropriate spatial and temporal scales using random sample designs. In the case of GRIPC, such validation would require independent data regarding the irrigation status of 500 m pixels during 2004–2006 at a random set of locations around the globe. No such data sets are in existence.

We therefore used two approaches for evaluating GRIPC: (1) a cross-validation approach to accuracy assessment based on iterative withholding of sites from the training data, and (2) comparison against existing independent data sets. For the comparison, we identified existing data sets in key locations to use as benchmarks to evaluate GRIPC. Specifically, we collected high-quality, independent data sets describing irrigated and paddy areas at global, regional, and local scales, and we assessed the accuracy of GRIPC relative to these data sets. Our comparisons focus on the Irrigated Cropland and Paddy Cropland classes in Table 1 because other data sets provide maps of global agriculture and the primary goal of this work is to improve information related to the global distribution and intensity of irrigated croplands. Our assessment includes three different scales of analysis:

(i) Global assessment: Country-level estimates of area actually irrigated are available from FAOSTAT for at least one year during the period 2004–2006 for 64 countries, which encompasses 59% of global irrigated area (FAO, 2011). Similarly, rice area harvested, which we use here as a surrogate for paddy area, is available from FAOSTAT for at least one year during 2004–2006 for 112 countries. For countries where multiple years of data are available from FAOSTAT, we compare our results against multiyear averages. While rice area harvested may overestimate paddy area in regions where dryland rice is grown, dryland rice is generally a small proportion (~7%) of total rice agriculture (Huke and Huke, 1997). Hence, this should be a relatively minor source of error. Further, rice area harvested counts double-cropped rice paddies twice, leading to overestimation of the land area devoted to rice-growing in FAOSTAT.

(ii) Regional assessment: We assessed GRIPC at regional scale using (1) non-gridded data for irrigation in Europe, (2) non-gridded data for paddy agriculture in India and China, and (3) gridded data sets describing irrigation in Australia and the United States. For Europe, we used the Farm Structure Survey (FSS), a census of agricultural holdings conducted by all European Union member states that is administered about once every three years (Eurostat, 2004). The FSS does not include small farms, which are defined as farms with less than one hectare and with production below a threshold determined by each member country. Here we used the finest level of FSS data available (Nomenclature of Territorial Units for Statistics (NUTS) level 2), which correspond to sub-national areas that are generally larger than county or district areas. In Asia, we used maps of paddy rice in China and India compiled by Frolking et al. (2002) and Frolking et al. (2006). In North America, we used

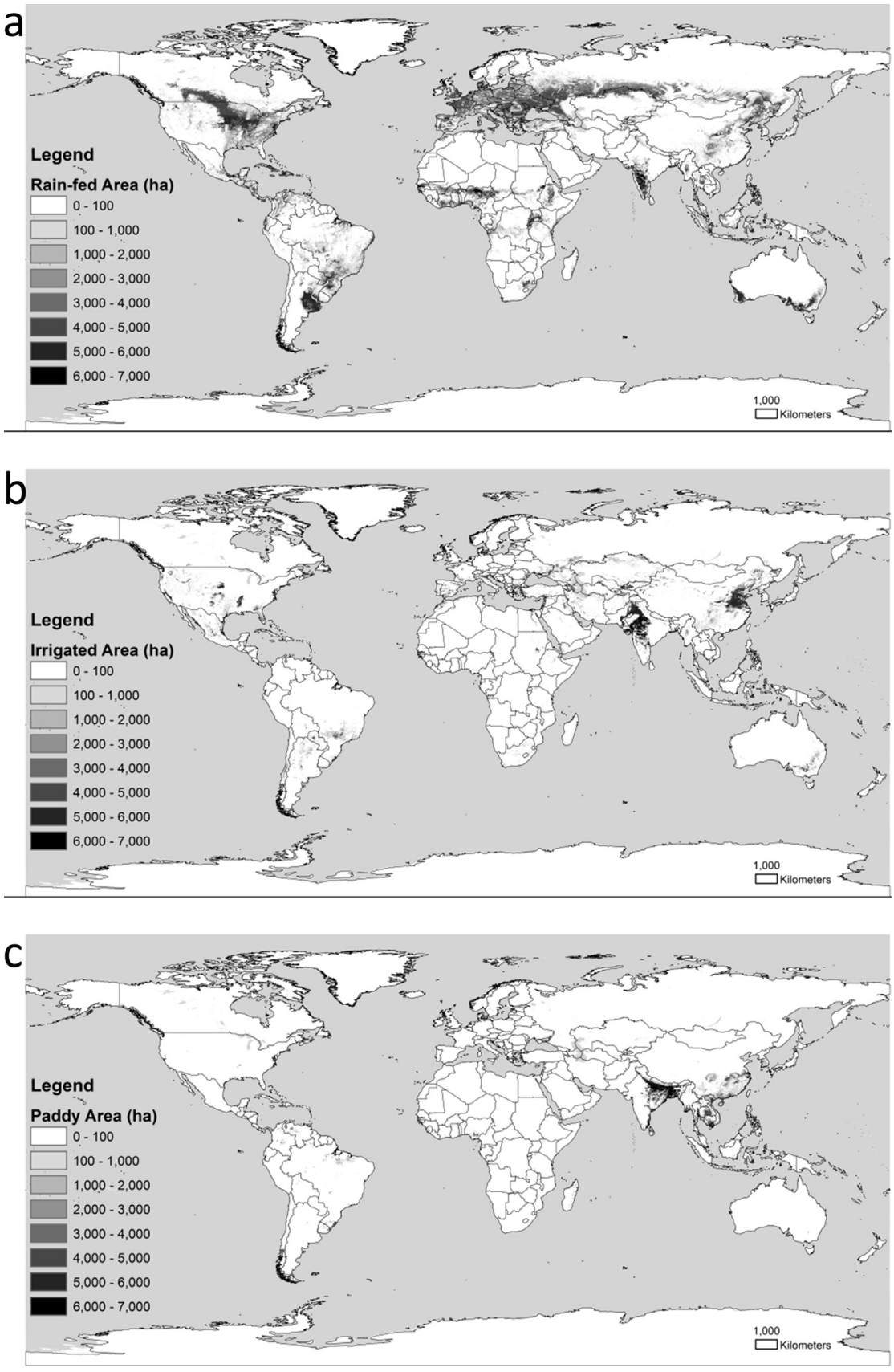

Fig. 3. Global maps of agricultural water management classes. (a) Rain-fed area; (b) Irrigated area; (c) Paddy area.

