Remote Sensing 2010, 2, 2388-2412; doi:10.3390/rs2102388

Remote Sensing
ISSN 2072-4292
www.mdpi.com/journal/remotesensing

Article

Mapping Irrigated Lands at 250-m Scale by Merging MODIS Data and National Agricultural Statistics

Md Shahriar Pervez 1,* and Jesslyn F. Brown 2

1 Stinger Ghaffarian Technologies, contractor to the U.S. Geological Survey, Earth Resources Observation and Science Center, 47914 252nd St, Sioux Falls, SD 57198, USA
2 Earth Resources Observation and Science Center, U.S. Geological Survey, 47914 252nd St, Sioux Falls, SD 57198, USA; E-Mail: jfbrown@usgs.gov

* Author to whom correspondence should be addressed; E-Mail: spervez@usgs.gov; Tel.: +1-605-594-6838; Fax: +1-605-594-6529.

Received: 26 August 2010; in revised form: 14 October 2010 / Accepted: 15 October 2010 / Published: 19 October 2010

Abstract: Accurate geospatial information on the extent of irrigated land improves our understanding of agricultural water use, local land surface processes, conservation or depletion of water resources, and components of the hydrologic budget. We have developed a method in a geospatial modeling framework that assimilates irrigation statistics with remotely sensed parameters describing vegetation growth conditions in areas with agricultural land cover to spatially identify irrigated lands at 250-m cell size across the conterminous United States for 2002. The geospatial model result, known as the Moderate Resolution Imaging Spectroradiometer (MODIS) Irrigated Agriculture Dataset (MIrAD-US), identified irrigated lands with reasonable accuracy in California and semiarid Great Plains states with overall accuracies of 92% and 75% and kappa statistics of 0.75 and 0.51, respectively. A quantitative accuracy assessment of MIrAD-US for the eastern region has not yet been conducted, and qualitative assessment shows that model improvements are needed for the humid eastern regions where the distinction in annual peak NDVI between irrigated and non-irrigated crops is minimal and county sizes are relatively small. This modeling approach enables consistent mapping of irrigated lands based upon USDA irrigation statistics and should lead to better understanding of spatial trends in irrigated lands across the conterminous United States. An improved version of the model with revised datasets is planned and will employ 2007 USDA irrigation statistics.
Keywords: irrigated area maps; irrigated agriculture; USDA irrigation statistics; MODIS NDVI; geospatial modeling

1. Introduction

Irrigated agriculture has played a vital role in the economic and social development of the United States. The current areal extent of irrigated lands in the United States is around 22.7 million ha, roughly 14% of the country’s total cropland [1]. The county-level irrigated area acreage has been well documented on a regular basis in the Census of Agriculture by the U.S. Department of Agriculture (USDA) National Agricultural Statistics Service (NASS). However, geographic information on the location and spatial distribution of irrigated areas within each county has not been regularly mapped with sub-county detail. Here we present a methodology for nationally consistent mapping of irrigated croplands directly tied to the 2002 USDA-NASS statistics at a 250-m resolution [2]. This resolution enables more detailed map results than previous irrigation mapping attempts and represents actual field sizes across much of the country.

Accurate, detailed, geospatial information on irrigated croplands is essential for answering many Earth system science, climate change, and water supply questions [3]. Irrigation increases the evapotranspiration and soil moisture and modifies properties that influence the interaction between the land and atmosphere [4], which in turn influences the energy budget at the land-atmosphere interface [5]. In a changing environment and with a growing population, fresh water may become a limited resource due to rising urban demand and water conservation efforts. In the United States requirements for irrigation water will likely increase to feed a rising population and support the production of biofuels for energy (e.g., production of ethanol from corn). Climate change may put further strain on a limited freshwater resource. Although the irrigation water use in the United States has been relatively stable since 1980, it remains the second largest use of water after thermoelectric. Irrigation water use accounted for 31% (484.5 million m³ per day) of the total water withdrawals in 2005 [6]. In a warmer climate, changes in precipitation will have a major impact on the hydrologic cycle and, subsequently, on the agricultural food production both from irrigated and rain-fed crops [7]. Future changes in precipitation are projected and many models predict increases in extreme events rather than a change in average precipitation amounts [8-10]. Accurate knowledge of the spatial distribution of irrigated lands at the national scale is essential to assess probable impacts of these extreme events on agricultural food production.

During the last decade, several national-scale irrigation maps have been created. Some were produced as part of global irrigated area mapping efforts using country-level statistics that were often outdated or mapped area equipped for irrigation rather than actual area irrigated [11,12]. Other maps used coarse-resolution (10-km to 500-m) multi-sensor satellite and climate data with methods optimized for global classification [3,13]. Although multiple global land cover classification efforts [14-16] have mapped national-level agricultural croplands, they were derived from coarse resolution (1-km or coarser) remotely sensed data and many did not differentiate irrigated lands from
general agricultural land cover. For the conterminous United States, a finer resolution (30-m) National Land Cover Dataset (NLCD) [17] also did not identify irrigated agriculture as a separate class.

Satellite observations provide reliable, economical, and synoptic data of the Earth’s surface. These data contribute to mapping land cover, including agricultural lands. Existing methods for agricultural land cover characterization have often been derived through image classification techniques. However, the variety of irrigated crops and the spatial patterns of their phenology require multi-temporal, consistent, composite vegetation growth information with sufficient spatial detail, along with a rich library of field reference training and ancillary data (e.g., climate and topography), to classify irrigated lands using satellite observations. Obtaining these parameters consistently at national scale is a major challenge. In this paper, we describe a method that assimilates reliable published acreage statistics for irrigated agriculture collected and documented by USDA-NASS with annual satellite-derived vegetation index information derived from 250-m Moderate Resolution Imaging Spectroradiometer (MODIS) to spatially map irrigated areas by county for the conterminous United States. We also present comparisons of the resulting irrigation map for 2002 with irrigation ground reference information for California and the central Great Plains.

