INTEGRATING SATELLITE OBSERVATIONS INTO PROCESS-BASED MODELS TO INFORM AGRICULTURAL WATER MANAGEMENT

By

Jillian M. Deines

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Environmental Geosciences—Doctor of Philosophy

2018

ABSTRACT

INTEGRATING SATELLITE OBSERVATIONS INTO PROCESS-BASED MODELS TO INFORM AGRICULTURAL WATER MANAGEMENT

By

Jillian M. Deines

Irrigation plays an important role in food production and the water cycle worldwide by enhancing agricultural yields, buffering climate variability, and appropriating 70% of total human freshwater use. Maintaining and even expanding irrigated areas is required to address increasing global food demand and climate-induced water stress. Over the latter half of the twentieth century, however, non-renewable groundwater use more than tripled to comprise ~1/5 of global irrigation water. As a result, key agricultural regions around the world are on unsustainable trajectories due to aquifer depletion. With limited water resources defining the 21st century, finding ways to maximize water use and operate within system boundaries is crucial. Crop and hydrology models can support decision making in the face of these challenges by simulating alternative management pathways under a range of resource conditions. In many cases, however, critical input datasets are missing or lack the precision and accuracy to fully parameterize landscape models. Recent rapid advances in large-scale satellite remote sensing can address these data gaps by quantifying landscape characteristics at previously infeasible spatial and temporal resolutions.

In this dissertation, I present new methodologies that translate Landsat satellite observations into annual irrigation maps needed to understand and manage agricultural water resources. Maps are then analyzed and integrated into crop models to better understand historic water use, evaluate novel stakeholder-driven groundwater management, and support future planning. I focused on the High Plains Aquifer (HPA) in the central United States, where a $20

billion agricultural economy is threatened due to extensive depletion over much of the aquifer.

In Chapter 1, I used Google Earth Engine and the full Landsat archive from 1999-2016 to generate annual, moderately high resolution (30 m) irrigation maps for the Republican River Basin portion of the HPA from 1999-2016. I found considerable interannual variability in irrigation location and extent, largely driven by annual precipitation, commodity prices, and increased irrigation efficiency over time. Chapter 2 extended this method to the full 450,000 km^2 HPA from 1984-2017, addressing additional challenges from satellite data gaps and a wider range of climate, crop types, and management. I estimated that up to 24% of currently irrigated area could be lost by 2100 if aquifer depletion continues along recent trends.

With increasing resource scarcity, a diverse set of groundwater management approaches have emerged across the HPA to slow depletion. In Chapter 3, I combined the satellite-derived irrigation maps, detailed well records, and national crop maps to assess the efficacy of innovative stakeholder-driven groundwater management in northwest Kansas referred to as the Local Enhanced Management Area (LEMA) program. I found that farmers surpassed targets for reduced water use without compromising irrigated area through adaptive cropping choices and increased irrigation efficiency. Chapter 4 extends the LEMA analysis with process-based crop models to robustly quantify impacts to the full water budget along with trade-offs in crop yield. Integrating remote sensing into this modeling framework allowed me to estimate quantities that are difficult or impossible to measure. As aquifer depletion threatens crop production in many parts of the world, approaches that integrate models with in-situ and remotely sensed data can improve understanding and help inform economically and hydrologically sustainable management strategies.

This dissertation is dedicated to my daughter, Riley Deines.
May we all keep learning.

ACKNOWLEDGEMENTS

It is a privilege to work with engaged colleagues on questions that both fascinate us and address challenges facing human society. I have been extremely fortunate to share my graduate experience with mentors and peers whose keen minds, generosity, and good humor have elevated my research and fanned my enthusiasm through successes and setbacks alike. First and foremost, I would like to thank my advisors. David Hyndman has guided and challenged me throughout this dissertation, lending his wide perspective and deep knowledge to improve the quality and relevance of each chapter. I am particularly grateful for his trust and support which enabled me to spend my final years remote, producing a daughter in addition to scientific works – many may talk about supporting trainees in finding balance among life's demands, but Dave nonchalantly gives his students the freedom to find solutions that work for them. His one possible failing is my continued satisfaction with boxed wine – sometimes intentional ignorance is too difficult to overcome. Jianguo "Jack" Liu laid the foundations upon which my dissertation evolved. Jack's relentless efforts to understand the whole picture in all of its complexities have helped me frame my understanding of agriculture as a coupled human-natural system. I aspire to his example of thoughtful, timely, and holistic research. Thank you, Jack, for your support in allowing me to pursue my scientific curiosity.

I would like to thank all of my committee members for their guidance. Bruno Basso provided helpful council in the development and analysis of my work, furthering my understanding of computational agronomy and how changing agricultural practices can impact the landscape. Jinhua Zhao helped frame the socio-economic context for my work, widening my perspective and providing insight into farmer decision making. Anthony Kendall deserves a page unto himself, but I'll try to restrain myself to this paragraph. My research would have stalled many times if not for

Anthony's fount of keen insights and ready availability for problem solving. More than this, Anthony has become a trusted sounding board for larger discussions on science, the academy, and life. I am a better scientist, programmer, and person due to his mentoring and friendship.

I had the good fortune to work closely with additional researchers outside of my committee and lab group, and my research is stronger for it. I'm particularly grateful to Jim Butler at the Kansas Geological Survey for his insight and keen analytical eye. Brian Baer and Lydia Rill in the Basso Lab provided crucial support in my crop modeling efforts, and I am endlessly appreciative of their timely responses and contributions to my modeling workflow.

The Hydro Lab provided the perfect incubator for my life as a researcher, largely due to the excellent collection of dedicated folks who enjoy research as much as they enjoy light pranks and walks for coffee. Thanks to Sherry Martin, Alex Kuhl, Erin Haacker, Jeremy Rapp, Blaze Budd, Kayla Cotterman, Travis Dahl, Chanse Ford, Quercus Hamlin, Bailey Hannah, Brent Heerspink, Xiao Liu, Troy Ludwig, Autumn Parish, Lisi Pei, Jake Roush, Sam Smidt, Ryan Vannier, Luwen Wan, and Tianfang Xu. Special thanks go to the EES department staff, including Pam Robinson, Dallas Coryell, and Brittany Walter.

Finally, I'd like to thank my parents, Dave and Mary Mueller, for instilling in me a fortuitous mix of perseverance and mirth well-suited to seeking answers in the unknown. I am endlessly grateful for their sacrifices, support, and friendship. I've also been fortunate to pick up some great family along the way, and I thank Robb and Laurie Deines in particular for their warm interest and support. Thanks also to the next generation of Deineses and Muellers – Riley, Macy, Claire, and Marshall – for providing joy and a pretty good reason for getting some of these grand challenges sorted out. And last and greatest, thanks to Andy, my partner on the path of life. You're awesome.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1:
ANNUAL IRRIGATION DYNAMICS IN THE US NORTHERN HIGH PLAINS DERIVED FROM LANDSAT SATELLITE DATA

Abstract

Sustainable management of agricultural water resources requires improved understanding of irrigation patterns in space and time. We produced annual, high resolution (30 m) irrigation maps for 1999-2016 by combining all available Landsat satellite imagery with climate and soil covariables in Google Earth Engine. Random forest classification had accuracies from 92-100% and generally agreed with county statistics (r^2 = 0.88-0.96). Two novel indices which integrate plant greenness and moisture information show promise for improving satellite classification of irrigation. We found considerable interannual variability in irrigation location and extent, including a near doubling between 2002 and 2016. Statistical modeling suggested precipitation and commodity price influenced irrigated extent through time. High prices incentivized expansion to increase crop yield and profit, but dry years required greater irrigation intensity, thus reducing area in this supply-limited region. Datasets produced with this approach can improve water sustainability by providing consistent, spatially explicit tracking of irrigation dynamics over time.

1. Introduction

Following rapid expansion in the late 20th century, global irrigated area is now relatively stable [*Wada et al.*, 2013]. Regional gains and losses, however, can be substantial [*Brown and Pervez*, 2014]. Dynamic crop prices, climate and precipitation variability, changing water policies, and crop rotations all drive considerable local interannual variability in irrigated area [*Ozdogan and Gutman*, 2008; *Wisser et al.*, 2008; *Brown and Pervez*, 2014]. Spatial irrigation datasets that accurately delineate irrigated areas annually would help constrain water budgets, improve hydrologic models, provide timely information to water managers and food security efforts, give insight into factors that influence irrigation behavior, and further clarify the effects of climate change on irrigation water demand and supply. Researchers have noted the need for routine mapping of irrigated lands [*Thenkabail and Wu*, 2012; *Brown and Pervez*, 2014; *Peña-Arancibia et al.*, 2014; *Teluguntla et al.*, 2017], yet satellite-derived annual datasets are rare [*Abuzar et al.*, 2015] due to historic computational limitations and inadequate ground reference data.

Quantifying temporal and spatial variations in irrigation is fundamental to the challenge of sustainable water management. Globally, irrigated agriculture accounts for approximately 70% of human freshwater use [*Rosegrant et al.*, 2009; *Wada et al.*, 2013]. Irrigation greatly enhances agricultural yields [e.g., *Smidt et al.*, 2016] and price stability, but overexploitation of water resources has depleted groundwater aquifers and reduced annual river discharge [*Postel*, 2003; *Rockström et al.*, 2012]. Moreover, incentives to expand irrigation continue to grow due to increased food demand [*Tilman et al.*, 2011], agricultural intensification [*Gleick*, 2003], and climate change [*Wada et al.*, 2013; *Aleksandrova et al.*, 2014]. Effectively managing limited water resources to meet future irrigation needs while remaining within regional and planetary

boundaries of sustainable freshwater use [*Rockström et al.*, 2012] is a major challenge.

Unfortunately, existing irrigation datasets are largely inadequate for this task, and the locations of irrigated areas remain uncertain [*Ozdogan and Gutman*, 2008; *Wisser et al.*, 2008; *Wada et al.*, 2011; *Peña-Arancibia et al.*, 2014]. Existing datasets are primarily based on administrative boundary statistics for irrigated area or land equipped for irrigation, which lack spatial precision and can contain self-reporting bias. Existing spatially explicit, satellite-derived datasets tend to have relatively low resolution (250 – 1000 m), particularly at regional scales. Critically, the vast majority of datasets are generally single year, static snapshots that overlook temporal irrigation dynamics.

Notable exceptions include recent work mapping annual irrigation for 14 years in Afghanistan [*Pervez et al.*, 2014] and 16 years in Australia [*Teluguntla et al.*, 2017], which provided insights into temporal trends and variability in irrigation. For example, Pervez et al. [2014] found irrigated area differed as much as 30% among years. Both studies were limited to the relatively coarse 250 m resolution of Moderate Resolution Imaging Spectroradiometer (MODIS) satellite products due to reported computing constraints. Although moderate resolution efforts are sufficient to capture broad scale patterns [*Wardlow and Egbert*, 2008], higher resolution imagery such as those from Landsat satellites (30 m) better resolve smaller or fragmented fields, provide precise field locations, and increase accuracy [*Velpuri et al.*, 2009]. Due to the corresponding increase in data volume and processing requirements, however, Landsat based annual datasets are rare and limited to local studies. For example, Ozdogan et al. [2006] produced nine annual 30 m maps for a 1500 km^2 area in Turkey using one Landsat scene per year. Irrigation dynamics compared across these early efforts in annual mapping differ in overall trend, yearly variance, and contextual drivers, suggesting that annual, spatial datasets

offer a refined picture of regional irrigation differences not well captured by static maps or aspatial data.

Here, we produced high resolution, annual irrigation maps from 1999 to 2016 across the greater Republican River Basin region in the central United States (Figure 1.1), hereafter termed the Annual Irrigation Maps – Republican River Basin (AIM-RRB) dataset (available at http://dx.doi.org/10.4211/hs.55331a41d5f34c97baf90beb910af070). We leveraged recent developments in cloud computing to utilize all available Landsat scenes each year, combining satellite imagery with climate and soil covariables in a random forest classification workflow that is readily applicable to future years for ongoing monitoring. Research using the full Landsat record is a relatively recent phenomenon [e.g., *Hansen et al.*, 2013] and to our knowledge not previously applied to irrigation mapping. We then used these maps to examine irrigation dynamics and associated drivers across this region.