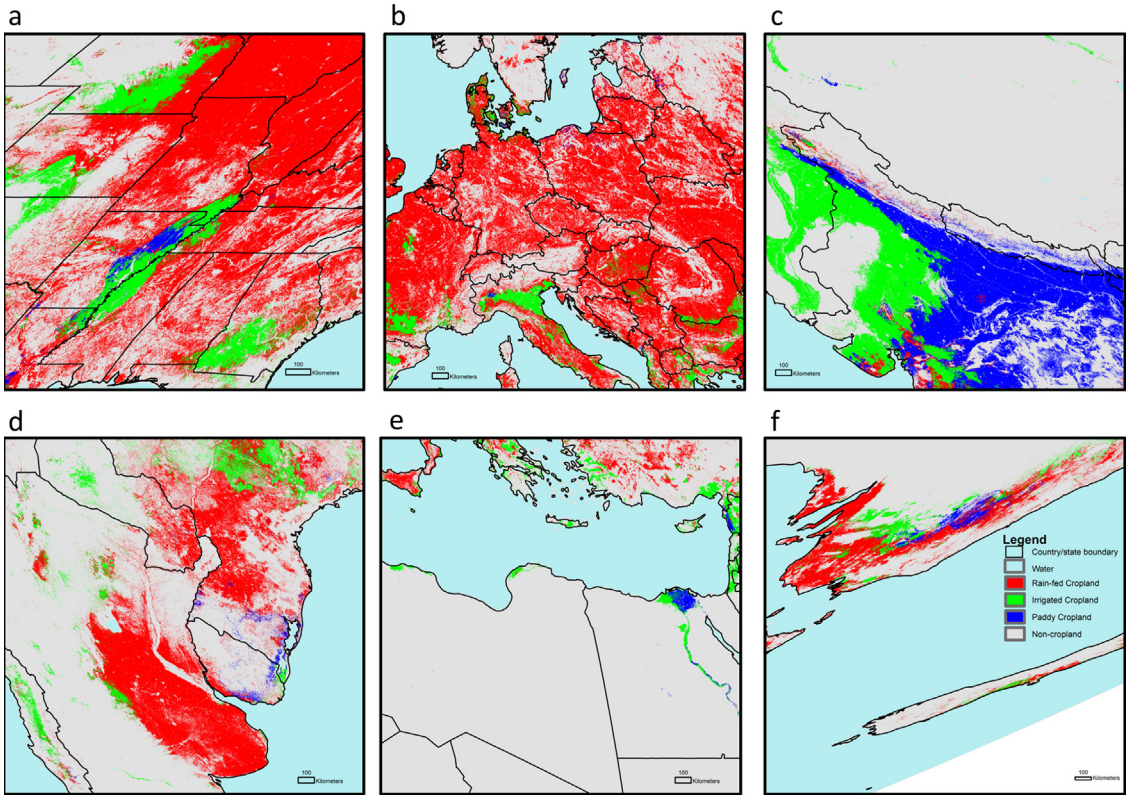

Fig. 4. Full-resolution maps of key agricultural regions. Rain-fed areas are shown in red, irrigated areas in green, and paddy areas in blue. (a) Central United States, (b) Central Europe, (c) Northern India, (d) Argentina, (e) Northern Africa, and (f) Southeastern Australia(For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.).

maps of irrigated area for the United States from the USGS at 250 m spatial resolution in 2007 (MIrAD-US) (Pervez and Brown, 2010). In Australia, we used version 4 of the "ACLUMP" database, which includes a map of irrigated agriculture at 50 m spatial resolution compiled by the Australian Bureau of Agricultural and Resource Economics (ABARES) using inventory statistics and remote sensing data from 2005–2006 (Bureau of Rural Sciences, 2010).

(iii) Local assessment: Estimates of irrigated area from GRIPC were compared against two data sets that map irrigation at the scale of individual fields. The first data set was compiled by the University of Georgia and Albany State University by merging databases for irrigation permits and water use meters with maps derived from aerial imagery in Georgia in 2007 and 2008 (Hook et al., 2009). The second data set provides information on field-scale irrigation in Nebraska, and was compiled by the University of Nebraska Center for Advanced Land Management Information Technologies (CALMIT) using multi-date Landsat imagery and aerial imagery (CALMIT, 2005).

Overview and assessment of GRIPC

Global patterns in rainfed, irrigated and paddy croplands

GRIPC includes 984.6 million hectares (Mha) of rain-fed croplands, 248.5 Mha of non-paddy irrigated croplands, and 128.2 Mha of paddies (Fig. 3; Table 4). The main features of global rain-fed agriculture are clearly captured by GRIPC, including geographically extensive belts of rain-fed croplands in central North America, the Pampas of South America, Europe, central Eurasia, southwestern India, and southern Australia (Fig. 4). Major regions of paddy and non-paddy irrigation are successfully identified in Asia (China, India, and Pakistan) and the United States (Midwest and Central Valley). Conventionally irrigated croplands are most common in the temperate mid-latitudes, while paddy croplands are most common in the tropics.

Comparing global estimates of irrigation is difficult due to differing definitions used by the major data sets. For instance, FAOSTAT does not provide a global estimate of area actually irrigated, since this quantity is not known for all countries. However, FAOSTAT does provide global and continental estimates of area equipped for irrigation, which is expected to be higher than the area actually irrigated. The GRIPC estimate of total irrigated land (314.1 Mha which includes irrigated paddies) is in close agreement with total area equipped for irrigation in 2005 from FAOSTAT (308.5 Mha), but is 21% less than the area included in IWMI–GIAM circa 2000 (398.5 Mha) (Table 4). Further, GRIPC includes ~50% more irrigated land than the total irrigated area in 2000 in MIRCA2000 (217.8 Mha). Finally, the area of global non-paddy irrigated area circa 2005 (248.5 Mha) in GRIPC is substantially higher (~50%) than the area of non-rice irrigated cropland in 2000 included in MIRCA2000 (165.1 Mha).

Table 4
Irrigated areas at global and continental scales from new map and previous studies. All values are in Mega hectares.

	FAOSTAT (area equipped)	MIRCA2000 (area cropped)	IWMI-GIAM[a]	GRIPC
Global	304.5	217.8	398.5	314.1
North America	23.5	27.5	35.4	36.3
South America	11.8	7.9	17.8	16.2
Europe	26.3	15.1	42.5	23
Asia	218.1	153.2	290.6	216.5
Africa	13.5	10.1	8.7	13
Oceania	3.1	2.7	12	7.5

[a] includes temporary fallow areas equipped for irrigation.

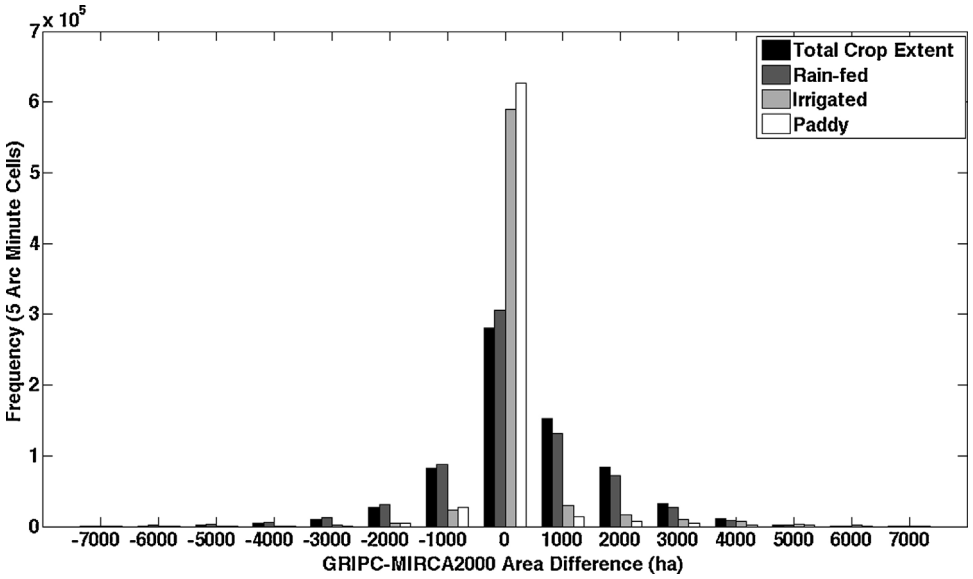

Fig. 5. Histogram of differences in GRIPC (summarized at the 5 arc-minute level) from MIRCA2000 in total cropland (black), as well as rain-fed (dark gray), irrigated (light gray), and paddy (white) croplands.

Comparison of GRIPC with MIRCA2000, which is included in the production of GRIPC via prior probability adjustment, shows how the finer-scale classification of the decision tree has influenced the map. Fig. 5 plots the frequency distribution for differences between GRIPC and MIRCA2000 at 5 arc-minute spatial resolution and demonstrates that GRIPC and MIRCA2000 show broad agreement regarding the amount and location of irrigated and paddy croplands worldwide. The MIRCA2000 estimate is expected to be lower, since it is calculated via the maximum monthly irrigated area, representing a lower limit of irrigated area in regions with multiple growing seasons. Indeed, GRIPC includes less rain-fed cropland and more irrigated cropland than MIRCA2000 in several key agricultural regions including Northwestern India, the North China Plain, the Southern Volga Delta, Morocco, Southeastern US, the upper Mississippi Delta, the US Great Plains, Sao Paulo, and the Murray-Darling basin. Conversely, GRIPC includes less paddy cropland than MIRCA2000 in Northeastern India, Northern Vietnam, and Southern Nigeria.