2. Background

2.1. Irrigation and Irrigation Water Use in the United States

In 2007, 18.3% of the harvested cropland in the conterminous United States was irrigated. However, irrigated cropland accounts for 22% of the major crops’ production and nearly half of the value of all crops sold [1,18]. Crop yields are usually higher when the crops are fully irrigated compared to yields from non-irrigated crops when water requirements are not met. For example, in 2007, irrigated barley and wheat yields were twice as much as yields of non-irrigated barley and wheat [1]. The density of irrigated lands varies greatly from west to east. Most agriculture in the West is irrigated because of its arid climate. Nearly 75% of irrigated lands are concentrated in 17 western states where average annual precipitation is less than 500 mm [1,6]. However, irrigation has been increasing in the southeastern United States because of intensified drought during the past decade [19].

In the western United States, average farm sizes are relatively large [2,20] and over 36% of the farms in the west had irrigated lands in 2002. These large western farms (with ≥ USD 250,000 in annual farm sales) account for 61% of irrigated lands, and consume 66% of the total farm water applied [21]. This implies that irrigation occurs in a relatively large spatial domain, or pattern, in the western United States; similar findings were also supported by [2]. The irrigation in the western United States is comprehensive, structured, and consistent, resulting in less temporal and spatial dynamics. Conversely, irrigation in the humid eastern and southeastern United States is generally supplemental and varies spatially.

Since 1950, the total U.S. irrigated area doubled, converting approximately 11 million ha of non-irrigated land into irrigated agriculture. Associated with population increase, the irrigated area steadily increased from 11.7 million ha in 1954 to 19.8 million ha in 1982. The irrigated area declined slightly during the 1980s then increased at a lower rate to 22.2 million ha in 1997. Irrigated area has
remained relatively constant since then (Figure 1(a)); however, the estimated area of total croplands has decreased nationally since 2000 [1,63-68].

Irrigation water use and total water withdrawal increased steadily from 1950 to 1980 (Figure 1(b)). Irrigation water use peaked in 1980 at 567.8 million m$^3$ per day and represented 35% of the total withdrawal (1.6 billion m$^3$ per day). Since 1980, irrigation water use has decreased, except in 2000. The prevailing dry condition through much of the country in 2000 led to the increase in irrigation water use and the total water withdrawal in 2000 [6]. Although the area of irrigated lands increased by 2.8 million ha since 1980, these lands were irrigated with 15% less water. The average application rate for irrigation water has declined steadily from 1.05 ha-m per ha in 1980 to 0.77 ha-m per ha in 2005 (application rates were calculated using irrigation area estimates of USDA-NASS and irrigation water use provided by [6]). This decline is attributed to climate, energy costs, shifts in crop genetics, improvements in irrigation systems, and optimized application of water [6,22]. Surface water historically has been the prime source for irrigation, but irrigation expansion in the central United States since 1970 has gradually increased the groundwater contribution to irrigation. In 2005, groundwater contributed to 42% of irrigation withdrawals nationally [6].

**Figure 1.** Trends in (a) Population and irrigated lands, and (b) Total water withdrawal and irrigation water use in the United States [1,6].

2.2. **Irrigation Statistics in the United States**

Our method of mapping irrigated areas is directly associated with the irrigated acreage reported by the 2002 USDA NASS Census of Agriculture. Historically, USDA NASS publishes the total number of acres of irrigated land by county administrative unit every 5 years (e.g., 1997, 2002, and 2007) in the Census of Agriculture. NASS surveyed all the farming operations that produce or would normally produce and sell USD 1,000 or more via questionnaire and compiled the collected information in the Census of Agriculture. They ensure that their statistics are reliable. A complete description of their methodology can be found in the 2002 Census (see Appendix C in the 2002 Census of Agriculture: Volume 1, Geographic Area Series). The questionnaire contained several questions addressing whether a farm had any irrigated acres and, if so, how many of those acres were harvested land and how many were pastureland or rangeland.
No exact error estimates for irrigated land statistics were published in the 2002 Census of Agriculture, so we are constrained to extrapolating from the general published errors. USDA publishes certain error estimates for U.S. totals and by state, but they are not available by county (Appendix C Table B of 2002 Census of Agriculture). For example, the relative RMSE for the land in farms (U.S. Totals) is 0.64%. From this, we have assumed that for the United States, a relative RMSE of 0.64% when applied to the total number of acres of irrigated lands for 2002 (22.4 million ha) results in a possible error up to 144,000 ha in the number of irrigated acres across the United States.

2.3. Existing Geospatial Irrigation Maps

Historical records of county-level irrigation acreage have been made available for many years by the USDA in the Census of Agriculture, but there have been very few attempts to map the spatial distribution of these irrigated areas at a higher level of precision than individual counties. The Kassel digital Global Map of Irrigated Areas (GMIA) [12] was the first raster dataset showing the percentage of each 0.5° × 0.5° cell area that was equipped for irrigation in 1995. The map was subsequently improved and upgraded to 5′ × 5′ cell size for 2000 [11].

When the high temporal resolution remotely sensed Earth observation datasets with continental and global coverage became available consistently, they were used in efforts to classify land cover (including agricultural lands). The U.S Geological Survey (USGS) Global Land Cover Characteristics (GLCC) dataset [14] identified four types of irrigated croplands for an 18-month period in 1992 and 1993—hot irrigated cropland, cool irrigated cropland, irrigated grass and cropland, and rice paddy and field—along with various other land cover types globally from 1-km Advanced Very High Resolution Radiometer (AVHRR) sensor data. Multi-temporal unsupervised classification method was used to produce the map. The Global Irrigated Area Map (GIAM) [13] was produced by the International Water Management Institute using a multi-resolution blend of satellite Earth observation, topography, and climate data using a robust unsupervised classification technique with post-classification refinement. The final map identified fractional irrigated areas for each 10-km unit presented in 28 irrigation classes for 1999. The United States was included in these datasets as part of their global application. A new optimized supervised classification technique [3] using MODIS time series and climate and agricultural extent data to map irrigated areas for the conterminous United States was identified, but the technique is also suitable for global application. Their map identified fractional irrigated areas for each 500-m area.