2. Methods

2.1. Study area

The Republican River Basin (RRB) overlies portions of Colorado, Nebraska, and Kansas, draining a large portion of the High Plains Aquifer (HPA) before leaving the aquifer near the downstream Nebraska-Kansas border (Figure A.1.1). The basin provides riparian surface-water irrigation and groundwater irrigation over the HPA. Annual cropping systems dominate the region, and the top five crops by area planted (wheat, corn, soy, alfalfa/hay, and sorghum) can be both irrigated or rainfed (Figure A.1.2). Due to litigation concerning interstate water use beginning in 1999, both groundwater and surface-water irrigation are regulated to preserve streamflow into Kansas in accordance with the Republican River Compact of 1942. Strategies to meet

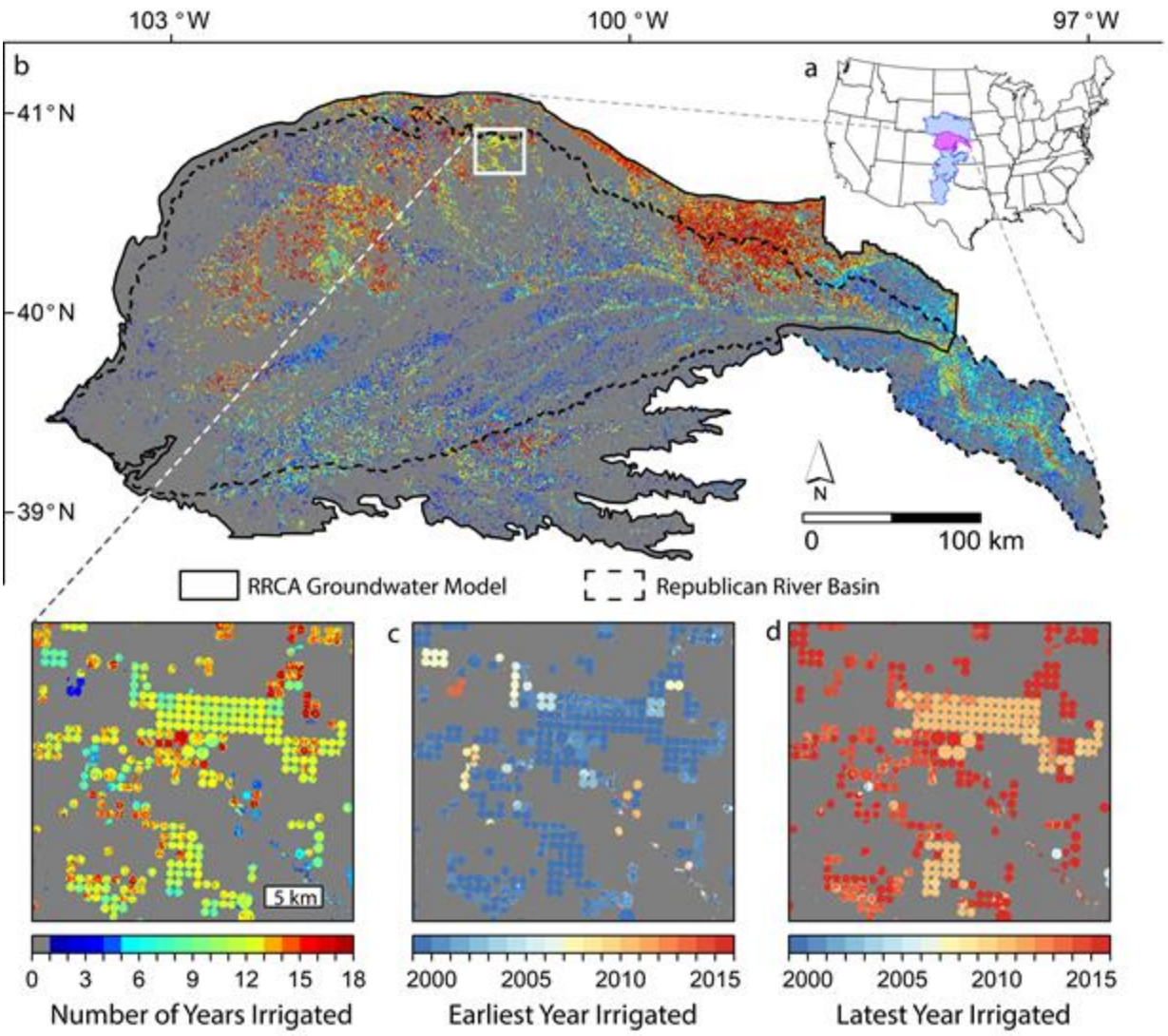

Figure 1.1. Study area location and map of irrigation frequency. (a) Study area (purple) in the context of the High Plains Aquifer (blue); (b) Number of years each 30 x 30 m map pixel was classified as irrigated between 1999-2016 across the Republican River Basin (dashed outline) and the associated Republican River Compact Administration's (RRCA) groundwater model (solid outline), with zoomed inset for enhanced resolution. Annual irrigation maps also demarcate novel and deactivated irrigated areas as demonstrated by mapping earliest (c) and latest (d) years irrigated during the study period.

streamflow targets vary widely across localized management districts, change over time, and include restrictions on pumping volume, well-drilling moratoriums, efforts to retire water rights, and expensive augmentation plans via engineered water transfers [see Griggs, 2017 for further discussion]. The Republican River Compact Administration assesses compliance with a groundwater model covering the groundwatershed upstream of Kansas [*RRCA*, 2003], an area

hereafter termed the RRCA. Therefore, our study domain is the 86,429 km^2 greater Republican Basin (GRB), defined as the union of the RRCA and RRB (Figure 1.1, Figure A.1.1). Annual irrigation maps and accuracy metrics are produced with a minimum 10 km buffer (total area: 141,603 km^2, Text A.1.1), though map results are presented solely for the GRB. In addition, we analyzed irrigation drivers in the portion of the RRB contained within the RRCA (RRB-RRCA) for 1999-2015 to capitalize on irrigation water volume data from the groundwater model (see section 3.4, *Drivers of irrigated area*). We defined the crop year as 1 November to 31 October to ascribe greenness from winter wheat to the year harvested. Mean annual precipitation increases eastward along a longitudinal gradient, ranging from 341 – 845 mm during the study period. Growing season precipitation (1 December – 31 August) ranged from 284 - 673 mm [*Abatzoglou*, 2013].

2.2. Satellite imagery, vegetation indices, and environmental variables

Landsat imagery is provided at nominal 30 m resolution in 182 x 185 km scene tiles, sixteen of which overlie the buffered study region (Figure A.1.3). Working in Google Earth Engine's (GEE) cloud computing platform [*Gorelick et al.*, 2017], we used all available Landsat Surface Reflectance Products [*USGS*, 2017b, 2017c] from 1 November 1998 to 31 October, 2016 (9592 scenes, Text A.1.2), as temporal resolution increased substantially in 1999 after Landsat 7 came online and image acquisition improved [*Pekel et al.*, 2016]. Concurrently operating Landsat satellites provided an 8-day overpass interval for all years except 2012, when only Landsat 7 was operational. This 8-day interval was simultaneously augmented by side-overlapping scene edges and reduced by clouds and acquisition inconsistencies, resulting in 99% of pixels having between 12 - 64 satellite observations per year except 2012 (mean including 2012: 28). This provided adequate temporal resolution to capture both baseline and peak

greenness for multiple crop calendars (Figure 1.2a). Detailed information about Landsat scenes, processing, and yearly statistics for pixel observation frequency can be found in Text A.1.2 and Figure A.1.4.

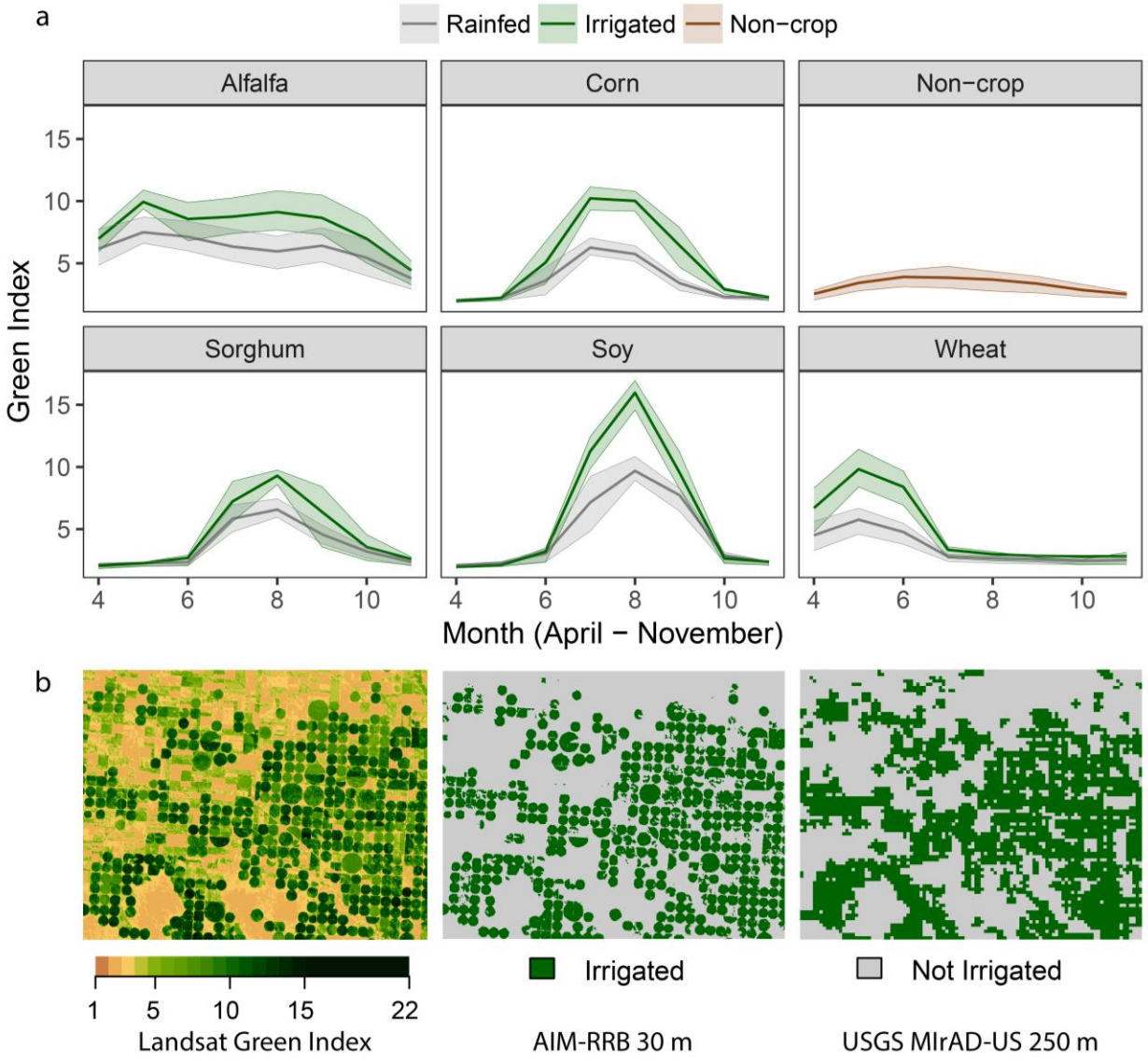

Figure 1.2. Seasonal greenness curves and qualitative assessment. (a) Mean and interquartile range (shaded) from Landsat monthly green index (GI) composites for 2010 training points. (b) Qualitative assessment of classification performance in 2002. Left: Landsat annual composite image of maximum greenness; Center: AIM-RRB classification. Right: Compared with 250 m resolution products [*Pervez and Brown*, 2010], AIM-RRB has similar patterns but higher spatial and temporal resolution (annual vs. five year).

In GEE, we produced composites of annual maximum and annual range for four vegetation indices: (1) the normalized difference vegetation index (NDVI); (2) the enhanced

vegetation index (EVI); (3) the normalized difference water index (NDWI), which is sensitive to plant water content [*Gao*, 1996]; and (4) a less common green index (GI) [*Gitelson et al.*, 2005] that is particularly sensitive to irrigation status [*Ozdogan and Gutman*, 2008]. Text A.1.3 provides detailed index calculations. Composites thus captured both peak growing season greenness and the magnitude of annual change per pixel regardless of crop phenology.