A cross-validation of GRIPC, produced by withholding each of the training sites and obtaining a class label for the pixels in the withheld site, shows that GRIPC has an overall accuracy of 69% within cropland areas (Table 5). This is similar to the accuracy of other global land cover maps based upon moderate-resolution data, due to the prevalence of mixed pixels (Philippe et al., 2006; Friedl et al., 2010). Moreover, irrigation, which occurs at the field scale, is particularly prone to causing mixed pixels, since agricultural fields in many places are smaller than a 500 m MODIS pixel. This cross-validation approach provides a conservative (lower bound) estimate of map accuracy of GRIPC, since each pixel in the confusion matrix was classified without the benefit of the full training data set used for the final map. Note that these accuracy estimates are not unbiased, since they are not based upon a random sample of locations, and therefore care should be taken with their interpretation.

Comparison with National irrigated cropland inventories

GRIPC estimates of irrigated area agree well with those from FAOSTAT for the 64 countries with FAOSTAT estimates available during 2004–2006 (Fig. 6a). Specifically, GRIPC estimates the percent of country area under irrigation with good agreement (R^2 = 0.89) and a slope close to 1.0 (Fig. 6a). For comparison, the total area available for irrigation (TAAI) from IWMI–GIAM circa 2000 shows similar agreement with FAOSTAT (R^2 = 0.84), but has a slope of 1.4 (i.e., IWMI over-estimated irrigated area relative to FAOSTAT by 40%). Thus, both maps capture country-to-country level variation in agricultural water management, but GRIPC is less prone to overestimating irrigated areas than IWMI–GIAM.

India, China and Pakistan dominate the total global area of irrigation, encompassing ~75 percent of all irrigated area in the countries considered. Interestingly, both IWMI–GIAM and GRIPC estimate significantly more irrigated area in India than FAOSTAT; India contains 58.9 Mha according to FAOSTAT, 101 Mha according to IWMI–GIAM (72 percent more), and 77.5 Mha according to GRIPC (32 percent more). IWMI–GIAM was extensively validated in India, implying that the higher estimates of irrigated area in IWMI-GIAM and GRIPC are reasonable. China, the second most-irrigated country, contains 53.9 Mha of irrigated cropland according to FAOSTAT, 112 Mha according to IWMI–GIAM, and 79.2 Mha according to GRIPC. In Pakistan, the data sets show relatively close agreement: FAOSTAT estimates 19.0 Mha of irrigated cropland, IWMI–GIAM includes 14.0 Mha and GRIPC estimates 15.6 Mha. Together, these three countries account for 84% of the difference (by area) between GRIPC and FAOSTAT, and 89% of the difference between IWMI–GIAM and FAOSTAT.

Irrigated croplands in Europe

Estimates for the irrigated area fraction in Europe from GRIPC show agreement with inventory data in 2004 across 204 FSS NUTS2-level units (R^2 = 0.82; Fig. 6b). However, the slope for a best fit line for this relationship is ~1.3, indicating overestimation of irrigated area fractions in Europe by GRIPC. Indeed, Fig. 6 shows that highly irrigated NUTS2 areas (areas with >0.2 Mha irrigated area in FSS; solid squares in Fig. 6b) often appear above the one-to-one line, indicating that GRIPC overestimates irrigated areas in some irrigated regions, including Northern Italy, Greece, and Southern Spain. The specific reasons for this are unclear. However, some of these semi-arid regions are ideal for both dryland wheat farming and irrigated farming of cash crops, including olives, which presents significant challenges for remote sensing-based discrimination of irrigation (Wriedt et al., 2009a).

For comparison, Fig. 6b shows that the IWMI–GIAM map provides similar results, but tends to underestimate irrigation in areas

Table 5
Confusion matrix produced with cross-validation of GRIPC using training data. Pixels were excluded which were unclassifiable due to missing data or were removed from GRIPC with the cropland mask.

	GRIPC rainfed	GRIPC irrigated	GRIPC paddy	Total	Omission rate	Comission rate	Mapping accuracy
STEP rainfed	1642	162	75	1879	15%	40%	62%
STEP irrigated	699	1273	336	2308	45%	13%	49%
STEP paddy	50	147	371	568	35%	72%	38%
Total	2391	1582	782	4755			69%

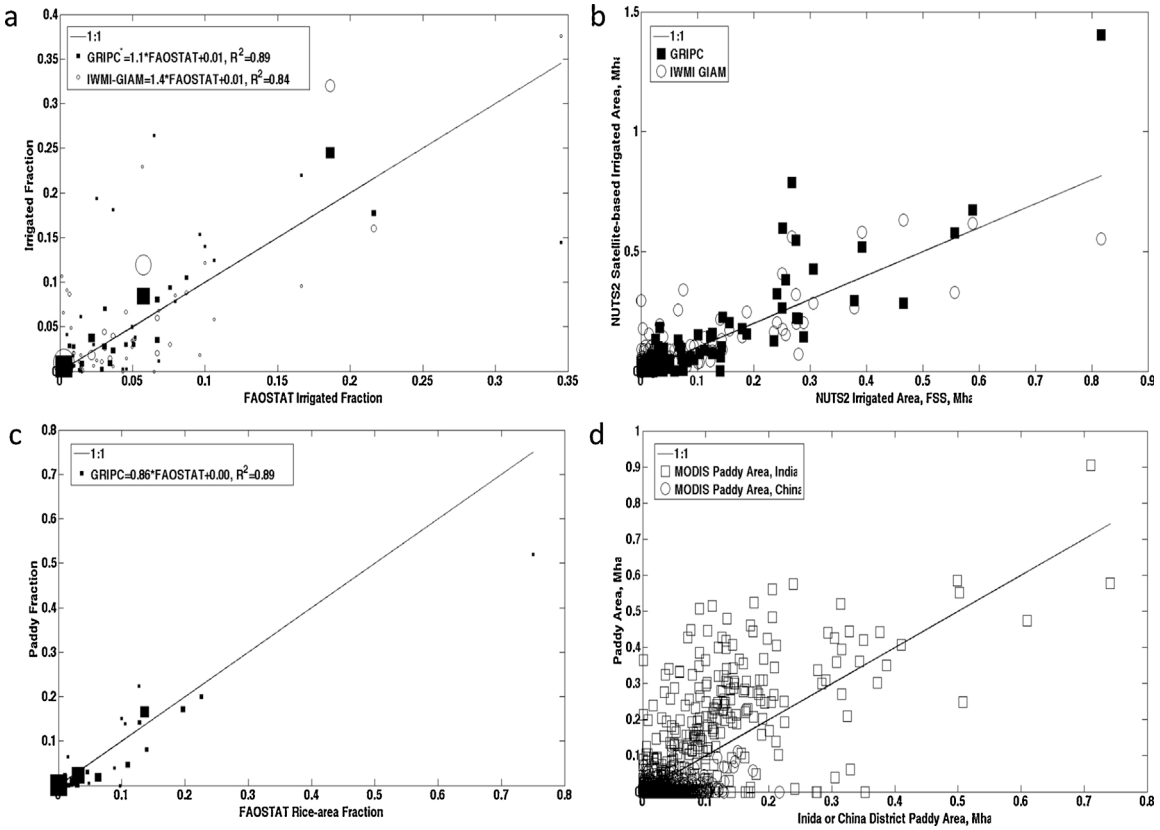

Fig. 6. (a) Comparison of country-level estimates of irrigated area fraction from the new data set presented in this paper and IMWI against corresponding data from FOASTAT. (b) Comparison of satellite-based irrigated area estimates with those from FSS at the NUTS-level. (c) Comparison of country-level estimates of paddy area fraction from the new data set presented in this paper against corresponding data from FOASTAT. (d) Comparison of GRIPC irrigated area estimates with those from Frolking et al. (2002) and Frolking et al. (2006) at the district- or county-level. In panels (a) and (c), Symbols are sized according to country area. Similarly, regression results shown in the legend reflect linear model weighted by country area *includes fraction of paddy area expected to be irrigated, using country-level estimates from Huke and Huke (1997).