These datasets present irrigated areas as percentage of the pixel’s unit area (unit area varies between 100 km² and 250,000 m²), which does not provide information on the subpixel location of the irrigated lands. Although estimation of fractional irrigated areas may satisfy quantification of spatially averaged parameters over large areas for Earth science research, it is not detailed enough to answer questions about local level water, carbon, or energy cycles requiring identification of irrigated areas with respect to their true spatial locations. Another problem for datasets developed using image classification techniques is inadequate reference training data. GIAM used 1,790 sites surveyed mostly in Asia and Africa for post-classification refinement and accuracy assessment, which makes the map highly optimized for these regions but less reliable for the rest of the area [3]. Both the GIAM- and MODIS-based maps heavily relied on high-resolution remote sensing Earth observations for class
labeling and obtaining training datasets for spectral classification and regression analysis. These Earth observations of the landscape are from one point in time and may not capture irrigation that occurred at other times, and this may lead to errors in class labeling in those maps. In terms of area estimates, these maps overestimated irrigated areas for the conterminous United States compared to the estimate provided by USDA [2,3]. Irrigated area estimates mapped by GIAM were even higher than the area equipped for irrigation for the United States (estimates were provided in Table 3 of [13]). Although there was a time lag between the two estimates, the mixed classes with ambiguous members in an unsupervised classification may have caused overestimation of irrigated areas in GIAM. The more recent MODIS-based map overestimated the irrigated areas by 10% for the conterminous United States (estimates were provided in Table 3 of [3]).

The method presented in this article differs from the above methods by incorporating published irrigation areal statistics into our modeling approach rather than deriving irrigated lands using image classification techniques. Our objective was to map irrigated lands using well-validated, statistically robust, survey-based data integrated with satellite remote sensing techniques to reduce the uncertainties in the derived map and provide a repeatable, consistent, and cost-effective means of mapping irrigated areas at a national scale.

3. Materials and Method

We present a robust yet simple and easy to implement geospatial model to create a map of irrigated areas, hereafter referred to as the MODIS Irrigated Agriculture Dataset for the United States (MIrAD-US). The method uses three primary data inputs [2]:

a. USDA county-level irrigation area statistics for 2002
b. Annual peak MODIS Normalized Difference Vegetation Index (NDVI) (a proxy for maximum vegetation growth; [2])
c. A land cover mask for agricultural lands derived from NLCD 2001 [17]

A schematic diagram of our mapping methodology is shown in Figure 2. The success of our modeling approach is tied to the following three hypotheses, which are further discussed below.

a. Irrigated crops have higher annual peak NDVI values than non-irrigated crops in the same county.
b. The growing season peak NDVI, at any time it occurs, will vary for each crop and for each geographic region of the United States.
c. The difference in NDVI between irrigated and non-irrigated crops will be enhanced under non-optimal precipitation conditions (e.g., drought).
3.1. Input Data

County Irrigation Statistics

The county irrigation statistics were the most important input variable in the MIrAD-US model. The 2002 Census of Agriculture [20] provided the number of irrigated acres for each of the 3,114 counties in the conterminous United States. The modeling methodology essentially reconstructs the Census of Agriculture irrigated acreage in a spatially-distributed fashion guided by the peak MODIS NDVI and the NLCD agriculture land cover. The Census of Agriculture statistics are the best available estimates of irrigated acreage for the entire United States from a single source. However, as discussed in section 2.2, there is the potential for error within the Census statistics. The reported 0.64% RMSE of the total statistical database for land in farms was a potential source of uncertainty in the derived map of irrigated areas. However, as a federally published national database, we consider the USDA statistics to be the best consistent available data on irrigation.

MODIS Annual Peak NDVI

The NDVI is a commonly used vegetation index that has been demonstrated in the remote sensing literature as a proxy measure for absorbed photosynthetically active radiation. The NDVI represents a dimensionless, radiometric measure shown to be correlated with the relative condition and amount of green vegetation [23,24] by means of the differential response of incident visible red (absorbed by leaf chlorophyll) and near-infrared (reflected by spongy mesophyll and green leaf biomass) reflectance.
properties of the vegetation canopy. Prior studies have shown time-series NDVI observations are linked to phenological signals [25-27] and biophysical vegetation characteristics over different land cover types (e.g., leaf area index and biomass) [28-32]. Positive correlations between NDVI and precipitation [33-36] have indicated that increasing available moisture for vegetation also increases the NDVI over many different cover types, including grasslands, shrubs, and crops [25,37]. Consistent with this legacy research, a maximum NDVI in an annual time series for a location is a proxy for the peak level of photosynthetic activity, the highest biomass, and possibly the densest vegetation cover in the canopy [38-40] but over certain land cover types (e.g., deciduous forest and crops), the NDVI often saturates over high-density vegetation. It has also been observed that irrigated crops exhibit higher (maximum) NDVI than non-irrigated crops, especially for corn (maize) and wheat [41,42]. Based on this heritage research and our own investigations, we assume that the highest annual peak NDVI for any agricultural crop is the result of consistent adequate soil moisture as is delivered by irrigation throughout the growing season; therefore, the maximum NDVI for irrigated crops generally exceeds the peak NDVI for non-irrigated crops. An analysis of crop specific time-series NDVI for irrigated and non-irrigated corn, dry beans, millet, pasture, winter wheat, and alfalfa for 2002 and 2006 in western Nebraska confirms a higher peak NDVI of irrigated crops compared to the peak NDVI of non-irrigated corps. Figure 3 presents examples of the crop NDVI time-series analysis for six sites and supports this assumption.

Single cropping (i.e., one crop grown during one annual period) is the prevalent agricultural practice in the United States. However, double cropping has been practiced in California, the Great Plains and some Midwestern states (e.g., winter wheat followed by soybeans). Even triple/quad cropping also can be found for high value crops on selected fields in California and Arizona. In our method, the calculation of the annual peak NDVI simply selected the highest value irrespective of cropping intensity.

The MODIS instrument has radiometric and geometric properties designed to collect and construct global science-quality remotely sensed data with high temporal frequency (standard 8-day and 16-day surface reflectance products) [43,44]. Since the first MODIS instrument was launched aboard the Terra platform in 1999, MODIS data have been used by many researchers for agricultural applications, including mapping crops, estimating crop yields, describing crop phenology, disaster monitoring and mapping irrigated agriculture [3,41,45-49]. Another feature of MODIS is its cost effectiveness. Data are available free-of-charge to the users.