Climate, soil, and slope information can improve classification accuracy by refining cases of potential irrigation and providing context for vegetation greenness. We assembled variables related to plant growth including precipitation, plant available water, slope, and aridity (Text A.1.4). We also developed two novel combination indices that integrate moisture information with greenness levels to exaggerate differences by irrigation status and facilitate regional-scale classification across climate gradients. We called these the water-adjusted green index (WGI), calculated from Landsat as NDWI * GI, and aridity-normalized green index (AGI), calculated as GI / growing season aridity derived from meteorological data. In total, we generated 9 Landsat variables and 11 covariables for use in machine learning classification (Table A.1.1).

2.3. Training data

We developed a robust training dataset using high-resolution (1 m) aerial imagery [*NAIP*, 2017], Landsat GI and EVI times series (Text A.1.5), and crop type maps (CDL) [*USDA-NASS*, 2017]. To maximize sampling of climate conditions, we created a multi-year training dataset using a wet year (2010) and a dry year (2012) and sampled across three Koeppengeiger climate zones [*Peel et al.*, 2007]. In GEE, we manually located points for the top five crops plus non-crop grassland, determining irrigation status from multiple lines of evidence such as irrigation infrastructure and seasonal greenness patterns (Text A.1.5). We defined irrigation as the use of supplemental water during the growing season and did not differentiate between partial or fully

irrigated fields or cropping intensity. Expert-selected training points, used here, can reduce data needs and improve performance compared to stratified random approaches [*Zheng et al.,* 2015]. The final training dataset consisted of 1401 points (see Figure A.1.5 for locations and Table A.1.2 for breakdown by crop and climate region). Crop-specific seasonal greenness curves show good separation by irrigation status (Figure 1.2a). Points were then aggregated into "irrigated" and "non-irrigated" types for binary classification.

2.4. Classification

We used the full training dataset to train both Classification and Regression Tree (CART) and random forest [*Breiman,* 2001] classifiers in GEE. A random forest classifier with 500 trees that omitted rainfed soy training points performed best on validation data used to evaluate classifiers (see Text A.1.6). We applied the classifier to the 1999-2016 period after masking urban, forest, and wetland areas using National Land Cover Dataset (NLCD) maps [Fry et al., 2011]. We did not mask other non-crop areas because this inhibited classification of dynamic irrigation changes among years.

Following initial classification, we performed two cleaning operations: (1) a 3x3 majority filter and (2) removal of pixels irrigated only once during the 18 year period, since infrastructure requirements make single-year irrigation unlikely. To understand the relative contribution of input variables to classification accuracy, we ran permutation tests and GINI index metrics in R [*R Core Team,* 2014] with an identically parameterized classifier since GEE does not currently output variable importance measures (Text A.1.7).

2.5. Accuracy assessment and analyses

Assessing multi-year classification efforts across large regions is challenging since limited ground truth data are available. We sought to evaluate accuracy with test datasets across

a wide range of years from multiple data sources. First, we used two sets of national county statistics for six years (2002, 2007, 2012: NASS Agricultural Census [*NASS*, 2017]; 2000, 2005, 2010: USGS water use data [*USGS*, 2015]) to compare total irrigated area for 35 counties contained within the buffered GRB. Second, we randomly generated points across Nebraska in 2002 and the full study area in 2015 (Figure A.1.6) and marked them as "irrigated" or "not irrigated" as described for training points. Cases where no clear determination could be made (24 of 2266 points) were marked as "uncertain" and omitted from accuracy assessments. We chose 2002 (dry year) and 2015 (wet year) to include all three Landsat sensors and two precipitation extremes in our assessment. Table A.1.3 gives point breakdowns among years and classes. We then analyzed annual maps to provide summary statistics of irrigated area, overall and regional trends, and exploratory analyses of irrigation drivers.

2.6. Data limitations

Although we leveraged several GIS, satellite, and aerial imagery datasets, our method relied on manually-produced training and test datasets well suited to identify areas where irrigation clearly enhances greenness. Locations where irrigation may have more subtle effects on greenness, such as sub-irrigation or where limited irrigation is used to prevent crop failure, were not selected. AIM-RRB could therefore be described as a map of "certainly irrigated" locations but may underrepresent some marginal irrigation areas.

3. Results and Discussion

The random forest classifier using all available Landsat scenes produced 18 annual irrigation maps from 1999-2016 (AIM-RRB). Figure 1.1b shows the number of years each pixel in the study region was classified as irrigated during these 18 years. We found that 24.3% of the

GRB was irrigated at some point during the study period, and of that area, 28.1% was perennially irrigated, which we defined as having irrigation over 80% of years to allow for periodic crop rotations and fallowing. Only 8.4% of irrigated land was classified as irrigated for the entire study period. In general, perennial irrigation was concentrated near major rivers (Platte and Republican, Figure A.1.1) and in well-established groundwater areas. Non-perennial irrigation includes irrigated fields added, deactivated, or intermittently rainfed/fallowed during the study period (see section 3.3, *Irrigation trends*). Figure 1.1c and Figure 1.1d demonstrate how AIM-RRB can resolve years in which irrigation of individual fields began and/or ceased. As this zoomed area highlights, irrigated areas were both added and deactivated throughout the study period.

3.1. Classification performance

Qualitatively, there was good visual agreement between Landsat composites, AIM-RRB, and previously published USGS MIrAD-US products at lower resolution (250 m) for 2002, 2007, and 2012 [*Pervez and Brown*, 2010; *Brown and Pervez*, 2014] (Figure 1.2b). Using our point test dataset, we found overall accuracies of 98.6% and 97.6% for 2002 and 2015, respectively. For the irrigated class, we had omission errors from 6.1-7.6%, and commission errors from 0-6.3%. Table A.1.3 shows a full breakdown of accuracy by class type for 2002 and 2015.

County-level comparisons with NASS and USGS irrigation statistics showed good agreement with AIM-RRB estimates (Figure A.1.7). We found r^2 values between 0.88 – 0.96 for the six available years, with similar agreement between years used to train the classifier (2010 and 2012) and non-training years as well as robust performance across high and low precipitation years. AIM-RRB slightly underestimated irrigated area per county compared to the county

statistics, which can be seen in relation to the 1:1 lines in Figure A.1.7. USGS data is derived from state-specific statistical models with associated uncertainties, so it remains unclear if AIM-RRB underestimates irrigation or if USGS estimates are high. NASS census data is self-reported but anonymized to minimize inaccurate reporting. However, there may be underlying incentives to report inflated numbers to preserve water rights. Alternatively, NASS may better reflect partial irrigation while AIM-RRB likely favors fully irrigated fields (see section 2.6, *Data limitations*). Finally, the Landsat dataset likely missed peak greenness in some locations due to cloud cover, resulting in occasional maximum greenness values similar to non-irrigated cropland. Because the MIrAD-US methodology uses the NASS county area statistics to allocate pixels to the irrigated class, AIM-RRB is the only independent multi-year data source in the region for this period

3.2. Variable importance

Our novel AGI and WGI indices, which combine GI with moisture indicators, ranked highest for both importance metrics used (permutation test: AGI; GINI Index: WGI; Text A.1.7, Figure A.1.8). GI contributed to the top three variables identified through both metrics, supporting previous findings that GI is more sensitive to irrigation status than conventional indices such as NDVI and EVI [*Ozdogan and Gutman*, 2008; *Ozdogan et al.*, 2010a]. The annual GI range scored higher than the maximum for both metrics, suggesting that the change in greenness over the year conveys more information than peak greenness alone, corroborating conclusions in Ozdogan et al. [2010]. Interestingly, no climate-related variables ranked in the top eight according to the GINI Index, despite the high relative importance of AGI in the permutation test. Climate-related variables may gain importance for continental scale applications with larger climatic ranges. Slope and soil-related variables scored lowest,

indicating they do not enhance accuracy in this region. The high importance of AGI and WGI suggests these indices warrant further study for use in irrigation classification in other agricultural regions.

3.3. Irrigation Trends

Irrigated area in the GRB increased during the study period at an average rate of 0.37% per year ($r^2 = 0.72$, $p < 0.0001$, Figure 1.3a), with a lower rate of 0.26% for the more regulated RRB-RRCA sub-domain ($r^2 = 0.62$, $p < 0.0001$; Figure 1.4a). We found considerable variability around this trend, including multiple years in which irrigated area decreased from the previous year. The range in irrigated area among years was large; for example, irrigated area in the GRB increased by 92% between the low in 2002 and the high in 2016. Given this variability, datasets lacking high temporal frequency could generate disparate conclusions based on the years sampled. For example, a five-year product such as MIrAD-US, which is based on NASS data for 2002, 2007, and 2012, would suggest a non-significant 0.02% increase per year ($p = 0.90$).

Linear regression of irrigated area over time by 4 km^2 aggregated grid cells revealed that the highest rates of increase were concentrated in the eastern, non-aquifer region and near the Platte and Republican Rivers, while western groundwater-dominated regions had relatively flat to decreasing trends (Figure 1.3b). This is likely due to groundwater allocation reductions, expanded well-drilling moratoriums, and retirement of water rights in Nebraska and Colorado to comply with the Republican River Compact. These efforts to protect streamflow have perhaps enabled the expansion of irrigation evident in the lower Kansas RRB (Figure 1.3a, Figure 1.3b). Irrigated area was both added and deactivated across the study region (Figure 1.3a). Surface-water dominated regions such as the lower Kansas RRB deactivated negligible irrigated area over the study period but did suffer large temporary reductions in irrigated area during drought

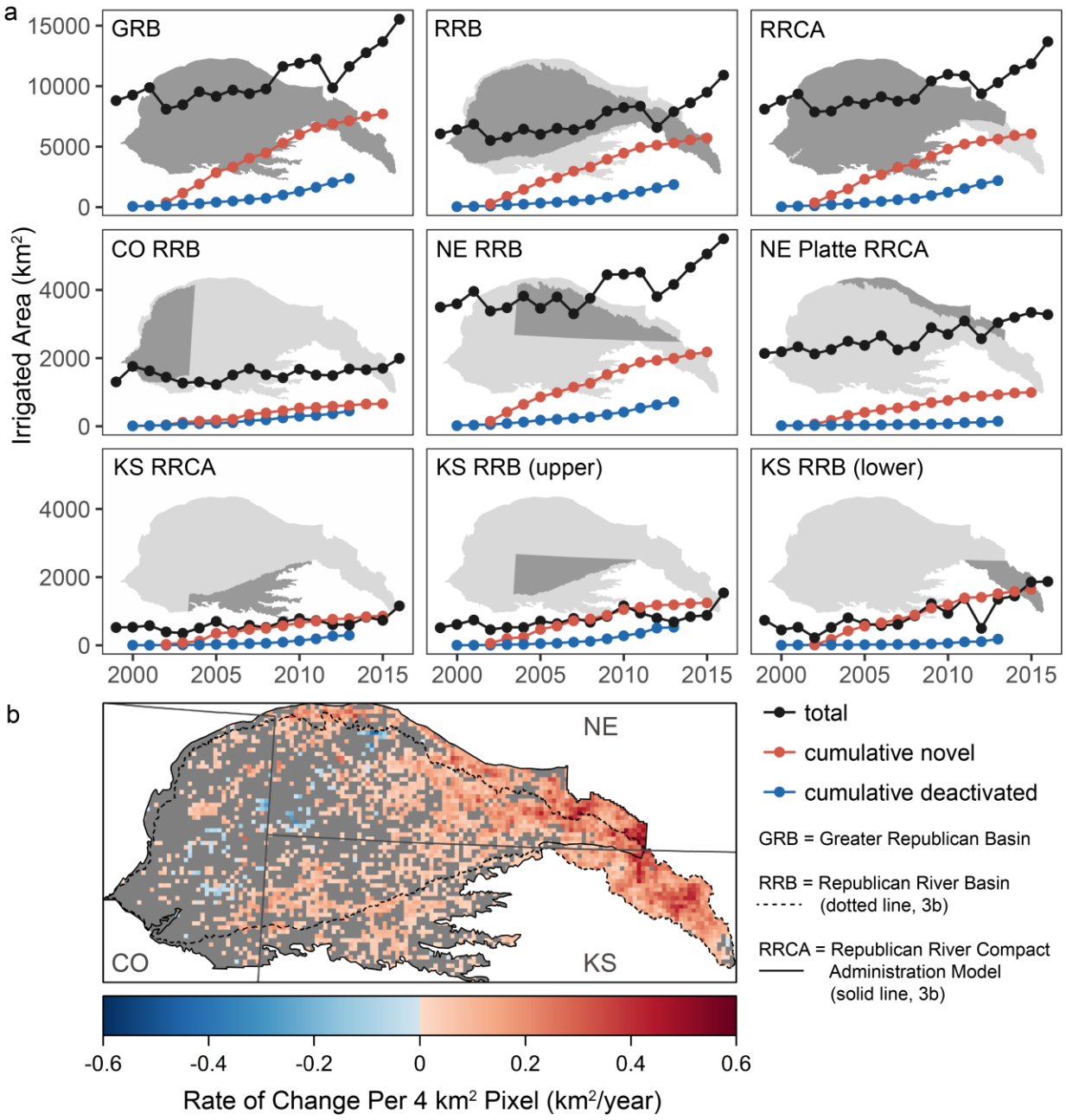

Figure 1.3. Sub-regional irrigation trends. (a) Irrigated area (black) by region. Spatial locations of each region are depicted in map backgrounds (dark gray). Cumulative novel area (red) summarizes newly irrigated pixels each year (2002-2015); cumulative area deactivated (blue) tracks pixels not irrigated in subsequent years (2000-2013). Omitted years buffered against consecutive fallow periods at the study period start or end; (b) Rate of change over time from linear regression. Cells with non-significant trends (alpha $\geq$ 0.05) are in gray.

years such as 2012. Groundwater dominated regions such as the CO RRB and the KS RRCA were less perturbed by drought but had the lowest net gain in novel irrigated areas (novel –

deactivated area). Changes in total irrigated area not accounted for by net gains or losses likely were due to reduced dryland crop rotations and/or fallowing frequency in existing irrigation areas.