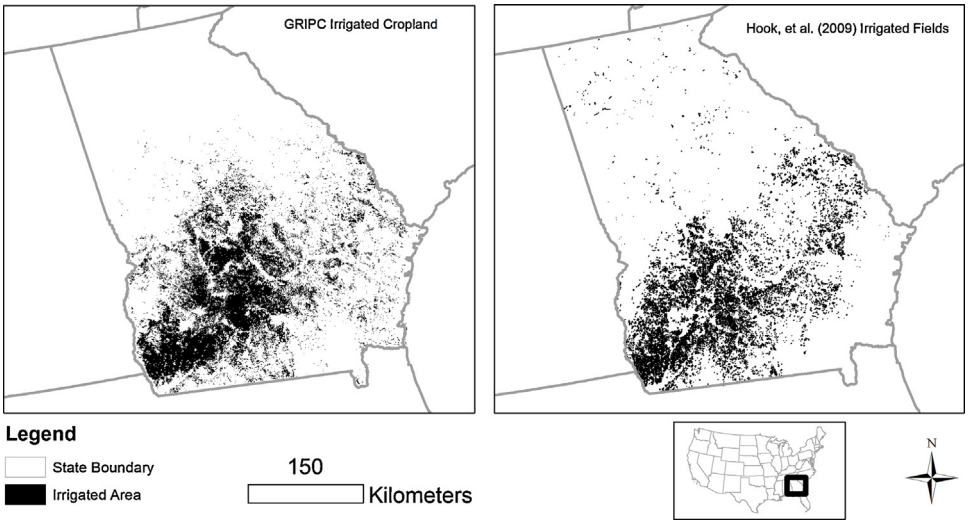

Fig. 7. Maps of irrigation in Georgia, according to GRIPC (left; irrigated croplands are black) and field-level data for the state of Georgia from Hook et al. (2009) (right: irrigated fields are black).

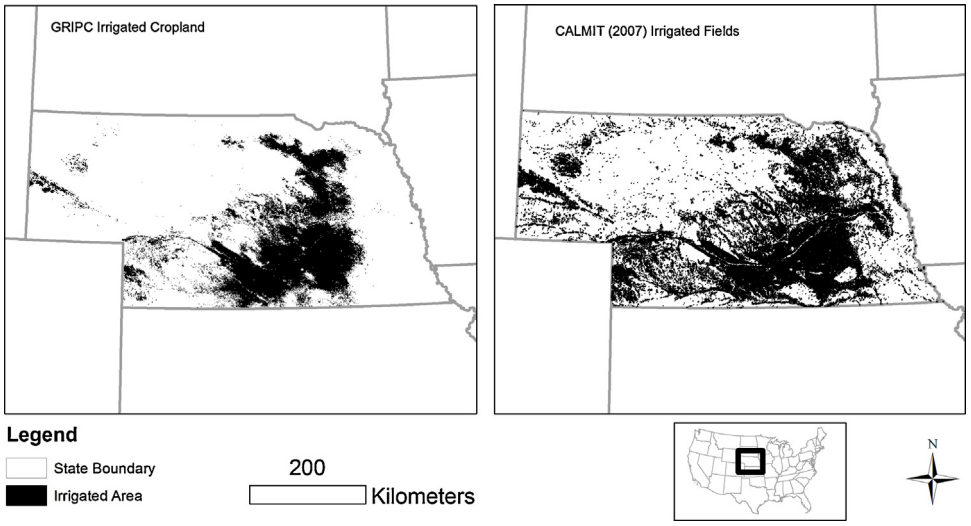

Fig. 8. Maps of irrigation locations in Nebraska, according to GRIPC (left; irrigated croplands are black) and field-level data for the state of Nebraska from CALMIT (2005) (right; irrigated fields are black).

with more irrigation and overestimate irrigation in areas with less irrigation. Overall, the IWMI–GIAM map tends to underestimate irrigation in Europe (slope of linear regression is 0.85), and is also less strongly correlated with the inventory data than GRIPC ($R^2 = 0.71$). Interestingly, the correlation between the GRIPC and IWMI data sets ($R^2 = 0.60$) is less than that between each satellite-based dataset and the inventory data, which demonstrates that different remote sensing observations (e.g., MODIS vs. AVHRR) and mapping approaches (e.g., supervised vs. unsupervised classification) can yield rather different results. For example, although both GRIPC and IWMI–GIAM overestimate irrigated areas in some regions of Europe relative to FSS data, the locations where this occurs tend to be different in each map; whereas GRIPC overestimates irrigated areas in heavily irrigated regions, IWMI–GIAM overestimates irrigated areas in regions with little-to-no irrigated area (areas with <0.1 Mha irrigated area in FSS; open circles in Fig. 6b).

In many European nations, irrigation is used in a supplementary fashion, where water is applied to fields only during droughts (Wriedt et al., 2009b). This style of irrigation is especially common in Central and Northern Europe, where less agricultural area is irrigated. Thus, the inventory data from FSS, which is representative of irrigation in 2004, may not be representative of irrigation patterns in other years, especially in less heavily irrigated areas. GRIPC, on the other hand, appears to accurately reflect the irrigated areas in FSS in these NUTS2 regions, possibly because climatic conditions in Europe were relatively normal from 2004–2006.

Irrigated croplands in North America

In the continental US, GRIPC closely agrees with the representation of irrigation according to the MIrAD-US 2007; 96% of 500 m cells mapped as irrigated in GRIPC are also irrigated in the MIrAD-US data set (Table 6). The User's Accuracy, on the other hand (44%), reflects substantial errors of commission in GRIPC. These errors are most prevalent in humid areas of the Southeastern United States, such as Florida and Southern Texas, as well as the Willamette Valley in Western Oregon. Detection of irrigation with remote sensing in humid areas such as these is complicated by the presence of naturally occurring surface moisture and green vegetation during the crop-growing period. While MIrAD-US is similarly based on MODIS data fused with census data sets, it is more tightly constrained by the census data, obscuring the difficulties that remote sensing data have with detecting irrigation in humid areas. At the same time, errors of omission in GRIPC are also prevalent in some semi-arid regions of the Great Plains, specifically in Eastern Nebraska.

Table 6
Accuracy assessment of GRIPC compared with high-resolution land use maps and field-level irrigation data sets.

Location	Overall accuracy	Producer's accuracy	User's accuracy
Continental United States	0.96	0.59	0.44
Georgia	0.86	0.75	0.26
Nebraska	0.84	0.59	0.68
Australia	0.98	0.39	0.16

In Georgia, GRIPC estimates of irrigated area show strong agreement with field-level data from 2007 and 2008 (overall agreement = 86%), but unlike Eastern Nebraska, tended to over-predict irrigation (user's accuracy of 26%; Fig. 7; Table 6). As we describe above, comparison with the MIrAD-US data set demonstrated that GRIPC tends to over-estimate irrigated area in humid parts of the United States. In spite of the humid subtropical climate, many cropland areas in Georgia require irrigation because of fast-draining soils in the Coastal Plains and the high water requirements of cotton, which is the largest crop in Georgia by area. Soil type and crop type data were not explicitly included in the methods and data we used to create GRIPC. In the future, it may be possible to improve mapping of irrigated croplands by incorporating these types of information in classification models. In the short term, however, accurate detection of irrigation in humid regions will remain a challenge.

GRIPC also showed good agreement with data from field level inventory data in Nebraska in 2005; the overall agreement between inventory and irrigated areas from GRIPC was 84% (Fig. 8). Producer's and user's accuracies (59% and 68%, respectively) suggest that GRIP does tend to underestimate irrigated areas for this region. In particular, GRIPC fails to capture irrigated areas in Northeastern Nebraska. This part of Nebraska possesses a humid continental climate where farmers use deficit irrigation to avoid water stress during critical crop growth stages to minimize unsustainable water withdrawal from the Ogallala Aquifer. This style of irrigation, however, is not well-captured by remote sensing, and thus it is prone to under-representation in GRIPC.