The annual peak (or maximum) NDVI for 2002 was extracted from annual time series of temporally smoothed MODIS 16-day composite NDVI data. The MODIS data products of MOD13Q1 MODIS/Terra Vegetation Indices 16-Day L3 Global 250m SIN Grid V005—Collection 5 were downloaded from the Land Processes Distributed Active Archive Center (LP DAAC), NDVI data layers were extracted and reprojected to Lambert Azimuthal Equal Area projection and an annual time series of NDVI was created. The NDVI is affected by number of phenomena including cloud contamination, atmospheric perturbations, variable viewing geometry of the sensor, and imperfect sensor calibration. All of these tend to enhance the difference in NDVI between two composite periods that often not reflect a real change in vegetation condition. To minimize this disturbance in temporal profile of vegetation growth signal, a temporal smoothing was applied on the NDVI time series data.
The smoothing employs a weighted least-square technique \([51]\) and uses a moving temporal window to calculate a family of regression lines that are associated with each observation. The family of lines is then averaged at each point and interpolated between points where a weighting factor favoring local peaks (high value) to produce a continuous temporally smoothed NDVI signal. The annual peak NDVI layer was then created by calculating the maximum NDVI from the annual time series for each pixel.

**Figure 3.** Examples of crop specific smoothed NDVI time series and peaks for 2002 and 2006 at three sites in Scotts Bluff and Banner counties in western Nebraska, (a) irrigated and non-irrigated corn in both years at 103.44W 41.75N irrigated, 103.44W 41.73N non-irrigated sites, (b) irrigated and non-irrigated dry beans in 2002 and irrigated dry beans and non-irrigated corn in 2006 at 103.45W 41.81N irrigated, 103.73W 41.98N non-irrigated sites (c) irrigated and non-irrigated pasture in 2002 and irrigated corn and non-irrigated millet in 2006 at 103.82W 41.55N irrigated, 103.65W 41.43N non-irrigated sites, (d) is the mean time-series NDVI for a, b and c. The sites were selected on the basis of irrigated and non-irrigated ground reference points surveyed in 2006 and the crop types were obtained from NASS Cropland Data Layers (CDL) \([50]\).
Our methodology incorporated a land cover mask to constrain the selection of irrigated areas within agricultural land cover. The land cover mask was derived from the 2001 NLCD [17]. The 30-m cell size of NLCD was resampled to 250-m cell size, and the dominant land cover type in each cell was identified through block majority geospatial techniques. Tests determined that annual peak NDVI from forest and woody wetlands could potentially be higher than the annual peak NDVI from croplands including Pasture/Hay (a list of the peak annual MODIS NDVI by land cover type is provided by [2]). Therefore, a land cover mask was required to mask out non-agricultural land cover areas to avoid mapping of irrigation over vegetative non-agricultural lands such as forest, woody wetlands, or golf courses.

From NLCD land cover classes, only Pasture/Hay and Cultivated crops were considered agricultural lands, and areas or cells that were not Pasture/Hay or Cultivated crops were successively eliminated or masked out (made null) at the MODIS annual peak NDVI data layer before ingesting them into the model. The overall thematic accuracy of 2001 NLCD was 85.3% for Anderson Level 1 classes with user’s accuracy of 82% for cropland category [69]. However, a modest agreement ($r^2 = 0.65$) of area by county between 2002 Census estimates of farmlands and 2001 NLCD estimates of agricultural lands suggests a possible source of uncertainty in the modeled result in the irrigated area map.

3.2. Description of the Geospatial Model

In a geospatial modeling framework, we identified cells using a threshold based on the annual peak NDVI within agricultural land cover that comprise an equivalent target area by county provided by the Census of Agriculture (Figure 4). The model was implemented at a county spatial domain and initiated by creating an ordered list of unique annual peak NDVI values from within the extent of agricultural land cover. The peak NDVI value list was sorted in descending order. In the first iteration, cells with the highest annual peak NDVI value were identified and mapped and the accumulated area covered by those pixels was calculated. The accumulated area was then compared with the target area provided by the Census. If the total area of the selected cells was less than the target area, then cells with the next highest peak NDVI were taken from the sorted list and corresponding cells having that value were identified and appended with the previous map. The cumulative area of the cells identified by the two highest peak NDVI values was calculated. The cumulative area was again compared against the target area. These steps were repeated until the target area for the county was exceeded and a final county map of irrigated areas was created.
Figure 4. Processes illustration of the Moderate Resolution Imaging Spectroradiometer (MODIS) Irrigated Agriculture Dataset (MIrAD-US) geospatial model; the USDA irrig. acg. refers to the county irrigation acreage tabular data at 2002 USDA Census of Agriculture; value of n determines number of iterations.

In principle, we have calculated model estimates of irrigated areas to match the irrigated area estimates of USDA by varying NDVI threshold for individual counties. The basic strategy relies on the selection of locations with higher peak NDVI for irrigated crops. The size of a typical center pivot irrigation is about 500,000 m$^2$ or 50 ha [52]. Thus, the MODIS resolution of 250-m (62,500 m$^2$) offers adequate precision to identify all major irrigated cells. However, model dependency on the USDA irrigation estimates means that inaccuracies in these estimates will lead to inaccuracies in the model results.

Since, the annual peak NDVI will vary by crop and may be further influenced by topographic, and climatic, and management factors, the optimal NDVI threshold to separate irrigated from non-irrigated cells needs to be locally selected. Improvements to the model might be realized by introducing a crop type map for 2002. However, no wall-to-wall U.S. crop map is currently available for 2002 at the 250-m resolution.

The irrigation mapping year of 2002 was climatologically a dry year with below average precipitation across the conterminous U.S. [53,54]. This provided the enhanced difference in peak NDVI between irrigated and non-irrigated crops suitable for this modeling approach.

In the final step, all the county outputs were mosaicked together to create a seamless 250-m raster data product of MIrAD-US. All lone pixels from the raster product were filtered out based on the
assumption that in the United States irrigated fields are generally larger than 62,500 m². This model will therefore not resolve very small, isolated irrigated fields.

4. MIrAD-US for the Conterminous United States

4.1. Results

The MIrAD-US dataset (Figure 5) is a two-class data product presenting irrigated and non-irrigated classes. The map presents a geospatially specific rendition of the reported irrigated areas from the 2002 Census of Agriculture at 250-m resolution, and each cell identified as irrigated in the data product was considered entirely irrigated. Unlike other remotely sensed irrigation maps, the MIrAD-US overcomes the spatial issues presented in subpixel fractional irrigated areas, but at the same time, it may have trouble identifying irrigated fields that are smaller than a 250-m cell.