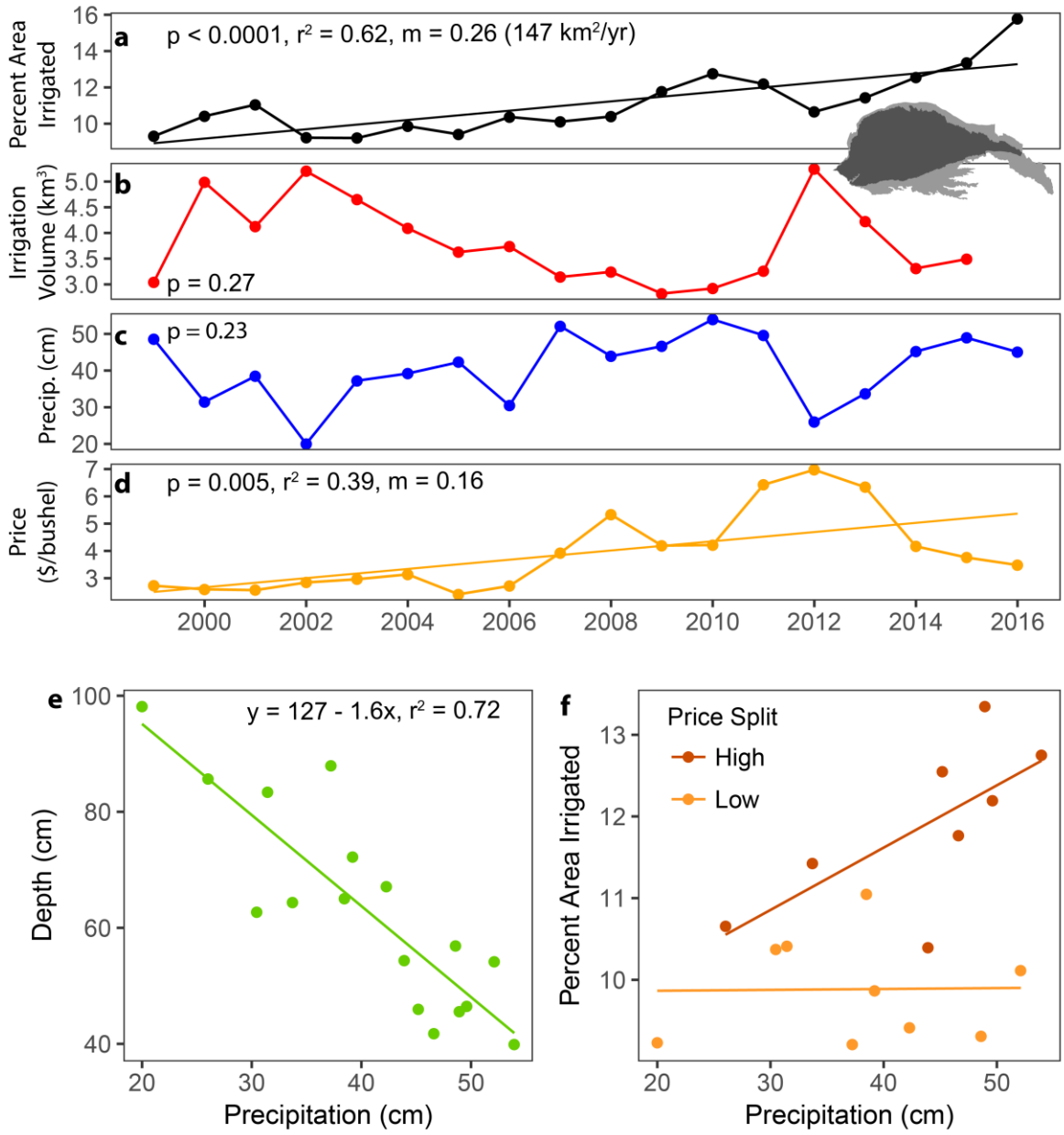

Figure 1.4. Irrigated area over time and associated drivers. For the portion of the Republican River Basin overlying the High Plains Aquifer (RRB-RRCA, shown in dark gray in the map inset of the Greater Republican Basin study area): (a) Percent irrigated area from AIM-RRB. Rate of change (m) is given in percent and actual area; (b) Irrigation water volume [*RRCA*, 2003]; (c) Precipitation from December 1 – August 31 (Text S3) [*Abatzoglou*, 2013]; (d) Corn price in 2016 dollars [*NASS*, 2017]; (e) Irrigation application depth (volume / area) vs. precipitation linear regression; (f) Trends in irrigated area vs. precipitation for years with high and low prices (split determined from CART (Text A.1.5, Figure A.1.9)).

3.4. Drivers of irrigated area

Spatiotemporal irrigation dynamics detailed in AIM-RRB result from farmer irrigation decisions made within the context of annual climate variation, crop commodity prices, water management, and water supply. Utilizing irrigation water volume estimates from the RRCA model for 1999-2015 (Figure 1.4b) [*RRCA*, 2003], we investigated how these drivers might interact to influence irrigation within the RRB-RRCA. Correlation matrices revealed that irrigated area was positively correlated with the previous year's crop prices ($r = 0.55$, $p = 0.02$) but not with precipitation, irrigation volume, or current year price (Figure A.1.9a). Instead, irrigated area likely was linked to precipitation through the depth of irrigation water applied, calculated as irrigation water volume divided by area. Both irrigation water volume and irrigation depth had strong negative correlations with precipitation ($r = -0.89$ and -0.86, respectively; $p < 0.0001$). In years with low precipitation, such as the 2002 and 2012 droughts, irrigation volume and depths were elevated while irrigated area was reduced (Figure 1.4a, Figure 1.4b, & Figure 1.4e), indicating farmers irrigated more intensely over less area to compensate for lack of rainfall. The inability to maintain or even expand irrigated area during dry periods when yield advantages are greatest suggests farmers are limited in either water supply, access rights, or delivery capability [e.g., *Foster et al.*, 2014]. Without complementary datasets on irrigation volume and spatial extent made possible by annual map products, it is not possible to discern if increased water use was due to areal expansion, application depth increases, or both.

Commodity prices also influence irrigation decisions by determining the return on investment for irrigation water use. Corn price approximately doubled between 2003 and 2012 (Figure 1.4d). CART analysis suggested that price and precipitation interacted to influence annual irrigation extent (model $R^2 = 0.78$; Text A.1.8 and Figure A.1.9b). When price was low,

irrigated area was low regardless of precipitation, likely due to poor return on irrigation costs (Figure 1.4f). In contrast, high prices incentivized irrigation expansion but was modulated by annual precipitation; low precipitation years required increased irrigation water depth ($r^2 = 0.72$, Figure 1.4e), limiting the amount of water available for areal expansion.

Although quantification is outside the scope of this paper, policy, management decisions, and groundwater depletion [e.g., *Basso et al.*, 2013; *Cotterman et al.*, 2017] also influence irrigation dynamics. This can include efficiency incentives and/or technological improvements that can increase area per unit volume, new use restrictions reducing irrigated area or, conversely, areal expansions in anticipation of future regulation [*Pervez and Brown*, 2010].

4. Conclusions

Our approach produced annual irrigation maps that provide consistent, spatially explicit tracking of irrigation, revealing temporal dynamics even in this heavily regulated system. Our use of the full Landsat record for each year allowed us to capture peak greenness values for multiple crops despite asynchronous crop maturation schedules and to quantify the annual range in greenness. We developed two new indicators combining remotely sensed plant greenness with moisture information (WGI and AGI) that show promise for improving satellite classification of irrigation. Because our approach utilizes satellite and derived climate datasets made available in near real time through Google Earth Engine, it can be applied to future years immediately following the growing season to provide updated and timely information to managers and scientists. The approach is transferable to other non-humid regions dominated by annual crops given region-specific training data. These annual maps provide critical insight into behavioral responses to irrigation drivers and document annual irrigation dynamics with high precision, thus providing vital information to inform agricultural water use models and management decisions.

Acknowledgements

This chapter was co-authored by Anthony Kendall and David Hyndman. AIM-RRB in GeoTIFF format and training/test point data sets are available on Hydroshare at http://dx.doi.org/10.4211/hs.55331a41d5f34c97baf90be b910af070. Input data are available in the cited references, and code may be obtained from JMD (jillian.deines@gmail.com). We thank Haoyang Li, Tianfang Xu, and Andrew Deines for guidance on analyses; Burke Griggs for guidance on water regulations; Jeremy Rapp for assistance developing training point methodology; and Alex Kuhl and our reviewers/Editor for manuscript comments. Funding was provided by NSF grant WSC 1039180, USDA NIFA grant 2015-68007-23133, and continuing support from the USDA-NIFA/NSF INFEWS program. Jillian Deines was partially supported by NASA Headquarters under the NASA Earth and Space Science Fellowship Program grant 14-EARTH14F-198. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of NSF, USDA NIFA, or NASA.

APPENDIX

Text A.1.1. Expanded study area description

Much of the Republican River Basin (RRB) overlies the High Plains Aquifer (HPA), one of the largest and most stressed aquifers in the world [*Gleeson et al.*, 2012b]. Average water demand exceeds supply over much of the region [*Devineni et al.*, 2015], leading to extensive groundwater depletion [*Haacker et al.*, 2016; *Mcguire*, 2017]. Similar to trends in the larger HPA, groundwater irrigation in the RRB expanded rapidly over the last half of the 20th century. Arguing that such large-scale groundwater development was depleting surface streamflow into the state, Kansas sued Nebraska in 1999 for violating the Republican River Compact of 1942, which allocated each state a portion of the unaltered basin water supply in perpetuity. Although groundwater was not explicitly addressed in the Compact, the US Supreme Court ruled that groundwater use was restricted under the Compact if it depleted transboundary streamflow and established a framework for using groundwater modeling to assess compliance on 5 year running averages under the Republican River Compact Administration [RRCA; *Peck*, 2007; *Kuwayama and Brozović*, 2013]. Griggs [2017] details various compliance strategies among actors in the basin.

To capture irrigation dynamics in the full Republican River Basin system, our study region is the union of the RRB and the RRCA groundwater model boundary (Figure A.1.1). To properly model the groundwater system, the RRCA boundary extends northward beyond the RRB to use the Platte River to set the boundary conditions along the northern model border, and extends beyond the RRB in the southeast in order to include the full aquifer in this region. The RRCA groundwater model uses 1-mile grid cells which results in a jagged, irregular boundary; thus we modified the RRCA boundary by using the actual borders of the aquifer on the east and west as well as the Platte River in the north. The RRB is fully contained within the RRCA with

the exception of the eastern, downstream tail, which exits the aquifer near the Nebraska-Kansas border and continues to flow through Kansas until it joins the Kansas River. Therefore, our study domain is the 86,429 km^2 greater Republican Basin (GRB), defined as the union of the 79,371 km^2 RRCA and the partially overlapping 64,521 km^2 RRB (Figure 1.1, Figure A.1.1).