Irrigated cropland in Australia

Accuracy assessment of GRIPC in Australia using the ACLUMP data set shows overall agreement of 98% (Table 6). In general, GRIPC agrees with the ABARES data regarding the general geographic distribution of irrigated lands in Australia in 2005–2006. However, GRIPC also significantly overestimates irrigation in many parts of Australia, which is reflected by a low user's accuracy (16%). Additionally, some areas of dryland wheat were incorrectly labeled as irrigated croplands in GRIPC. As a result, in the GRIPC map, the extent of irrigation in these regions tends to be much larger.

It is also important to note that irrigated agriculture in Australia changed dramatically during the 2004–2006 study period represented by GRIPC. Specifically, the National Water Initiative was adopted in 2004, and was aimed at restoring environmental flows to the Murray–Darling River System through reduced irrigation and water trading. This initiative successfully reduced irrigation by 30% from 2004–2006. Hence, it's possible that some of the positive bias in irrigated area represented by GRIPC is an artifact of changes in irrigated areas that occurred on the ground during the study period.

Paddy areas

Country-level estimates of paddy area from GRIPC are strongly correlated with those from FAOSTAT ($R^2 = 0.89$; Fig. 6c). Compared with rice harvested area from FAOSTAT in 2005, GRIPC appears to slightly underestimate paddy areas (slope = 0.86). However, FAOSTAT includes double-counting of areas that undergo two rice harvests per year, leading to overestimation of rice-growing area in the FAOSTAT data. For example, in Bangladesh, the FAOSTAT estimate of harvested rice area is about 75% of the country's land area, which is probably higher than the land area devoted to rice growing, due to multi-cropping practices. In this context, the GRIPC estimate of paddy area is about 50% of the land area of Bangladesh.

To evaluate the representation of paddy croplands in GRIPC at regional scale, we used the data sets compiled by Frolking et al. (2002), (2006) to compare county and district-level estimates of paddy area in 2000 in China and India, which possess the majority of paddy croplands worldwide. Fig. 6d shows modest overall agreement ($R^2 = 0.54$), but GRIPC tends to underestimate paddy areas in China while overestimating them in India. In both cases, persistent cloud cover in the humid tropics, especially during the monsoon season when much of the rice in Asia is grown, causes over-reliance on ancillary and inventory data sources, which fail to capture finer-scale gradations in paddy area. This likely contributes to under-prediction of paddy croplands in China and over-prediction of paddy croplands in India in GRIPC.

Discussion

Sources of uncertainty in GRIPC

The GRIPC data set was produced by fusing several data sources using a supervised classification methodology. These data include remotely sensed imagery, agricultural census data, meteorological measurements, and surface moisture fields from models. Each of these data sources possess uncertainties and deficiencies that limit their utility as stand-alone sources of information related to irrigation. Our approach was designed to reduce the individual effects of these uncertainties by exploiting joint information related to agriculture and irrigation in each data source. However, in the absence of perfect predictors with low uncertainty, the properties of the data used to create GRIPC inevitably affect our results. In particular, remote sensing data sets include uncertainties due to misregistration, topographic effects, sensor noise, and inadequate screening for clouds and cloud shadows (Lunetta et al., 1991). Uncertainties in agricultural census data sets are less well characterized, and data availability to support robust statistical summaries is a key issue, especially in developing countries (FAO, 2012). Uncertainties in climatological data sets include sparse and missing surface meteorological measurements (New et al., 1999), and the modeled soil moisture data sets used to produce GRIPC include uncertainties associated with the forcing data used to create these modeled fields, as well as the assumptions and errors imbedded in the models used to generate them.

Among these varied sources of uncertainty, it is likely that those influencing the remotely sensed data had the strongest effect on GRIPC. A comparison of the fraction of cropland irrigated in one-degree cells (Fig. 9) shows that the final product of GRIPC is more influenced by the results of the decision tree classifier ($R^2 = 0.64$) than by the information in MIRCA2000 ($R^2 = 0.33$). The majority of the data provided to the decision tree classifier is based on remotely sensed observations, suggesting that these are exerting the strongest influence on the final designation of water management status in GRIPC.

An additional source of error in our approach is the agricultural mask we used to distinguish natural land cover and urban areas from agricultural land. For GRIPC, we used the MODIS Land Cover Type Product, which contains uncertainties due to land cover changes and classifier error. By using multi-year time series of the MODIS Land Cover Type product, our methodology reduces uncertainty associated with errors in this product. However, our approach does not eliminate all errors in the mask, and remaining errors are propagated into GRIPC. In particular, land cover and land use in areas with heterogeneous mixes of agriculture and natural vegetation are difficult to map from MODIS, and are therefore likely to have the higher frequencies of classification error in the MODIS product.

Assessment of uncertainty in GRIPC

The utility of land use and land cover maps is greatly improved with the addition of an accuracy assessment using independent data at randomly distributed sites or clusters of sites. The benefits of such an exercise include the addition of uncertainty estimates and refinement of any area estimates produced from the map. However, assessing a global map in this way is a very labor intensive undertaking, sometimes requiring as much time and effort as the original map production. Therefore, most global cropland maps rely on alternate methods of assessment (e.g., Siebert et al. (2005), (2013); Portmann, et al. (2010); Ramankutty et al. (2008)).

Among land cover and land use characteristics, irrigation status is particularly difficult to assess using available data. The dynamic nature of agricultural management heightens the necessity of obtaining data during the time period of interest, which is not always feasible for randomly selected sites. For instance, Google Earth imagery may not be available during the time period of interest, and signs of irrigation may not be present outside of the growing season and year during which it was used. As an example, we were unable to ascertain the irrigation status of 55% of the original 328 STEP sites assessed in this study for training data. This level of missing data would substantially decrease the utility of any accuracy assessment that relies on Google Earth. The alternative is to acquire high spatial resolution data following a strategy such as that described by Olofsson et al. (2012). However, acquisition and analysis of this type of imagery is far beyond the scope of what is feasible for most studies. In this study, we instead used a cross-validation analysis based on the training data as well as comparison with a variety of available independent data sets produced at global, regional, and local scales to assess the accuracy of GRIPC.