In the MIrAD-US, most of the major irrigation-dominated areas across the conterminous United States are located in the central valley of California, the Snake River Basin in Idaho, the Columbia Basin of the interior Northwest, the Ogallala Aquifer in the central Plains, and the Mississippi Flood Plains, with more sparsely scattered irrigation located along the east and southeast coasts (Figure 5). Most of the irrigated areas were concentrated in the western United States, but sparsely scattered irrigated fields were identified throughout the more humid eastern seaboard and Mississippi Flood Plain.

Figure 5. The Moderate Resolution Imaging Spectroradiometer (MODIS) Irrigated Agriculture Dataset (MIrAD-US) showing the spatial distribution of irrigated lands in light gray color identified in a geospatial modeling framework using USDA irrigation statistics for 2002, MODIS annual peak NDVI, and 2001 NLCD

The MIrAD-US mapped the targeted irrigated areas well in the western and Great Plains regions, but the difference between the targeted irrigated area and mapped irrigated area was relatively high in the humid Eastern regions (Table 1), suggesting that MIrAD-US may be more accurate in the West than in the East. Although the county statistics published by the Census of Agriculture provided the
target for the MIrAD-US model, the actual area identified in MIrAD-US was approximately 1.5% higher for the continental United States. This is because the model selects the NDVI peak threshold for each county that just exceeds the irrigated area indicated by the USDA statistics. This reduces the precision of the model especially in counties with small amounts of irrigation. Table 1 presents a state-by-state comparison of irrigation area estimates between USDA and MIrAD-US. The differences between the estimates were low for most of the highly irrigated western states, but they were relatively high for the eastern states. The amounts of irrigated lands were comparatively low and scattered in small patches in most of these eastern states, and county sizes were relatively small. The filtering of single cells identified as irrigated also contributed high differences between the estimates for few eastern states. At the cell size of 250-m, the NDVI threshold-based separation of irrigated and non-irrigated cell areas are more effective when the target mapping area is approximately 200,000 ha or more by state and in the regions where differences in NDVI between irrigated and non-irrigated crops are more distinct.

<table>
<thead>
<tr>
<th>States</th>
<th>Irrigated lands by source in ha</th>
<th>% Difference</th>
<th>States</th>
<th>Irrigated lands by source in ha</th>
<th>% Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>USDA</td>
<td>MIrAD-US</td>
<td></td>
<td>USDA</td>
<td>MIrAD-US</td>
<td></td>
</tr>
<tr>
<td>California</td>
<td>3,524,550</td>
<td>3,507,406</td>
<td>Illinois</td>
<td>158,169</td>
<td>190,113</td>
</tr>
<tr>
<td>Nebraska</td>
<td>3,085,797</td>
<td>3,192,381</td>
<td>Wisconsin</td>
<td>156,169</td>
<td>164,919</td>
</tr>
<tr>
<td>Texas</td>
<td>2,053,633</td>
<td>2,058,038</td>
<td>Indiana</td>
<td>126,719</td>
<td>143,713</td>
</tr>
<tr>
<td>Arkansas</td>
<td>1,679,351</td>
<td>1,781,975</td>
<td>North Carolina</td>
<td>106,860</td>
<td>89,850</td>
</tr>
<tr>
<td>Idaho</td>
<td>1,330,818</td>
<td>1,352,225</td>
<td>North Dakota</td>
<td>82,077</td>
<td>97,669</td>
</tr>
<tr>
<td>Kansas</td>
<td>1,083,860</td>
<td>1,110,800</td>
<td>Iowa</td>
<td>57,509</td>
<td>85,325</td>
</tr>
<tr>
<td>Colorado</td>
<td>1,048,400</td>
<td>1,041,769</td>
<td>Alabama</td>
<td>44,023</td>
<td>28,506</td>
</tr>
<tr>
<td>Montana</td>
<td>799,704</td>
<td>818,656</td>
<td>Virginia</td>
<td>40,029</td>
<td>30,700</td>
</tr>
<tr>
<td>Oregon</td>
<td>771,989</td>
<td>772,456</td>
<td>Delaware</td>
<td>39,322</td>
<td>35,463</td>
</tr>
<tr>
<td>Washington</td>
<td>737,805</td>
<td>743,531</td>
<td>New Jersey</td>
<td>39,211</td>
<td>36,400</td>
</tr>
<tr>
<td>Florida</td>
<td>734,575</td>
<td>746,100</td>
<td>South Carolina</td>
<td>38,705</td>
<td>27,356</td>
</tr>
<tr>
<td>Wyoming</td>
<td>623,899</td>
<td>602,763</td>
<td>Maryland</td>
<td>32,710</td>
<td>30,400</td>
</tr>
<tr>
<td>Mississippi</td>
<td>475,720</td>
<td>501,081</td>
<td>New York</td>
<td>30,215</td>
<td>25,644</td>
</tr>
<tr>
<td>Utah</td>
<td>441,516</td>
<td>438,700</td>
<td>Tennessee</td>
<td>24,774</td>
<td>19,681</td>
</tr>
<tr>
<td>Missouri</td>
<td>418,029</td>
<td>444,300</td>
<td>Pennsylvania</td>
<td>17,206</td>
<td>12,381</td>
</tr>
<tr>
<td>Louisiana</td>
<td>379,935</td>
<td>396,569</td>
<td>Ohio</td>
<td>16,465</td>
<td>14,106</td>
</tr>
<tr>
<td>Arizona</td>
<td>377,060</td>
<td>376,156</td>
<td>Kentucky</td>
<td>14,873</td>
<td>10,081</td>
</tr>
<tr>
<td>Georgia</td>
<td>352,404</td>
<td>318,831</td>
<td>Massachusetts</td>
<td>9,599</td>
<td>6,838</td>
</tr>
<tr>
<td>New Mexico</td>
<td>341,878</td>
<td>337,675</td>
<td>Maine</td>
<td>7,974</td>
<td>7,131</td>
</tr>
<tr>
<td>Nevada</td>
<td>302,160</td>
<td>291,369</td>
<td>Connecticut</td>
<td>4,103</td>
<td>2,938</td>
</tr>
<tr>
<td>Oklahoma</td>
<td>209,446</td>
<td>207,000</td>
<td>Rhode Island</td>
<td>1,604</td>
<td>931</td>
</tr>
<tr>
<td>Michigan</td>
<td>184,649</td>
<td>197,419</td>
<td>Vermont</td>
<td>945</td>
<td>500</td>
</tr>
<tr>
<td>Minnesota</td>
<td>184,071</td>
<td>212,450</td>
<td>New Hampshire</td>
<td>928</td>
<td>356</td>
</tr>
<tr>
<td>South Dakota</td>
<td>162,313</td>
<td>187,619</td>
<td>West Virginia</td>
<td>802</td>
<td>431</td>
</tr>
</tbody>
</table>