As standard best practice to ensure the Annual Irrigation Maps – Republican River Basin (AIM-RRB) dataset generously covers the region of interest for downstream efforts such as hydrological modeling, we applied a buffer to the RRCA (10% of RRCA width, or ~26 km) and to the RRB (10 km) prior to performing a GIS union to combine the areas into the GRB (Figure A.1.1). We used the resulting 141,603 km^2 buffered region to 1) train the classifier, 2) produce maps for the full buffered region, and 3) report accuracy statistics. For example, the buffered study area fully contains 35 counties, while the non-buffered union of the Republican River Compact Administration and the Republican River Basin would only contain 21 counties, limiting this accuracy assessment for the six years in which county data are available.

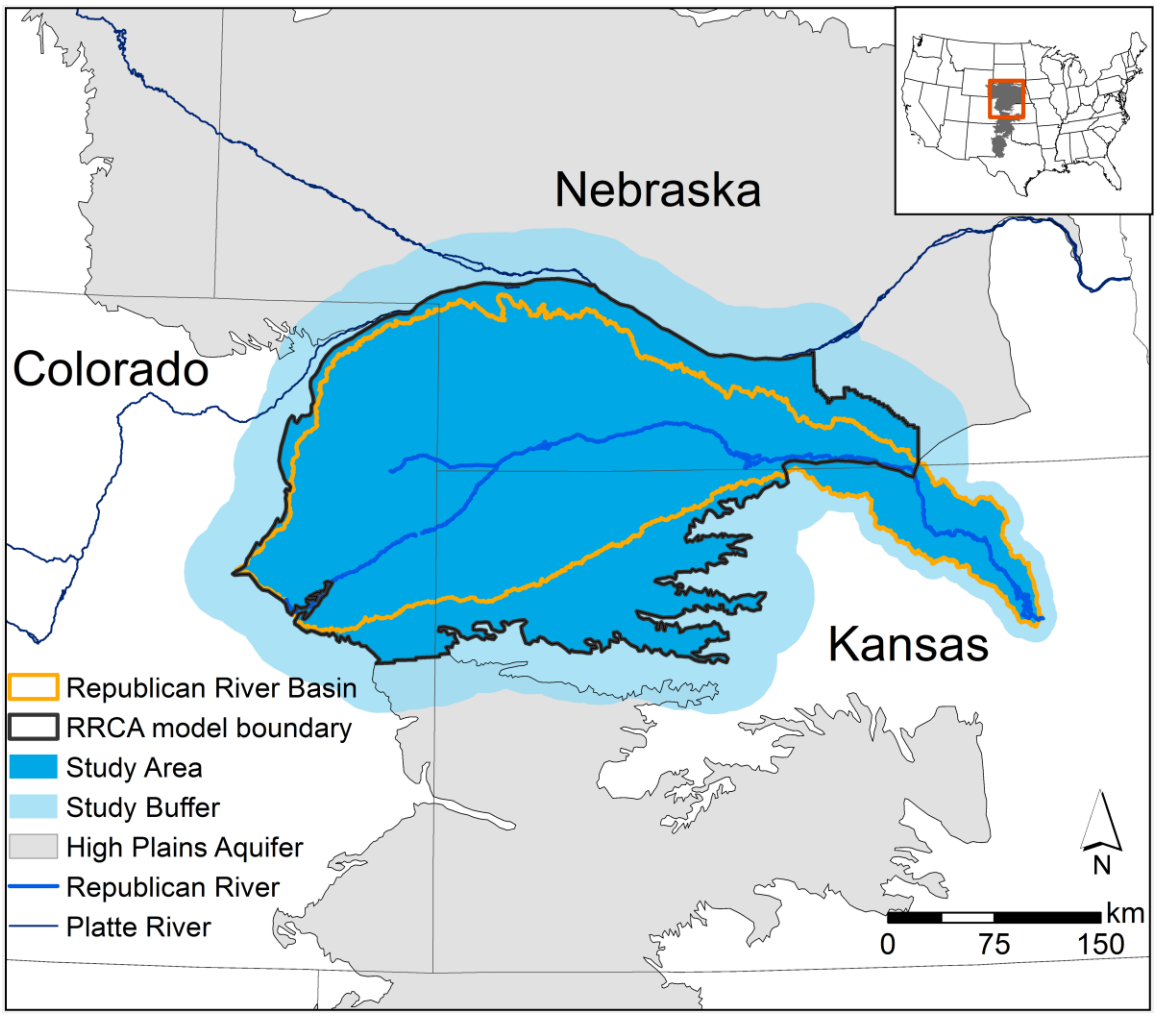

Figure A.1.1. Detailed study area map. The study region (blue) is the union of the Republican River Compact Administration (RRCA) groundwater model domain (black outline) and the Republican River Basin (RRB; orange outline). Annual irrigation maps and accuracy metrics are produced for the full study are plus buffer (light blue), which encompasses the study region with a minimum 10 km buffer. Exploratory analyses of irrigation drivers are conducted on the portion of the RRB contained within the RRCA. Lines demark the Republican River and the Platte River.

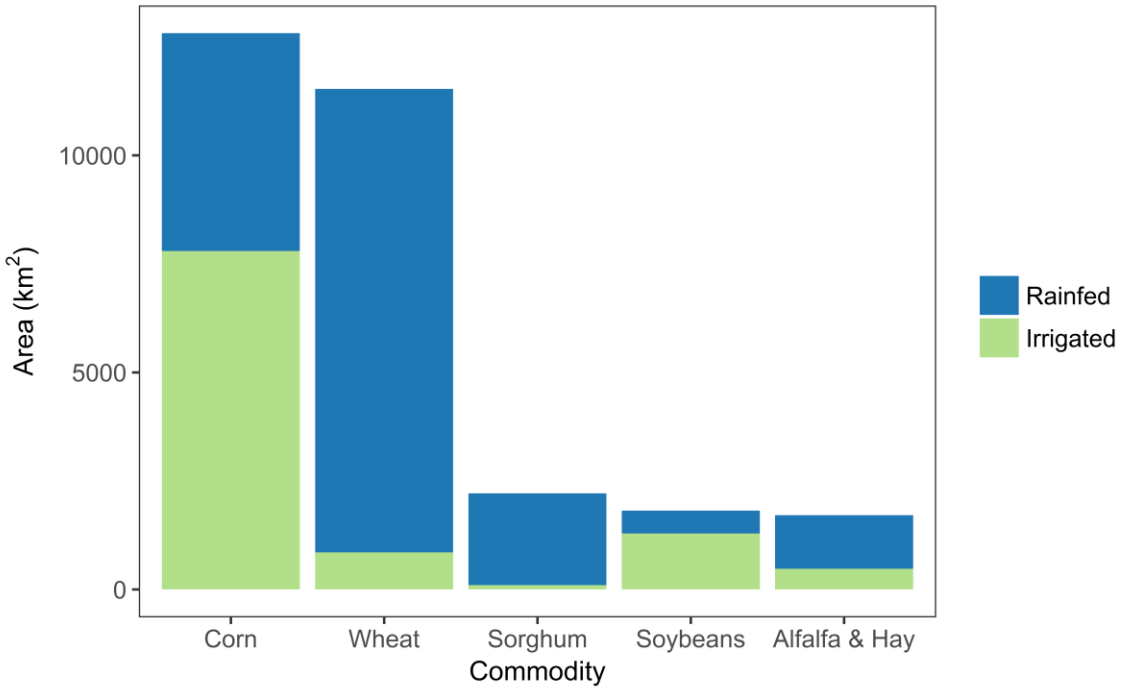

Figure A.1.2. Dominant crops in the study area. The summed areas for all 35 counties completely contained within the buffered study region are shown. Colors indicate irrigation status. Data from USDA NASS county statistics for 2007 [*NASS*, 2017].

Text A.1.2. Satellite Images

Sixteen Landsat scenes overlie the study region: Paths 28-29, rows 32-33, and paths 30-33, rows 31-33 (Figure A.1.3). Three Landsat satellite sensors were operational during the study period: Landsat 5 Thematic Mapper (TM), Landsat 7 Enhanced Thematic Mapper-plus (ETM+), and Landsat 8 Operational Land Imager (OLI). Vegetation metrics derived from the three sensors have been found to be comparable without modification [Vogelmann et al., 2015], particularly for Surface Reflectance (SR) products that have been terrain, radiometrically, and atmospherically corrected [*USGS*, 2017b, 2017c]. All available SR scenes from the three satellites covering the study area between 1 November 1998 and 31 October 31 2016 were used, accessed through Google Earth Engine's (GEE) [*Gorelick et al.*, 2017] LANDSAT/LT5_SR, LANDSAT/LE7_SR, and LANDSAT/LC8_SR image collections. Because winter wheat has an

initial green-up stage in November prior to peak harvest season in April and May, we considered each nominal crop year to run from 1 November of the previous year through 31 October of the nominal year. This ensured that annual composites for each year included only greenness related to crops harvested during the nominal composite year.

Clouds and cloud shadows were masked using the "cfmask" quality band provided with SR data products. We applied a negative 3 km buffer to all scenes prior to mosaicking to remove bad pixels along scene edges. Each satellite has a 16 day revisit cycle. There were two operating satellites orbiting at an 8-day offset in all years except 2012. This 8-day interval was simultaneously augmented by side-overlapping scene edges (imaged one week apart) and reduced by masked pixels due to cloud contamination, occasional poorly registered or missing scenes, and Landsat 7's scan line corrector failure in 2003, which results in the loss of approximately 22% of each scene [*Chen et al.*, 2011]. Satellite systems, scene totals, and pixel observation frequencies after buffering and cloud masking for each year are shown in Figure A.1.4.

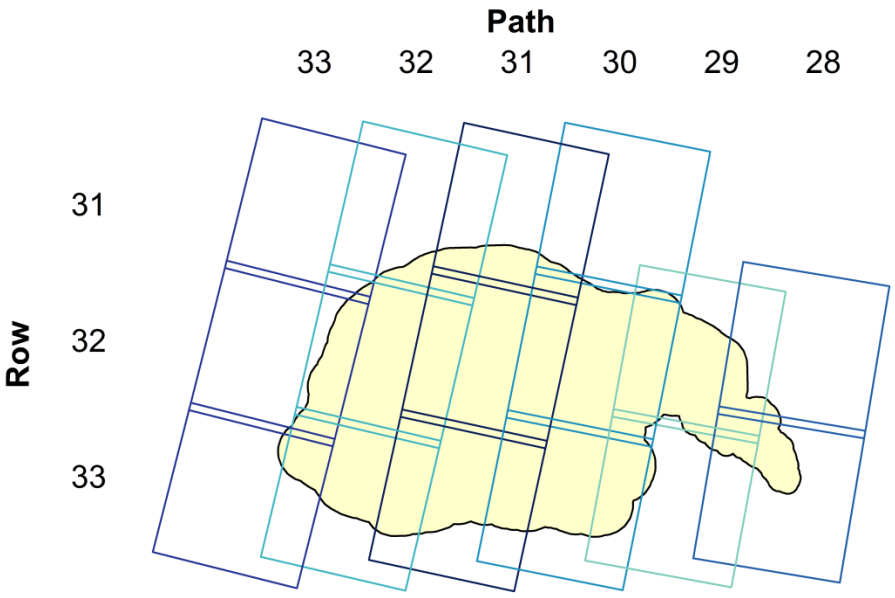

Figure A.1.3. Landsat scenes overlying the buffered study region. The Landsat Worldwide Reference System (WRS) 2 is used to locate each scene with a unique row-path identifier. Sixteen scene footprints are displayed over the buffered study area, colored by path number for clarity. Adjacent, side-overlapping paths are imaged seven days apart. Overlap on the top and bottom of scenes in the same path are duplicate data points; duplicates were removed for pixel observation statistics (Figure A.1.4).