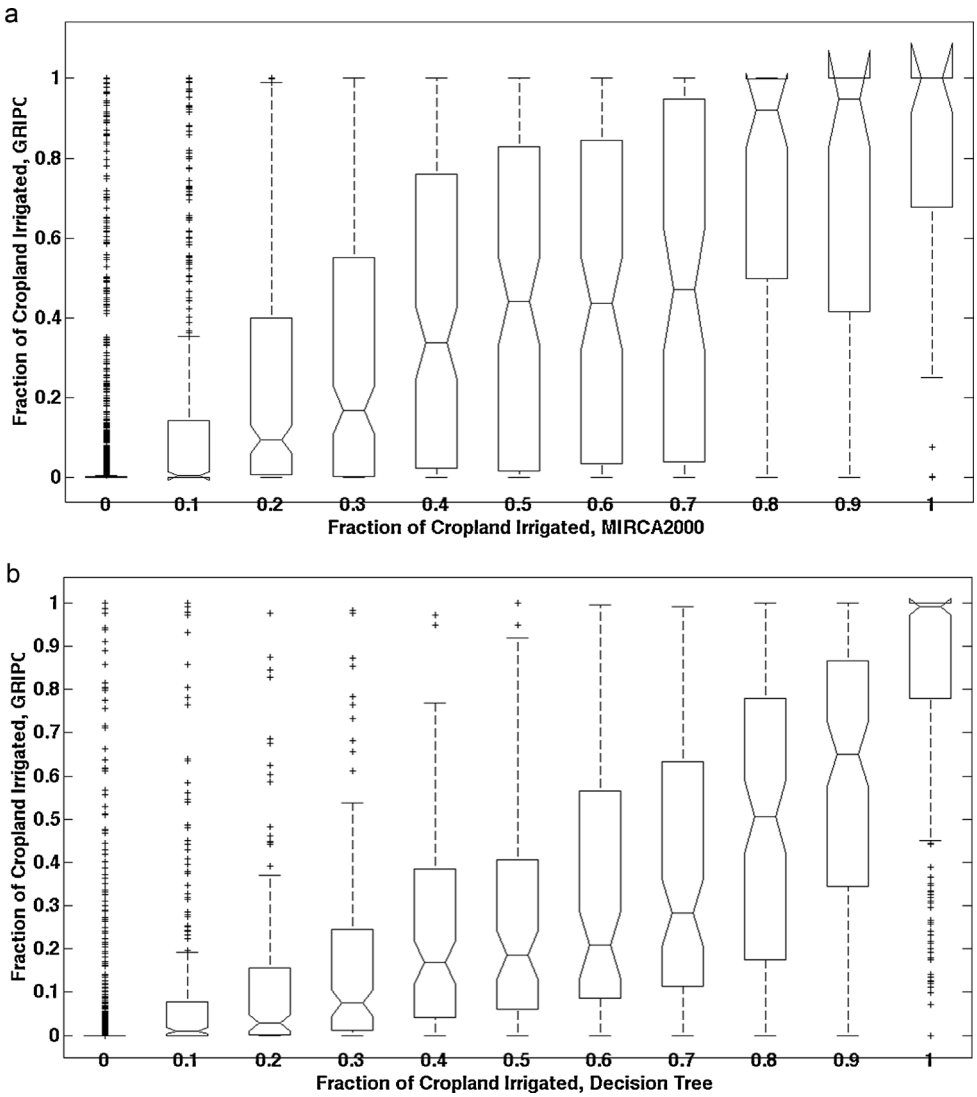

Fig. 9. Global comparison of the area fraction of irrigated cropland irrigated in one-degree cells for MIRCA2000 vs. GRIPC (upper panel) and decision tree classifier vs. GRIPC (lower panel). See Fig. 2 for full classification algorithm. GRIPC estimates are correlated more strongly with those from the decision tree classifier than those from MIRCA2000. Horizontal lines are median, boxes represent 25–75th%; whiskers extend to the most extreme points not considered outliers, and + are outliers.

Relationship of GRIPC to MIRCA2000 and IWMI–GIAM

Unlike GRIPC, MIRCA2000 is based almost exclusively on agricultural census and inventory data. To spatialize these data onto a regular grid, MIRCA2000 uses the Cropland2000 data set (Ramankutty et al., 2008), which is derived from older global land cover datasets (GLC2000 and Version 4 of the MODIS Land Cover Type Product) to map cropped areas within administrative census units. Thus, while GRIPC and MIRCA2000 both rely upon satellite-based land cover maps to identify the global distribution of croplands, they differ in terms of the data sets used to mask agriculture and how these data are merged with other data sources. Specifically, MIRCA2000 uses several older sources of land cover data, and relies on census estimates to determine the total cropland area in each administrative unit. GRIPC, on the other hand, uses a single remote-sensing derived source of land cover data (the MODIS Land Cover Product) and does not rely on census data to determine the location of croplands. Equally important, MIRCA2000 and GRIPC differ in the data sources used to determine the irrigation status of croplands. Whereas MIRCA2000 relies entirely on census data to do this, GRIPC uses a combination of remote sensing imagery, climate and model derived data sets, and agricultural census information. As a result, uncertainties and errors in MIRCA2000 reflect limitations of the FAO agricultural census data, while errors and uncertainties in GRIPC are less dependent on a single data source.

The IWMI–GIAM data set, on the other hand, relies entirely on remote sensing data sources to determine both the distribution of croplands and their irrigation status. As a result, uncertainties and limitations of the remote sensing data used to produce this dataset (e.g., missing data, misregistration, gridding errors, topographical effects, and classifier error) affect the quality of the final product. GRIPC also relies heavily on remote sensing data sources, but incorporates a variety of other data sources. Hence, while both data sets are influenced by limitations of remote sensing, GRIPC attempts to reduce errors introduced by these limitations by exploiting information from agricultural inventory, climate and modeled surface moisture fields.

Paddy croplands

Paddy agriculture is an enormously important class of croplands (Maclean and Hettel, 2002). However, because paddy croplands

are mostly located in the humid tropics and are frequently multi-cropped, they are difficult to map using remote sensing. Specifically, persistent cloud cover precludes observations through much of the year in many parts of the tropics. As a result, MODIS time series contain high proportions of missing data over many areas with the highest density of paddy croplands. For example, in a major rice-growing area of Uttar Pradesh in Northern India, 44% of MODIS NBAR observations were missing. Similarly, rice-growing areas in Jiangsu in Eastern China have about 61% of MODIS observations missing in an average year. Because of this, remote sensing-based results tend to be of lower quality, and our methodology to map the distribution of paddy croplands relies more heavily on other data sources, especially crop inventories data, which are generally reported for political units, not high-resolution grids, and may not capture the full dynamics of agriculture in rapidly changing economies (e.g., Deng et al., 2009). Comparisons of paddy croplands in GRIPC with existing maps of paddy agriculture suggest that the representation of paddy croplands in GRIPC is consistent with these other data sets. However, it is important to note there is clearly room for improvement in this element of the GRIPC data set. Moving forward, additional data sources derived from microwave remote sensing (e.g., Bouvet et al., 2009) or methods that exploit time series properties of remote sensing more effectively (e.g., such as those described by Xiao et al. (2005) and Shao et al. (2001)) may provide avenues to help resolve the limitations of optical remote sensing and supervised classification of paddy croplands.

Conclusions

We developed a new map of global irrigated, rain-fed, and paddy croplands circa 2005 at 500 m spatial resolution by merging information from satellite remote sensing, climate, and census-based inventory data sets. To generate this map, a database of sites characterizing the geographic distribution of global croplands was used to estimate a supervised classification of global cropland classes. Results from this classifier were merged with available inventory data to create the final map, which we call GRIPC. GRIPC includes 314 Mha of irrigated cropland, which agrees closely with national inventory statistics from FAOSTAT (305 Mha), but includes 24% less area of irrigated cropland than the IWMI–GIAM map of global irrigation (399 Mha). The map is available for download at https://dl.dropboxusercontent.com/u/12683052/GRIPCmap.zip.

The results from this effort demonstrate the potential of merging remote sensing and existing inventory data sets to improve information regarding water management of global croplands. At the same time, there is clear room for refinement of the methods and results we present in this paper. For example, because GRIPC uses the MODIS Land Cover Type Product to discriminate natural vegetation and urban areas from agricultural land areas, errors in this product will propagate into GRIPC. In addition, GRIPC does not include several important classes of irrigated agriculture including deficit irrigation (irrigation occurring less than once per year), permanent crops (orchards and vineyards), and un-harvested pastures. Future efforts could further refine irrigation estimates by identifying areas under these types of management.

A key goal of this effort is to provide an improved basis for regional to global scale modeling and analysis studies that require information on irrigated agriculture. Potential applications that might benefit from GRIPC include regional-to-global estimation of green and blue water use by agriculture (e.g., Liu and Yang (2010)), assessment of irrigation impacts on regional hydrology and climate (e.g., Lobell et al. (2008)), and analyses of global yield gaps (e.g., Licker et al. (2010)). In addition, because GRIPC provides information on crop water management at 500 m spatial resolution, it should be particularly useful for high resolution land surface modeling at global scale (e.g., Wood et al., 2011).

As climate change impacts become more pronounced and worldwide demand for agricultural commodities increases in the coming decades, agricultural practices are likely to adapt and change rapidly. In particular, water demand by agriculture is projected to increase, but will only be met if sustainable sources of water are available. Hence, knowledge regarding the global distribution and intensity of irrigation is critical to balancing water needed for food production with water needed for industry, households, and the environment. Results from this study can support this type of analysis by providing a new high-resolution global map rain-fed, irrigated and paddy croplands. In this context, GRIPC has two key innovations relative to available global crop type and management databases (e.g., Monfreda et al., 2008; Portmann et al., 2010). First, GRIPC has relatively high spatial resolution (500 m vs. 5 arc-minute). Second, it was produced using a consistent methodology that does not change across administrative boundaries. More generally, the data and methods presented in this paper provide a foundation for ongoing monitoring of global water use by agriculture in the coming decades.