Total U.S. | 22,354,552 | 22,698,700 | 1.5
4.2. Validation

The validation of MIrAD-US was a major challenge because of inadequate comprehensive and timely historical ground observations. While technically feasible, intercomparison with existing remote sensing-derived maps of irrigated lands available for the United States would not necessarily provide validation of the MIrAD-US, because those maps themselves typically have not been comprehensively validated against actual ground observations. In addition, potential discrepancies would be expected because of underlying differences in processing techniques, input data, methods, and map legend definition implemented in creation of the referenced geospatial data products [55]. A comprehensive ground verification survey was outside the scope of this effort; therefore, we relied on the best available geospatial irrigated area information based on ground surveys from multiple sources for the accuracy assessment of MIrAD-US. We have evaluated MIrAD-US both quantitatively and qualitatively. Quantitative assessment was based on land use data provided by California Department of Water Resources (DWR) and irrigation ground reference points collected by the University of North Dakota (UND) in Nebraska, Kansas, Oklahoma, and Texas. No irrigation ground reference points for 2002 were found for the eastern United States, thus a qualitative evaluation was performed by comparing MIrAD-US with August 2002 Landsat images for the eastern states of Missouri and Florida.

Accuracy Assessment of MIrAD-US in California

The California DWR conducts detailed land use surveys in California to map agricultural lands each year. In these surveys, field-level data on crop categories, irrigation methods, and water sources are well documented. However, not every county is surveyed every year. Each survey begins with delineation of agricultural field boundaries from aerial photographs and high-resolution satellite images. Then, over 95% of these fields are visited by a DWR staff member for visual identification of the land use. Once the field work is complete and land use attributes with irrigation information are added with the spatial (vector polygon) field unit, a digital composite vector field map of the survey area is created with appropriate metadata. After quality checking, these field maps are posted at a DWR Web site (http://www.water.ca.gov/landwateruse/lusrvymain.cfm). We downloaded vector field boundaries for 19 counties (Figure 6(a)) that were considered representative for the state. The irrigation and land use information for the fields in these counties were acquired in a 5-year period (2000 to 2004). As total irrigated area did not vary substantially over these 5 years in California, we assumed that ground observations collected across a 5-year period bracketing the year 2002 would have a minimal influence on the assessment.

We used a random sampling technique to quantify the agreement between MIrAD-US and DWR irrigated areas. We converted DWR surveyed polygon fields into grids using an irrigation attribute (1 for irrigated and 0 for non-irrigated) matching the spatial properties (projection, extent, and cell size) of the MIrAD-US so that both data products fully overlay each other on a cell-by-cell basis. Thus, the MIrAD-US and DWR raster maps were made to have the same attribute values (e.g., 1 if the cell is irrigated and 0 if the cell is not irrigated). We used a random point generator to create 5 sets of sample points constrained by the common areas from both data products and with a minimum spacing of 1 km between any two points. These points were overlaid on both data products, and irrigation information
was extracted from the cell that contained the point. Approximately 22% of the points were in the irrigated area and 78% of the points were in the non-irrigated areas. The numbers of these sample points from each type (irrigated and non-irrigated) were directly proportionate to the area they cover in the landscape. Finally, an error matrix was generated using the extracted irrigation information from MIrAD-US and DWR for each of those set of points. The results from the error matrix are presented in Table 2. Five sets of points were generated to understand the influence of the number of samples in a random sampling. The numbers of samples varied from 1,000 to 5,000 with point density varying 1 point per 96.7 km$^2$ to 19.3 km$^2$. Test results implied that the total number of samples in random sampling had a negligible impact on the derived statistics.

Table 2. Error matrix summary of irrigated lands between MIrAD-US and DWR fields map for California.

<table>
<thead>
<tr>
<th>Sample # / density per point in km$^2$</th>
<th>Classes</th>
<th>Producer’s accuracy</th>
<th>Errors of omission</th>
<th>User’s accuracy</th>
<th>Errors of commission</th>
<th>Overall accuracy</th>
<th>Kappa stat</th>
</tr>
</thead>
<tbody>
<tr>
<td>3000/32.2</td>
<td>Irrigated</td>
<td>0.75</td>
<td>0.25</td>
<td>0.86</td>
<td>0.14</td>
<td>0.92</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>Non-Irrigated</td>
<td>0.97</td>
<td>0.03</td>
<td>0.94</td>
<td>0.06</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Results presented in Table 2 suggest that the irrigated areas in the MIrAD-US agreed well with DWR irrigated fields with an overall accuracy of 92%. Another measure of agreement, especially for remotely sensed data product, is the Kappa statistic [56]. Kappa values range between 1 and −1, where 1 is perfect agreement, 0 is no agreement beyond agreement chance and −1 is complete disagreement. The Kappa index calculated from the error matrix for California ranged from 0.75 to 0.77, which can be considered a substantial agreement. The producer’s accuracy of around 75% and the user’s accuracy of around 86% suggest that the MIrAD-US have omitted some irrigated areas. In terms of total area, the DWR county field-based maps showed 24% more irrigated area than the irrigated areas mapped in MIrAD-US for these 19 counties. If an irrigated field was isolated and less than 250-m in size or around 6 ha, then it is obvious that the vegetation greenness signal would not be strong enough from that field to be picked up by MODIS annual peak NDVI, and consequently may not be identified as irrigated areas in MIrAD-US. In DWR field surveys, as little as 2,000 m$^2$, or 0.2 ha, of irrigated fields were identified and mapped. Many of these small fields were generalized when DWR vector fields were converted to raster. This difference in spatial detail between the compared data sets might have contributed to this omission error. Another cause for the omission error is the time lag between MIrAD-US and DWR field-based data composite map. MIrAD-US mapped irrigated areas for 2002, whereas the field information acquisition dates of the DWR field map range from 2000 to 2004. Despite very little change in total irrigated area estimates over these 5 years, any crop rotation or changes in irrigation requirements that took place across the landscape over this period would change the spatial distribution pattern of irrigated areas across the landscape which might have resulted in omission error in MIrAD-US map.
**Accuracy Assessment of MIrAD-US in Great Plains**