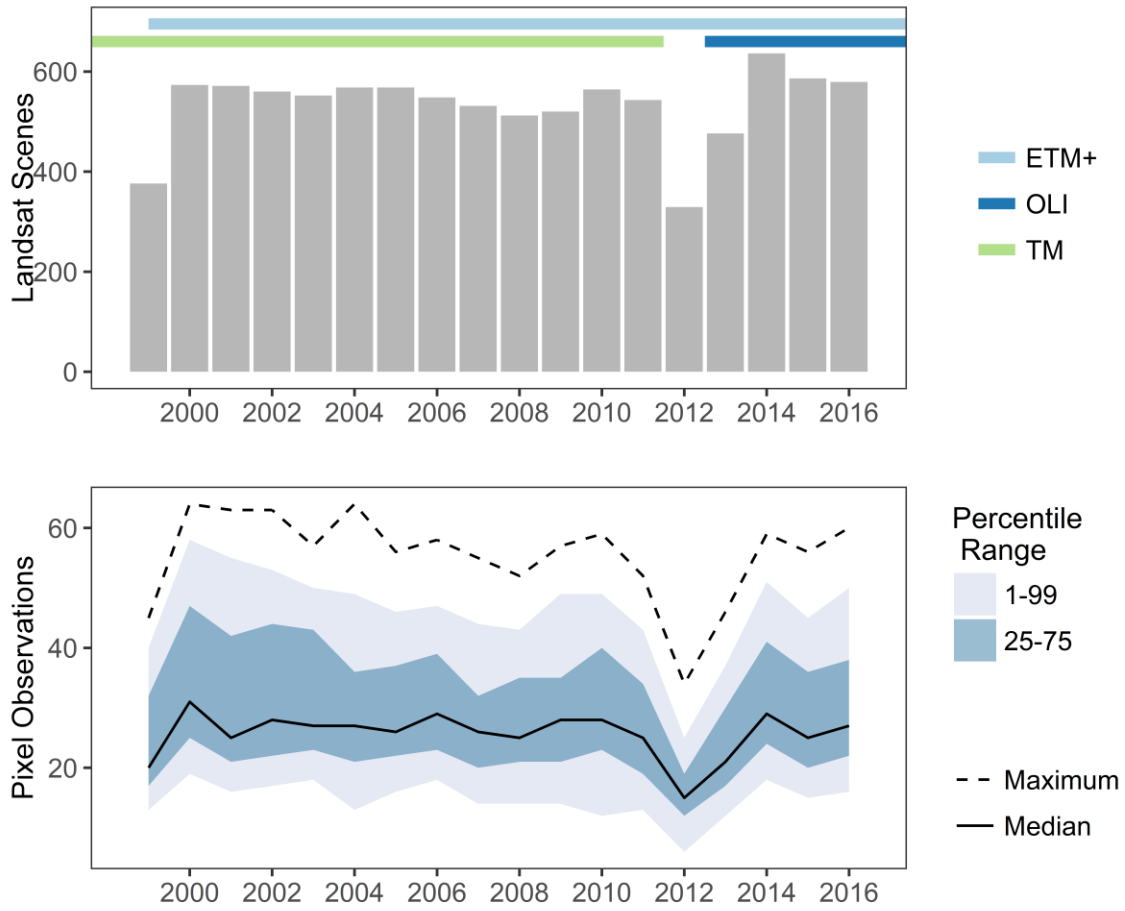

Figure A.1.4. Landsat imagery statistics. Top: Number of Landsat Surface Reflectance scenes used each year. Colored bars denote the active life of each satellite sensor. Bottom: Median, maximum, interquartile range (25-75%), and 1-99% range for number of satellite observations per pixel across the buffered study region over time. Duplicate observations in adjacent rows along the same path were removed for pixel observation statistics.

Text A.1.3: Vegetation Indices

Four vegetation indices were used: (1) the Normalized Difference Vegetation Index (NDVI), a consensus index for irrigation mapping that has performed well in previous studies [see review in *Ozdogan et al.*, 2010]; (2) the Enhanced Vegetation Index (EVI), which is known to have a wider functional range than NDVI before saturation in dense canopies [*Huete et al. 2002*]; (3) the Normalized Difference Water Index (NDWI),which is sensitive to the water content in plants [*Gao*, 1996]; and (4) a less common Green Index (GI) [*Gitelson et al.*, 2005] , which has been found to be particularly sensitive to plant greenness differences due to irrigation [*Ozdogan and Gutman*, 2008; *Ozdogan et al.*, 2010a]. The following equations were used:

$$NDVI = (NIR - Red) / (NIR + Red) \quad (1)$$

$$EVI = 2.5 * (NIR - Red) / (NIR + 6 * Red - 7.5 * Blue + 1) \quad (2)$$

$$NDWI = (NIR - SWIR) / (NIR + SWIR) \quad (3)$$

$$GI = NIR / Green \quad (4)$$

where NIR is the near-infrared band, SWIR is the short-wave infrared band 1 (band 5 for TM/ETM+, band 6 for OLI), and Red, Green, and Blue are the visible red band, visible green band, and visible blue band, respectively. Equation (2) uses constants optimized for MODIS sensors following Huete et al. [2002].

Text A.1.4: Environmental Covariables

We used ten environmental covariables to provide climate, soil, and terrain context for irrigated and non-irrigated locations for the random forest classification. Variables fell into two categories: (1) static variables describing terrain and soil properties used for all classification years, and (2) yearly variables describing precipitation and dryness conditions for each year. Terrain slope was calculated from the 1/3 arc-second resolution USGS National Elevation Dataset [*USGS,* 2012], averaged over a 150 m radius moving kernel to capture mean slope for agriculturally relevant field sizes, and resampled to 30 m Landsat resolution. Total plant available water storage (PAW) was obtained from the USDA SSURGO Web Soil Survey [*NRCS,* 2016], calculated as field capacity minus wilting potential, and was included as both the volume fraction in the top 25 cm (PAW-Vol) and as cm in the top meter (PAW cm). Daily gridded precipitation and potential evapotranspiration (PET) data at 4 km resolution was obtained from GRIDMET [*Abatzoglou,* 2013] and processed into 5 input layers: (1) total annual growing season precipitation (Ppt-Grow), defined as precipitation from 1 December of the previous year through 31 August of the nominal year to capture precipitation relevant to that year's soil replenishment and crop growth; (2) total late season precipitation (Ppt-Late), defined as precipitation from 1 May – 31 August for each year to capture precipitation from primary green-up through harvest; (3) total early season precipitation (Ppt-Early), defined as precipitation from 1 December of the previous year through 30 April of the nominal year as a proxy for soil water replenishment; (4) the ratio of early precipitation to PAW cm, a similar indicator of water availability in the soil prior to the growing season (Ppt-PAW); and (5) growing season aridity (Aridity), calculated as the total precipitation divided by PET from 1 May – 31 August. Finally, we obtained mean Palmer Drought Severity Index numbers [*Abatzoglou et*

al., 2014] for both the primary green-up through harvest time period (PDSI-Late, 1 May – 31 August) and the full growing season timeframe (PDSI-Grow, previous1 December – 31 August) time periods. Table A.1.1 summarizes these 10 input variables along with the 8 Landsat composites and 2 novel indices (AGI and WGI). All inputs were accessed through Google Earth Engine (GEE)'s data archive except the SSURGO soil data, which was manually uploaded to GEE for classification.

Table A.1.1 Summary of variables used in random forest classification of satellite imagery. All inputs were accessed through Google Earth Engine (GEE)'s data archive except the SSURGO soil data, which was manually uploaded to GEE for classification.

Variable	Short Name	Type	Time Period	Res (m^2)	Source
Maximum annual EVI	EVI max	yearly	1 Nov - 31 Oct	30	Landsat SR
Maximum annual GI	GI max	yearly	1 Nov - 31 Oct	30	Landsat SR
Maximum annual NDVI	NDVI max	yearly	1 Nov - 31 Oct	30	Landsat SR
Maximum annual NDWI	NDWI max	yearly	1 Nov - 31 Oct	30	Landsat SR
Annual range in EVI	EVI range	yearly	1 Nov - 31 Oct	30	Landsat SR
Annual range in GI	GI range	yearly	1 Nov - 31 Oct	30	Landsat SR
Annual range in NDVI	NDVI range	yearly	1 Nov - 31 Oct	30	Landsat SR
Annual range in NDWI	NDWI range	yearly	1 Nov - 31 Oct	30	Landsat SR
Growing season precipitation	Ppt-Grow	yearly	1 Dec - 31 Aug	4000	GRIDMET
Late season precipitation	Ppt-Late	yearly	1 May - 31 Aug	4000	GRIDMET
Early season precipitation	Ppt-Early	yearly	1 Dec - 30 Apr	4000	GRIDMET
Annual PDSI	PDSI-Grow	yearly	1 Dec - 31 Aug	4000	Abatzoglou et al. 2014
Growing season PDSI	PDSI-Late	yearly	May 1-Aug 31	4000	Abatzoglou et al. 2014
Terrain slope	Slope	static	NA	30	USGS NED
Soil Plant Available Water (fraction)	PAW-Vol	static	NA	30	SSURGO
Soil Plant Available Water (cm)	PAW-cm	static	NA	30	SSURGO
Precipitation:paw	Ppt-PAW	yearly	1 Dec - 30 Apr	30	Derived: Ppt-Early * PAW-cm^{-1}
Aridity	Aridity	yearly	1 May - 31 Aug	4000	Derived: Ppt-Late * PET^{-1}
Water-adjusted green index	WGI	yearly	1 Nov - 31 Oct	30	Landsat SR (NDWI*GI)
Aridity-normalized green index	AGI	yearly	1 Nov - 31 Oct	30	Derived: GI max * $Aridity^{-1}$

Text A.1.5: Expanded training data methodology

As noted in the main text (*Training data*), we used high-resolution (1 m) aerial imagery [*NAIP*, 2017], Landsat GI and EVI times series, and crop type maps (CDL) [*USDA-NASS*, 2017] to develop the training dataset. To identify irrigated and non-irrigated locations for target crop classes, we first produced a set of monthly imagery containing maximum GI and EVI per month for each training year (2010 and 2012) within GEE. These monthly time series were used to provide more information about greenness patterns on the landscape, including the timing of peak greenness to verify crop type. Pixels lacking valid observations in any month between April – November due to cloud contamination and Landsat's relatively long revisit cycle (text S2) were masked so that training points were targeted in areas with maximum information about seasonal greenness magnitudes and timing (Figure A.1.5). We then based irrigation status decisions on multiple lines of evidence such as irrigation infrastructure, seasonal greenness patterns, and crop type as indicated by CDL. For example, irrigated corn training points were generated by first locating candidate corn fields using CDL. The Landsat monthly time series was then checked to confirm CDL classification accuracy. Visual cues were then assessed, such as visible irrigation infrastructure and maximum greenness levels. A circular field with center pivot equipment, then, also had to have a temporal time series of monthly Landsat greenness that supported an irrigated status. Irrigated fields could also be square with center pivot equipment (with assumed corner extenders), square or other shapes with visible lateral move irrigation systems, or very green with evidence of irrigation ditches and surface irrigation. The expert selected method allowed us to target cases on the landscape that provided higher certainty, such as adjacent fields of the same crop type with vastly different greenness patterns and presence/absence of visible infrastructure. Figure A.1.5 provides training dataset locations and

Table A.1.2 provides numerical breakdown of training points by crop and climate region. Training points are available as indicated in the Acknowledgments and Data section of the main text.

With recent computational advancements such as GEE, the acquisition or generation of accurate training, validation, and test datasets that reflect local crop types, irrigation systems, climate, and non-crop ecosystems is becoming a key limiting step for remote sensing classification of irrigated area across wide regions. To apply our approach to new areas, one need not replicate our specific method to generate training points. One only needs a robust spatial dataset that accurately denotes irrigated and non-irrigated locations for specific growing seasons, ensuring that these points adequately sample the dominant crop types, non-crop ecosystems, and climate zones in the region of interest.