Acknowledgements

This work was supported by a NASA Earth System Science Fellowship to JMS (NNX09AO02H). Support was also provided by NASA grantsNNX07AW07G and NNX11AE75G, and through the National Science Foundation program in Water, Sustainability and Climate (EAR-1038907). The authors also gratefully acknowledge contributions of Damien Sulla-Menashe, for providing the multi-year cropland mask, and Damien Sulla-Menashe, Douglas Bolton and Adam Sibley for help with database development for cropland training sites.

References

AgRISTAR: Agriculture and Resources Inventory Surveys through Aerospace Remote Sensing, 1981. NASA, Lyndon B Johnson Space Center, Houston, Texas59, 110–123.

Arino, O., Gross, D., Ranera, F., Bourg, L., Leroy, M., Bicheron, P., Weber, J.L., 2007. GlobCover: ESA service for global land cover from MERIS. In: Geoscience and Remote Sensing Symposium, 2007. IGARSS 2007. IEEE International. IEEE, Barcelona, Spain, pp. 2412–2415.

Bartholomé, Belward, A.S., 2005. GLC2000: a new approach to global land cover mapping from Earth observation data. Int. J. Remote Sens. 26 (9), 1959–1977.

Biggs, T.W., Thenkabail, P.S., Gumma, M.K., Scott, C.A., Parthasaradhi, G.R., Turral, H.N., 2006. Irrigated area mapping in heterogeneous landscapes with MODIS time series, ground truth and census data, Krishna Basin, India. Int. J. Remote Sens. 27 (19), 4245–4266.

Biradar, C.M., Xiao, X., 2011. Quantifying the area and spatial distribution of double-and triple-cropping croplands in India with multi-temporal MODIS imagery in 2005. Int. J. Remote Sens. 32 (2), 367–386.

Bouvet, A., Le Toan, T., Lam-Dao, N., 2009. Monitoring of the rice cropping system in the Mekong Delta using ENVISAT/ASAR dual polarization data. Geosci. Remote Sens. IEEE Trans. 47 (2), 517–526.

Bureau of Rural Sciences, 2010. 'User Guide and Caveats: Land Use of Australia Version 4, 2005–06'. Bureau of Rural Sciences, Canberra.

Center for Advanced Land Management Information Technologies (CALMIT), 2005. Nebraska Center Pivot Irrigation Systems, 2007. University of Nebraska-Lincoln (accessed March 2010) http://calmit.unl.edu/2005landuse/statewide.shtml

Deng, J.S., Wang, K., Hong, Y., Qi, J.G., 2009. Spatio-temporal dynamics and evolution of land use change and landscape pattern in response to rapid urbanization. Landscape Urban Plan. 92 (3), 187–198.

Eurostat, 2004. General and RegionalStatistics Database, edited, URL:http://epp.eurostat.cec.eu.int [Accessed: May 2012].

Evans, L.T., 1998. Feeding the Ten Billion: Plants and Population Growth. Cambridge University Press, Cambridge, UK.

FAO, FAOSTAT. 2011. http://faostat.fao.org/default.aspx?lang=en (accessed November 2011).

FAO, 2012. Action Plan of the Global Strategy to Improve Agricultural and Rural Statistics. Food & Agriculture Org.

FEWS NET, 2012. http://www.fews.net/Pages/default.aspx (accessed August 2012).

Fischer, G., Shah, M., van Velthuizen, H., & Nachtergaele, F. O. (2001). Global agro-ecological assessment for agriculture in the 21st century.