The second set of ground observations was obtained from UND. In 2006, UND conducted a comprehensive field survey to visually identify irrigated and non-irrigated fields in several high-intensity irrigated areas of the Great Plains. A total of 336 sites were observed during the survey and coordinates of the sites were recorded using a Global Positioning System (GPS) along with multiple digital photographs with photo facing direction and land use as attributes. The survey was conducted between the fourth week of July and first week of August. Figure 6b shows the location of the ground observation sites surveyed by UND.

We digitized the area of the targeted fields for each of these points by overlaying them on August 2006 Landsat images and visually interpreting the image with the help of point attributes (Figure 6(c)). In this fashion, the ground observation points were transformed into area polygons. These area polygons were converted to raster grid with the same spatial properties of MIrAD-US (cell size and extent) so that both grids align perfectly on a cell-by-cell basis. Finally, these ground observations were compared with MIrAD-US in two ways. First, the grid with ground observation information was overlaid on MIrAD-US and compared on a cell-by-cell basis, and an error matrix was produced; second, each of the digitized area polygons was overlaid on the MIrAD-US, and the land use was visually identified from MIrAD-US based on a portion of the maximum coverage (e.g., if more than one half of the polygon was covered by irrigated cells from MIrAD-US, it was considered “irrigated” for MIrAD-US). This land use information from MIrAD-US was then compared with the original ground-observed land use, and a second error matrix was produced. The summary of these two error matrixes is presented in Table 3.

**Table 3.** Error matrix summary of irrigated lands between MIrAD-US and UND ground observation sites for the Great Plains.

<table>
<thead>
<tr>
<th>Comparison type</th>
<th>Producer’s accuracy</th>
<th>Errors of omission</th>
<th>User’s accuracy</th>
<th>Errors of commission</th>
<th>Overall accuracy</th>
<th>Kappa stat</th>
</tr>
</thead>
<tbody>
<tr>
<td>By grid cells</td>
<td>0.67</td>
<td>0.33</td>
<td>0.94</td>
<td>0.06</td>
<td>0.75</td>
<td>0.51</td>
</tr>
<tr>
<td>By polygons</td>
<td>0.79</td>
<td>0.21</td>
<td>0.94</td>
<td>0.06</td>
<td>0.81</td>
<td>0.56</td>
</tr>
</tbody>
</table>

Although lower than the agreement results for California, the agreement between MIrAD-US and ground observations was reasonable for the Great Plains with overall accuracy of 75% or 81% depending on the comparison methods. The Kappa statistic of 0.51 or 0.56 also suggests a fair agreement. Comparison by polygon may be relatively skewed because of generalization of the information obtained from the MIrAD-US for each of those ground observation polygons. Relatively high errors of omission and low errors of commission suggest that, in the regions where the ground data were collected, there may have been less irrigated fields mapped by MIrAD-US. This could be due to the discrepancy between MIrAD-US date (2002) and field survey dates (2006). Some of the high-intensity irrigation in the United States occurs in the Great Plains, which is primarily fed by the Ogallala Aquifer. The irrigated area did increase in Nebraska, Kansas, Oklahoma, and Texas by 392,890 ha, or 6.1% between 2002 and 2007 [1].
Figure 6. Showing the spatial distribution of irrigation ground truth data used in the accuracy assessment, (a) California Department of Water Resources field boundaries for 19 counties surveyed between 2000 and 2004, gray shaded fields were irrigated and void outlined fields were non-irrigated, (b) black dots were the ground truth points surveyed by University of North Dakota during July–August 2006, and (c) a polygon was drawn for each of the ground truth points by overlaying them on August 2006 Landsat images and using the point attributes taken during the survey, these polygon areas were used in the accuracy assessment instead of the points directly.

Qualitative Assessment of MIrAD-US

In Figure 7, we present the MIrAD-US irrigated lands side-by-side with Landsat images for six major irrigation sites (the MIrAD-US is in the first and third column and Landsat image views of the landscape is in the second and fourth column). The Landsat images were acquired during August 2002 and presented in false color combination (Thematic Mapper bands 7, 4, and 3 as RGB). The bright green cells were assumed to be irrigated in the Landsat images; however, they cannot be confirmed just by looking at one snapshot of the cropping system; nevertheless, they were helpful in providing qualitative observations of crop growth conditions.

In the Pacific Northwest, the Cascade Mountain range effectively causes most of the moisture from the atmosphere to precipitate onto the west side of the mountains, leaving the east side in a rain shadow. The eastern area, therefore, requires irrigation for profitable production of most crops [62]. Figure 7(a) shows the irrigation in the Snake River Basin in southern Idaho. MIrAD-US was successful in identifying the spatial distribution pattern of irrigation in this region.
Figure 7. MIRAD-US (first and third column) is qualitatively compared with a snapshot of the cropping system depicted by August 2002 Landsat images (second and fourth) in (a) Snake River Basin in southern Idaho, (b) Central Valley in California, (c) south-central Nebraska, (d) semiarid northwestern Texas, (e) humid Mississippi flood plains in Missouri, and (f) South Florida. Landsat bands 7, 4, and 3 are shown in red, green, and blue combination.
Figure 7(b) depicts the irrigation in the central valley in California. The central valley has the highest concentration of irrigation within the state. MIrAD-US seems to overmap irrigated areas by identifying most of the agricultural lands as irrigated, but the observation was based on just one Landsat image. Most of the agricultural lands in the central valley require an additional supply of water to grow crops because of very low summer precipitation and the Mediterranean climate.