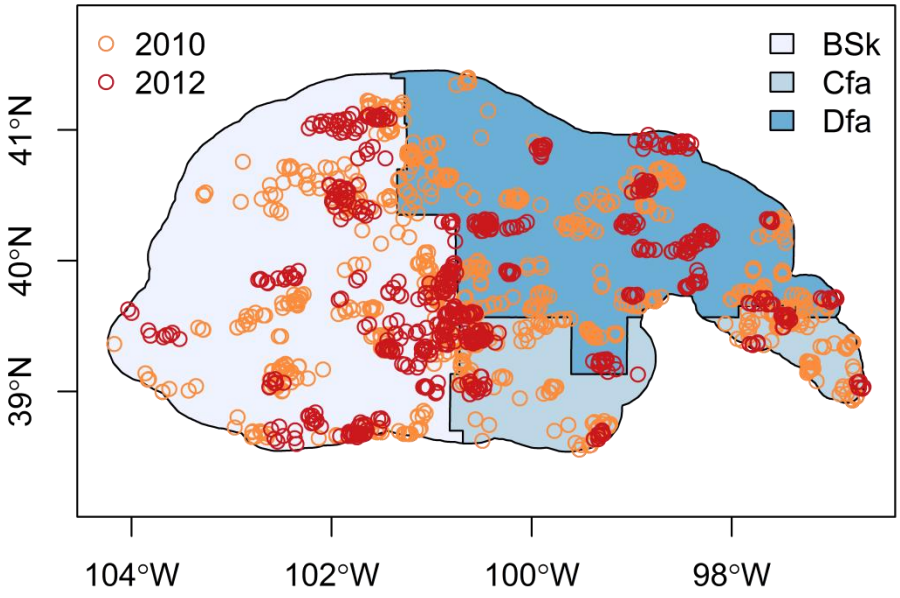

Figure A.1.5: Training point locations. Training point locations overlaid on Koeppengeiger climate regions [*Peel et al.*, 2007]. Climate types in the buffered study region include cool arid steppe (BSk) along with humid subtropical (Cfa) and humid continental (Dfa) temperate zones.

Table A.1.2. Number of training data by crop type, climate region, and year. In total, 1401 training points were sampled. Climate regions are described by Koeppengeiger codes [*Peel et al.*, 2007] and include cool arid steppe (BSk) along with humid subtropical (Cfa) and humid continental (Dfa) temperate zones.

	Climate	2010		2012	
Crop Type	Region	Irrigated	Not Irrigated	Irrigated	Not Irrigated
Alfalfa	BSk	30	17	32	18
	Cfa	15	24	3	5
	Dfa	34	30	24	22
Corn	BSk	30	30	31	30
	Cfa	20	20	12	11
	Dfa	30	30	30	31
Sorghum	BSk	14	30	7	30
	Cfa	14	20	6	9
	Dfa	12	30	6	18
Soy	BSk	30	4	30	0
	Cfa	21	8	8	2
	Dfa	30	26	28	9
Wheat	BSk	25	30	26	39
	Cfa	11	21	3	12
	Dfa	15	30	3	24
Fallow	BSk	-	15	-	16
	Cfa	-	7	-	5
	Dfa	-	15	-	7
Grassland /	BSk	-	30	-	27
Non-crop	Cfa	-	24	-	10
	Dfa	-	30	-	25
Total		331	471	249	350

Text A.1.6: Validation dataset for algorithm development

Initial classification algorithm development was performed in the Middle Republican Natural Resource District (NRD) in Nebraska within the GRB study area. We created a validation dataset for this purpose by generating random points stratified across crop types and grassland according to Cropland Data Layer (CDL) maps [*USDA-NASS*, 2017] for 2007 and 2010 within the NRD. Points were marked as "irrigated" or "non-irrigated" in the same manner as the training and test data (see Methods and Text A.1.5), except that point locations were randomly generated in contrast to the expert selected training points. Similar to the test point dataset (see *Accuracy assessment and analyses*), cases where no clear determination could be made were marked as "uncertain" and removed from the validation dataset. This resulted in 878 points for 2007 and 864 points for 2010. This provided a point dataset with higher point density for the smaller NRD region to evaluate classification decisions during algorithm development.

In this study, our objective was to identify irrigated area regardless of crop type. Due to the similarity in greenness between irrigated corn and some fields of rainfed soy (Figure 1.2a), the accuracy of the training algorithm was reduced when dryland soy training points were included. Because rainfed soy is relatively rare on the landscape of our study system (Figure A.1.2), we selected the random forest classifier that omitted rainfed soy training points, as it performed best in accuracy metrics and qualitative evaluation (not shown).

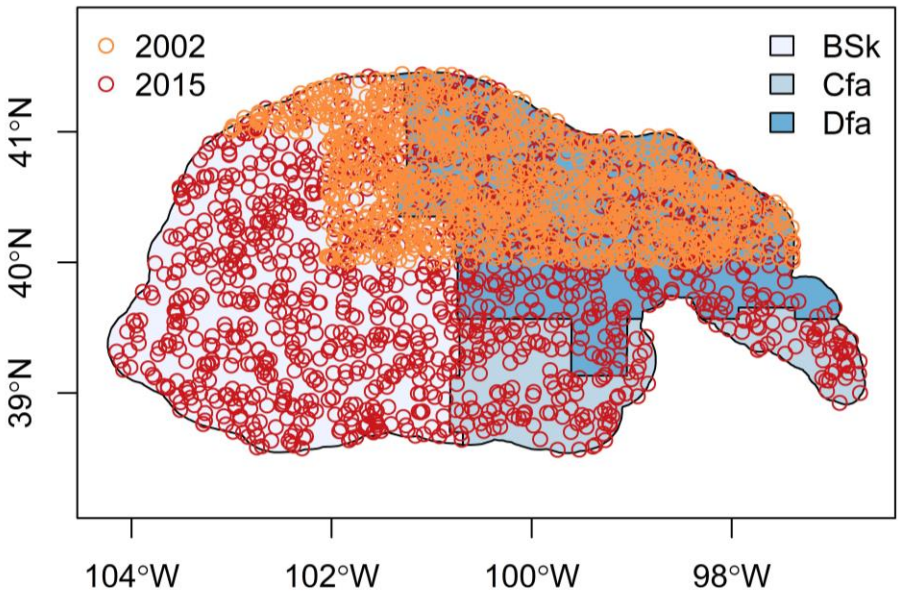

Figure A.1.6: Test point locations for accuracy assessment. Randomly generated test point locations overlaid on Koeppengeiger climate regions [*Peel et al.*, 2007]. Points in 2002 were restricted to Nebraska, since only Nebraska had crop type maps (NASS Cropland Data Layers, [*USDA-NASS*, 2017]) available before 2006. Climate types in the study region include cool arid steppe (BSk) along with humid subtropical (Cfa) and humid continental (Dfa) temperate zones.

Table A.1.3. Point-based accuracy of irrigation classification. Accuracy metrics for 2002 and 2015 test datasets. Parenthetical numbers give the number of points for each category. Precipitation represents December – August totals to capture agriculturally relevant precipitation (Text A.1.3). Omission errors describe the percentage of training points in each class that were not classified in that class (false negatives), while commission errors describe the percentage of training points that were not in that class but were predicted to be by the random forests classifier (false positives).

Year	Precip. (mm)	Class	Omission Errors	Commission Errors	Overall Accuracy
2002[+]	203	Irrigated	6.1% (14/229)	0% (0/215)	98.6% (1005/1019)
		Non-Irrigated	0% (0/790)	1.7% (14/804)	
2015[*]	484	Irrigated	7.6% (16/210)	6.3% (13/207)	97.6% (1194/1223)
		Non-Irrigated	1.3% (13/1013)	1.6% (16/1016)	
Combined		Irrigated	6.8% (30/439)	3.1% (13/422)	98.1% (2199/2242)
		Non-Irrigated	0.7% (13/1803)	1.6% (30/1820)	

[+]Active satellite sensors: Landsat 5 TM and Landsat 7 ETM+
[*]Active satellite sensors: Landsat 7 ETM+ and Landsat 8 OLI

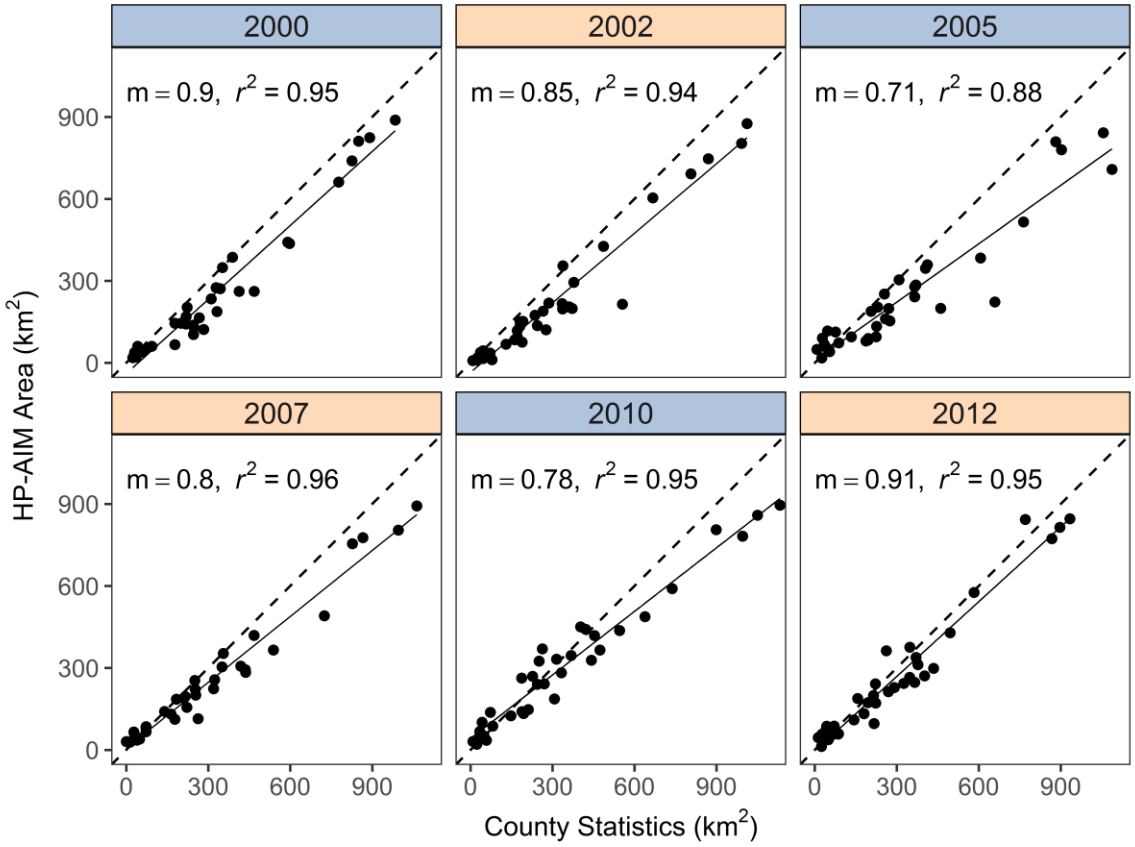

Figure A.1.7. County-level accuracy assessment. Irrigated area from county-level statistics compared with irrigated area from AIM classified maps against a 1:1 line (dashed). Agreement between the two datasets was assessed with r^2 metrics from simple linear regression (trend line = solid line). Data for blue panels are from the USGS [*USGS*, 2015], while data for tan panels are from the USDA National Agricultural Statistics Service [*NASS*, 2017] and include 35 counties fully contained within the buffered study region.

Text A.1.7. Variable importance

To understand the relative contribution of the 20 input variables to classification accuracy, we ran permutation tests and GINI index metrics in R since Google Earth Engine (GEE) does not currently support variable importance measures (GEE accessed May 2016 - July 2017). We used the randomForest package [*Liaw and Wiener*, 2002] with an identically parameterized random forest classification using 500 trees and identical training data. Agreement in predictions between the R and GEE classifiers was 99.91% for the point test data (2242 points), suggesting both classifiers functioned similarly and the variable importance metrics from

R likely mirror their importance in GEE. Because random forest algorithms use randomization in node optimization and bagging, we averaged importance scores for 20 runs with different random seeds to produce robust metrics. The permutation test measures variable importance by permuting the value of each variable over all trees and finding the resulting mean decrease in accuracy for class predictions made on the out-of-bag samples. If the variable is not important, randomly rearranging its values would have little effect on prediction accuracy. Larger decreases in accuracy following permutation are expected for variables with greater contributions to overall accuracy. The second metric used is the GINI Index, which measures the reduction in node impurities resulting from splitting on each variable. Scores for both metrics are shown in Figure A.1.8.