Foley, J.A., Ramankutty, N., Brauman, K.A., Cassidy, E.S., Gerber, J.S., Johnston, M., Zaks, D.P., 2011. Solutions for a cultivated planet. Nature 478 (7369), 337–342.
Friedl, M.A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., Huang, X., 2010. MODIS Collection 5 global land cover: algorithm refinements and characterization of new datasets. Remote Sens. Environ. 114 (1), 168–182.
Frolking, S., Qiu, J., Boles, S., Xiao, X., Liu, J., Zhuang, Y., Qin, 2002. Combining remote sensing and ground census data to develop new maps of the distribution of rice agriculture in China. Global Biogeochem. Cycles 16 (4), 1091.
Frolking, S., Yeluripati, J.B., Douglas, E., 2006. New district-level maps of rice cropping in India: a foundation for scientific input into policy assessment. Field Crops Res. 98 (2), 164–177.
Gao, B.C., Goetzt, A.F., 1995. Retrieval of equivalent water thickness and information related to biochemical components of vegetation canopies from AVIRIS data. Remote Sens. Environ. 52 (3), 155–162.
Godfray, H.C.J., Beddington, J.R., Crute, I.R., Haddad, L., Lawrence, D., Muir, J.F., Toulmin, 2010. Food security: the challenge of feeding 9 billion people. Science 327 (5967), 812–818.
Klein Goldewijk, K., Van Drecht, G., Bouwman, A.F., 2007. Mapping contemporary global cropland and grassland distributions on a 5×5 min resolution. J. Land Use Sci. 2 (3), 167–190.
Green, R.E., Cornell, S.J., Scharlemann, J.P., Balmford, A., 2005. Farming and the fate of wild nature. Science 307 (5709), 550–555.
Gumma, M.K., 2011. Mapping rice rreas of south Asia using MODIS multitemporal mata. J. Appl. Remote Sens. 5 (1), 53547, http://dx.doi.org/10.1117/13619838.
Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. Int. J. Climatol. 25 (15), 1965–1978.
Hook, J.E., Hoogenboom, G., Paz, J., Risse, M., Bergstrom, J., Mullen, J., 2009. Assessing agricultural groundwater needs for the future: identifying irrigated area and sources. In: Proceedings of the 2009 Georgia Water Resources Conference, held April 27–29, 2009, University of Georgia.
Huke, R. E. and Huke, E.H.1997. Rice Area by Type of Culture: South, Southeast, and East Asia, a Revised and Updated Data Base (September 18): 1–53.
IIASA/FAO, 2012. Global Agro-Ecological Zones (GAEZ v3.0). IIASA, Laxenburg, Austria and FAO, Rome, Italy.
Jang, J.D., Viau, A.A., Anctil, F., 2006. Thermal-water stress index from satellite images. Int. J. Remote Sens. 27 (8), 1619–1639.
Jiang, Z., Huete, A.R., Didan, K., Miura, T., 2008. Development of a two-band enhanced vegetation index without a blue band. Remote Sens. Environ. 112 (10), 3833–3845.
Lemly, A.D., Kingsford, R.T., Thompson, J.R., 2000. Irrigated agriculture and wildlife conservation: conflict on a global scale. Environ. Manage. 25 (5), 485–512.
Licker, R., Johnston, M., Foley, J.A., Barford, C., Kucharik, C.J., Monfreda, C., Ramankutty, N., 2010. Mind the gap: how do climate and agricultural management explain the 'yield gap'of croplands around the world? Global Ecol. Biogeogr. 19 (6), 769–782.
Liu, J., Yang, H., 2010. Spatially explicit assessment of global consumptive water uses in cropland: green and blue water. J. Hydrol. 384 (3), 187–197.
Lobell, D.B., Bonfils, C.J., Kueppers, L.M., Snyder, M.A., 2008. Irrigation cooling effect on temperature and heat index extremes. Geophys. Res. Lett. 35 (9), L09705.
Lobell, D.B., Cassman, K.G., Field, C.B., 2009. Crop yield gaps: their importance, magnitudes, and causes. Annu. Rev. Environ. Resour. 34 (1), 179.
Loveland, T.R., Reed, B.C., Brown, J.F., Ohlen, D.O., Zhu, Z., Yang, L.W.M.J., Merchant, J.W., 2000. Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. Int. J. Remote Sens. 21 (6–7), 1303–1330.
Lunetta, R.S., Congalton, R.G., Fenstermaker, L.K., Jensen, J.R., McGwire, K.C., Tinney, L.R., 1991. Remote sensing and geographic information system data integration: error sources and research issues. Photogramm. Eng. Remote Sens. 57 (6), 677–687.
MacDonald, R.B., Bauer, M.E., Allen, R.D., Clifton, J.W., Erickson, J.D., Landgrebe, D.A., 1971. Results of 1971 the corn blight watch experiment.LARS Technical Reports. Purdue University107.
MacDonald, R.B., Hall, F.G., Erb, R.B., 1975. The use of Landsat data in a large area crop inventory experiment (LACIE). In: Laboratory for Applications of Remote Sensing LARS Symposia, May 1975, Purdue University, pp. 1–25.
Maclean, J.L., Hettel, G.P. (Eds.), 2002. Rice almanac: Source Book for the Most Important Economic Activity on Earth. IRRI (free PDF download).
Philippe, Mayaux, Eva, Hugh, Gallego, Javier, Strahler, Alan H., Herold, Martin, Agrawal, Shefali, Naumov, Sergey, 2006. Validation of the global land cover 2000 map. Geosci. Remote Sens. IEEE Trans. 44 (7), 1728–1739.
McIver, Friedl, M.A., 2002. Using prior probabilities in decision-tree classification of remotely sensed data. Remote Sens. Environ. 81 (2), 253–261.
Monfreda, C., Ramankutty, N., Foley, J.A., 2000. Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. Global Biogeochemical Cycles 22 (1).
New, M., Hulme, M., Jones, P., 1999. Representing twentieth-century space-time climate variability. Part I: Development of a 1961–90 mean monthly terrestrial climatology. J. Clim. 12 (3), 829–856.
Ozdogan, M., Gutman, G., 2008. A new methodology to map irrigated areas using multi-temporal MODIS and ancillary data: An application example in the continental US. Remote Sens. Environ. 112 (9), 3520–3537.
Ozdogan, M., Rodell, M., Beaudoing, H.K., Toll, D.L., 2010a. Simulating the effects of irrigation over the United States in a land surface model based on satellite-derived agricultural data. J. Hydrometeorol. 11 (1), 171–184.
Ozdogan, M., Yang, Y., Allez, G., Cervantes, C., 2010b. Remote sensing of irrigated agriculture: opportunities and challenges. Remote Sens. 2 (9), 2274–2304.
Pervez, M.S., Brown, J.F., 2010. Mapping irrigated lands at 250-m scale by merging MODIS data and national agricultural statistics. Remote Sens. 2 (10), 2388–2412.
Olofsson, Pontus, Stephen Stehman, V., Curtis Woodcock, E., Damien Sulla-Menashe, Adam Sibley, M., Jared Newell, D., Mark Friedl, A., Herold, Martin, 2012. A global land-cover validation data set part I: fundamental design principles. Int. J. Remote Sens. 33 (18), 5768–5788.
Portmann, F.T., Siebert, S., Döll, P., 2010. MIRCA2000 – Global monthly irrigated and rainfed crop areas around the year 2000: a new high-resolution data set for agricultural and hydrological modeling. Global Biogeochem. Cycles 24 (1).
Quinlan, J.R., 1993. C4. 5: Programs for Machine Learning, Vol. 1. Morgan kaufmann.
Quinlan J.R., Bagging, boosting, and C4. 5. In AAAI/IAAI, vol. 1 (August), 1996 725–730.
Ramankutty, N., Foley, J.A., Norman, J., McSweeney, K., 2002. The global distribution of cultivable lands: current patterns and sensitivity to possible climate change. Global Ecol. Biogeogr. 11 (5), 377–392.
Ramankutty, N., Evan, A.T., Monfreda, C., Foley, J.A., 2008. Farming the planet: 1: Geographic distribution of global agricultural lands in the year 2000. Global Biogeochem. Cycles 22 (1), GB1003.
Rodell, M., Velicogna, I., Famiglietti, J.S., 2009. Satellite-based estimates of groundwater depletion in India. Nature 460 (7258), 999–1002.
Rosenzweig, C., Parry, M.L., 1994. Potential impact of climate change on world food supply. Nature 367 (6459), 133–138.
Rounsevell, M.D.A., Ewert, F., Reginster, I., Leemans, R., Carter, T.R., 2005. Future scenarios of European agricultural land use: II. Projecting changes in cropland and grassland. Agric. Ecosyst. Environ. 107 (2), 117–135.
Scanlon, B.R., Faunt, C.C., Longuevergne, L., Reedy, R.C., Alley, W.M., McGuire, V.L., McMahon, P.B., 2012. Groundwater depletion and sustainability of irrigation in the US high plains and central valley. Proc. Nat. Acad. Sci. 109 (24), 9320–9325.
Schaaf, C.B., Gao, F., Strahler, A.H., Lucht, W., Li, X., Tsang, T., Roy, 2002. First operational BRDF, albedo nadir reflectance products from MODIS. Remote Sens. Environ. 83 (1), 135–148.
Searchinger, T., Heimlich, R., Houghton, R.A., Dong, F., Elobeid, A., Fabiosa, J., Yu, 2008. Use of US croplands for biofuels increases greenhouse gases through emissions from land-use change. Science 319 (5867), 1238–1240.
Shao, Y., Fan, X., Liu, H., Xiao, J., Ross, S., Brisco, B., Staples, 2001. Rice monitoring and production estimation using multitemporal RADARSAT. Remote Sens. Environ. 76 (3), 310–325.
Siebert, S., Döll, P., Hoogeveen, J., Faures, J.M., Frenken, K., Feick, S., 2005. Development and validation of the global map of irrigation areas. Hydrology and Earth System Sciences Discussions Discussions 2 (4), 1299–1327.
Siebert S., Verena H., Karen F., Jacob B. 2013. Update of the Digital Global Map of Irrigation Areas to Version 5, October, 1–171.
Thenkabail, P.S., 1999. An Irrigated Area Map of the World (1999), Derived from Remote Sensing, Vol. 105. IWMI.
Thornthwaite, C.W., 1948. An approach toward a rational classification of climate. Geogr. Rev. 38 (1), 55–94.
Tilman, D., Fargione, J., Wolff, B., D'Antonio, C., Dobson, A., Howarth, R., Swackhamer, 2001. Forecasting agriculturally driven global environmental change. Science 292 (5515), 281–284.
United States Department of Agriculture Foreign Agriculture Service (USDA FAS). 2012. http://www.pecad.fas.usda.gov/(accessed September 2012).
Wan, Z., 2008. New refinements and validation of the MODIS land-surface temperature/emissivity products. Remote Sens. Environ. 112 (1), 59–74.
Willmott, C.J., Feddema, J.J., 1992. A more rational climatic moisture index*. Prof. Geographer 44 (1), 84–88.
Wood, E.F., Roundy, J.K., Troy, T.J., Van Beek, L.P.H., Bierkens, M.F., Blyth, E., Whitehead, P., 2011. Hyperresolution global land surface modeling: meeting a grand challenge for monitoring Earth's terrestrial water. Water Resour. Res. 47, 5.
Wriedt, G., van der Velde, M., Aloe, A., Bouraoui, F., 2009a. A European irrigation map for spatially distributed agricultural modelling. Agric. Water Manage. 96 (5), 771–789.
Wriedt, G., Van der Velde, M., Aloe, A., Bouraoui, F., 2009b. Estimating irrigation water requirements in Europe. J. Hydrol. 373 (3), 527–544.
Xiao, X., Boles, S., Liu, J., Zhuang, D., Frolking, S., Li, C., Moore III, B., 2005. Mapping paddy rice agriculture in southern China using multi-temporal MODIS images. Remote Sens. Environ. 95 (4), 480–492.
Liangzhi, You, Wood, Stanley, Wood-Sichra, Ulrike, Wu, Wenbin, 2014. Generating global crop distribution maps: from census to grid. Agricultural Systems, 127. Elsevier, pp. 53–60.