Figure 7(c) depicts the highly irrigated parts in south-central Nebraska, mostly fed by the Ogallala Aquifer. The MIrAD-US mapped the irrigated areas significantly well because of the distinct difference in NDVI between irrigated and non-irrigated crops during the peak growing season in the relatively dry summer month of August. A recent study conducted by the University of Nebraska-Lincoln also confirmed a high degree of agreement (90.2%) between MIrAD-US and the Landsat-based Nebraska land use map [57].

In the semiarid northwestern Texas (Figure 7(d)), irrigation typically occurs on a large scale based on center pivot systems depending on withdrawals form the Ogallala Aquifer. Over 90% of Ogallala withdrawals in the High Plains are used in irrigation [58]. The irrigation presence at this site was clearly mapped by MIrAD-US with identification of the irrigated center pivots. The sharp contrast between large-scale irrigated and non-irrigated vegetation in this semiarid climate was key to the successful identification of irrigated lands in the Texas High Plains.

Unlike the west and the Great Plains, in the humid Mississippi flood plains in Missouri, irrigation occurs on small field sizes and on a supplemental basis. Despite an average annual precipitation of 1,000 mm, periodic summertime drought makes irrigation necessary to avoid crop failure and yield reductions. The comparison in Figure 7(e) shows that MIrAD-US was reasonably successful in mapping irrigated land in this highly heterogeneous region.

In Figure 7(f), irrigated areas were distinctive in this June 2002 Landsat image in south Florida. Despite high annual precipitation, the non-uniform distribution and porous soils with low water holding capacity makes irrigation necessary in Florida. Here, irrigation also reduces water stress for high-value specialty crops and enables environmental modifications, including freeze protection and crop cooling [59]. The irrigated lands mapped by MIrAD-US appear to align better with the Landsat-based irrigation in the northeastern part of the Landsat image but do not appear to be as accurate in the northwestern part of the image possibly because of the tendency for NDVI to saturate over high-density vegetation that may minimize the distinction between irrigated and non-irrigated crops.

5. Discussion and Conclusion

We mapped irrigated lands in the conterminous United States by spatially identifying the areas from the tabular data that were reported by USDA at the county level. The method was robust and easy to implement using datasets from secondary sources that were all publicly available. The resulting map of irrigated areas for 2002 was reasonably accurate in the western and central parts of the United States, with overall accuracy of 92% and 75%, respectively. The very distinct crop growth signals between irrigated and non-irrigated fields were the key in successful mapping of irrigated areas in these regions. MIrAD-US identified irrigation status for each 250-m cell without the fractional presentation of irrigation information. This is especially helpful for the quantification of parameters for water and energy budget with respect to the true spatial location.
Our method of identifying irrigated lands was tied to the USDA statistics for irrigation surface area and the spatial distribution of remote-sensing-based annual peak NDVI for spatial pattern. This makes it possible to spatially map the irrigated areas across the United States consistently for USDA statistics compilation years (i.e., years ending with 2 and 7). Routine implementation of the approach would facilitate the analysis and understanding of spatial trends in irrigated lands across the United States, which is an essential component in the assessment of national water use. The USGS Science strategy calls for a Water Census of the United States, including performance of regular and timely assessments of national water availability and use. As a contribution, the USGS Geographic Analysis and Monitoring program proposes to develop and implement a national irrigation water use monitoring system. This system will provide estimates of water use on irrigated lands across the country requiring detailed information on the location of irrigated fields, their surface area, seasonal cycles, crop types, and rates of evapotranspiration. Combining remote sensing data, including the MIrAD-US, with in-situ data will lead to a comprehensive assessment of current water use for agriculture.

One of our modeling assumptions was that there would be a distinct difference of NDVI between irrigated and non-irrigated crops. This assumption was found to be appropriate in the western and semi-arid regions of the United States, but for the humid regions, high precipitation and the supplemental nature of irrigation caused the difference in NDVI between irrigated and non-irrigated crops to be less distinct, which makes it difficult to map irrigated lands in the humid regions. So, in this approach, we categorized most of the irrigated lands well as they are concentrated in the West, but we may not have categorized it well in the humid eastern region. However, visual interpretation shows comparative improvements in the mapping of irrigated areas in MIrAD-US across the humid eastern regions of the U.S. over existing geospatial maps of irrigated areas by [3,11,13-16].

The accuracy of the MIrAD-US model results is strongly dependent on the accuracy of county irrigation statistics and the NLCD data layer. Thus, uncertainties in these source data products from secondary sources may compromise the performance of the model results.

The 2002 MIrAD-US model and geospatial results were encouraging and we plan to investigate ways to improve the methodology. Immediate plans include using this method to calculate a new version of the MIrAD-US for 2007. The 2007 county irrigation statistics are now available from the USDA. Other developments in the inputs that will be incorporated in the model will be a 2007 annual peak NDVI derived from eMODIS and the revised 2006 land cover from the National Land Cover Database which is expected to be available in 2010 [60]. A new MODIS-based North American land cover dataset at 250-m resolution will also be tested. Other plans include testing the model with alternative vegetation index inputs. For example, the Enhanced Vegetation Index (EVI) may provide a better indicator of peak vegetation than NDVI [61]. Also optimizing the model for the hierarchical selection of vegetation index by incorporating crop types may yield improved separation of irrigated and non-irrigated lands especially in the humid regions. Incorporation of topographic derivatives may provide improvements to the model.

The 2002 MIrAD-US is available for download from the USGS EROS Early Warning and Environmental Monitoring Program Web site at http://earlywarning.usgs.gov/USirrigation/. The dataset is provided with Environmental System Research Institute GRID and ENVI image file formats at two different cell sizes (250-m and 1-km) and includes appropriate metadata.
Acknowledgements

This work is performed under USGS contract 08HQCN0005. We would like to acknowledge the Kansas Applied Remote Sensing Program for downloading, mosaicking, and projecting the MODIS vegetation index data used in this project. We sincerely thank Santhosh K. Seelan and Bethany A. Kurz of the University of North Dakota for providing us the ground reference irrigation data for the Great Plains. We are also grateful to the anonymous reviewers for helpful comments. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

References


57. Dappen, P.; Merchant, J.; Ratcliffe, I.; Robbins, C. *Delineation of 2005 Land Use Patterns for the State of Nebraska*; Department of Natural Resources, Final Report; Nebraska Department of Natural Resources: Lincoln, NE, USA, 2007.


© 2010 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/3.0/).