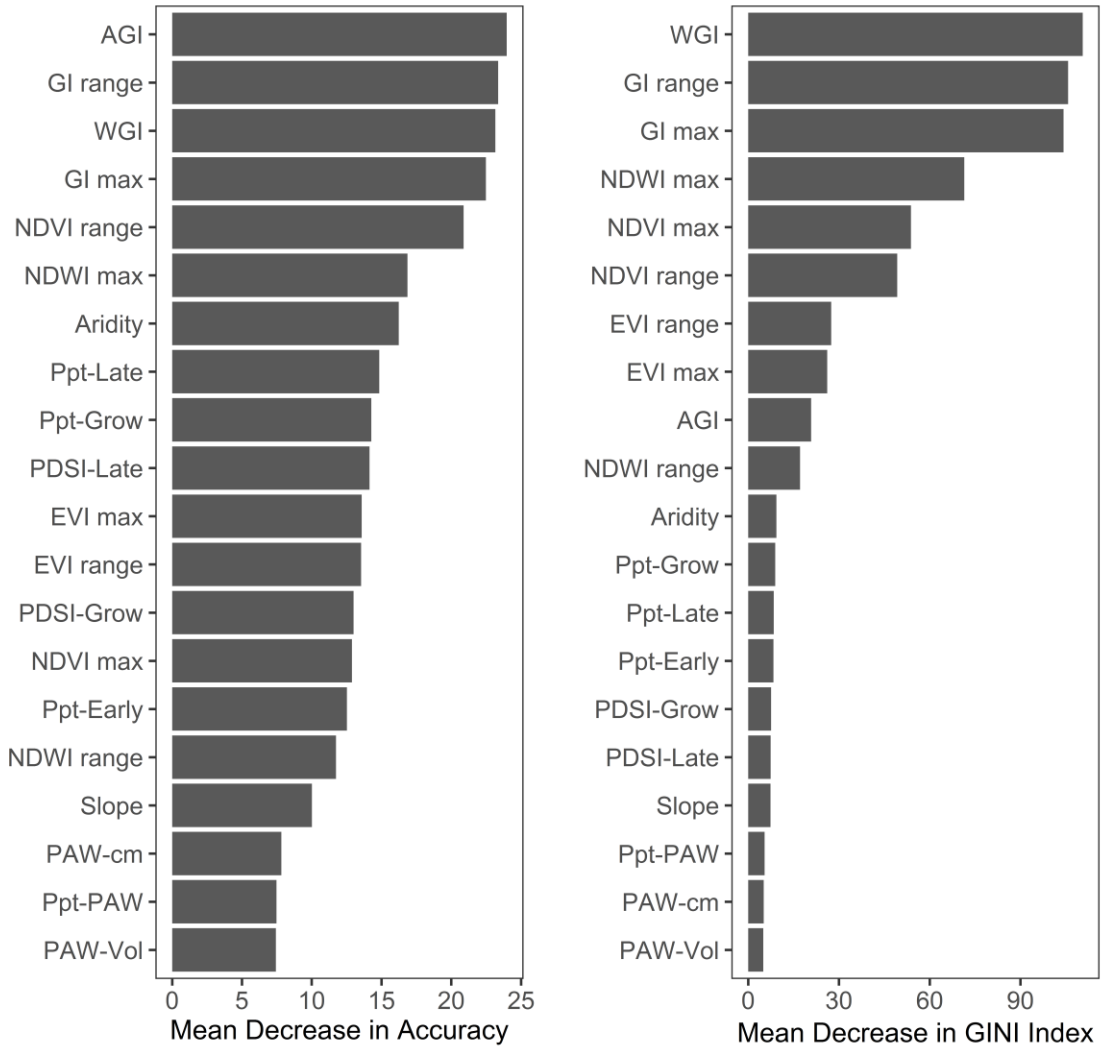

Figure A.1.8. Variable importance in the random forest classification. Variables are identified by their short name, which can be related to full names and source information using Table A.1.1. Left: The mean decrease in accuracy (standardized) across trees found through a permutation test. Right: The mean decrease in node impurity resulting from splits on each variable as measured by the GINI Index.

Text A.1.8. Drivers of irrigated area

We analyzed irrigation drivers for the RRB portion of the RRCA (RRB-RRCA, Figure A.1.1) for 1999-2015 because the irrigation volume dataset was restricted to this region of the RRB and time period. First, we ran a fully crossed correlation matrix in R on the following datasets, aggregated to the RRB-RRCA domain: (1) annual irrigated area from AIM-RRB; (2) annual irrigation volume from the RRCA groundwater model [*RRCA*, 2003]; (3) annual irrigation depth, found by dividing irrigation volume by irrigated area; (4) annual growing season precipitation (see Table A.1.1); (5) national annual corn price as a proxy for commodity prices [*NASS*, 2017], adjusted for inflation to 2016 dollars using the Consumer Price Index; and (6) annual crop price with a 1 year lag, referred to as price lag. Results are shown in Figure A.1.9a.

To get a sense of how variables with significant correlation to irrigated area interact, we built a preliminary predictive model using classification and regression tree analysis (CART) as CART is robust to auto-correlation often found in time series data and nonlinear interactions among predictor variables. CART was performed in R using the 'ctree' function in the party package [*Hothorn et al.,* 2006]. We set the split criteria at $p = 0.15$ with minimum node weights of 3 given our relatively short time series. Percent irrigated area was the dependent variable. We used price lag and precipitation for predictor variables, since correlation analysis and linear regression indicated that irrigated area was influenced by precipitation through the irrigation depth required (Figure A.1.9a & Figure 1.4e). The resulting CART tree is shown in Figure A.1.9b, and area vs. year trends given high or low prices based on the primary CART split is shown in Figure 1.4f.

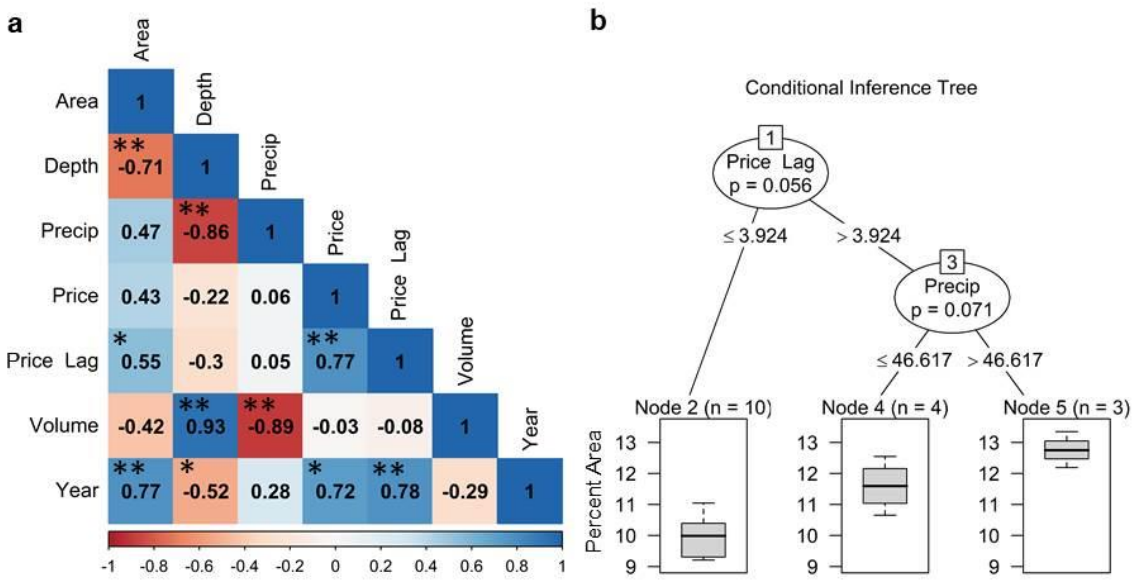

Figure A.1.9. Drivers of irrigated area. Correlation matrix for irrigated area and related variables. * indicates $p < 0.05$, ** indicates $p < 0.005$; (b) Conditional inference tree from CART model ($R^2 = 0.78$). Price lag refers to the previous year's corn commodity price in dollars per bushel from NASS [*NASS*, 2017]. Precipitation is 1 December – 31 August totals for each growing season.

CHAPTER 2:
MAPPING THREE DECADES OF IRRIGATION ACROSS THE HIGH PLAINS AQUIFER

Abstract

Understanding how irrigated areas change over time is vital to effectively manage limited agricultural water resources, but long-term, high-resolution, and spatially-explicit datasets are rare. The High Plains Aquifer (HPA) in the central United States is one of the largest and most stressed aquifer systems in the world. It supports a $20 billion economy, but groundwater use is unsustainable over much of the aquifer. Emerging cloud computing tools like Google Earth Engine (GEE) make it possible to leverage the full Landsat record to monitor regional systems like the HPA with high spatial and temporal resolution over multiple decades. Challenges remain, however, to develop irrigation classification methods that are robust to wide range of climate, crop types, and evolving crop varieties and management, along with missing data. Here, we address these challenges to produce annual, moderately high resolution (30 m) irrigation maps from 1984-2017 over the aquifer using random forest classification. Leveraging GEE's extensive data catalog, we combined Landsat imagery with climate and environmental covariables to create a single random forest classifier. A novel Neighborhood Greenness Index contributed to a 91.4% map accuracy across years. Spatially refined trend analysis of irrigated area through time identified regions of stable, expanding, and declining irrigated area. Given declining aquifer conditions, we estimate that up to 24% of irrigated area may be lost this century. The map dataset is the longest, most spatially-refined record of where and when irrigation occurs in the world. It is freely available for stakeholders, managers, and researchers to inform future policies and management, as well as for use in hydrology, agronomy, and climate models.

1. Introduction

Irrigation strongly affects food production and the water cycle worldwide, yet for such a critical water use, little is known about exactly where and when irrigation occurs. Globally, 85 to 90 percent of consumptive water use [*Shiklomanov,* 2000; *Doll,* 2009] goes towards irrigating approximately 20% of croplands, enabling irrigated land to contribute ~40% of global food production [*Abdullah,* 2006]. Extensive irrigation dramatically alters the water cycle, depleting aquifers [*Gleeson et al.*, 2012b] and surface water bodies [e.g., *Conrad et al.*, 2016] while increasing atmospheric moisture content with demonstrable downstream climate effects [*Lobell et al.*, 2006; *Pei et al.*, 2016]. Maintaining and even expanding irrigation is required to meet the twin challenges of increasing food demand from demographic changes [*Tilman et al.,* 2011] and water stresses in agricultural systems due to climate change and enhanced variability [*Döll,* 2002; *Wada et al.*, 2013]. At the same time, depleted water resources may curtail the future ability to irrigate existing croplands [*Elliott et al.*, 2014; *Cotterman et al.*, 2018].

Spatially-explicit knowledge of the historical trajectories of irrigated areas is needed to inform planning and management for sustainable local economies and food security [*Peña-Arancibia et al.*, 2014; *Abuzar et al.*, 2015], as well as to provide inputs necessary to improve surface energy balances in climate and earth systems models [*Ozdogan et al.*, 2010b]. Unfortunately, existing irrigation data do not meet the needs of scientists and resource managers alike due to a combination of factors, including low spatial and temporal resolution, limited spatial extent (localized studies), and/or potentially biased statistical summaries within political units.

Satellite observations provide an opportunity to detect and map irrigation at sub-field-scale resolution on an annual basis to inform agricultural water management. Until recently,

efforts to remotely sense irrigated lands at regional scales were limited to either coarse resolution (250 to 1000 m) sensors or short time frames, largely restricted to single or small subsets of years. Opening of the higher resolution (30 m) Landsat archive free-of-charge [*Woodcock et al.*, 2008; *Wulder et al.*, 2012] and new cloud computing tools like Google Earth Engine (GEE) [*Gorelick et al.*, 2017] allow estimation of land use tailored to regions of interest with high spatial and temporal resolution over multiple decades [*Azzari and Lobell*, 2017; *Wulder et al.*, 2018]. The resulting increased accessibility, computational efficiency, and ease of integration with supporting datasets via the co-located data catalog have created an unparalleled opportunity to increase our understanding of the spatial and temporal aspects of irrigation. Although barriers such as cost and processing time have been dramatically lowered, substantial technical challenges remain due to data gaps in the Landsat record as well as the diversity of irrigated characteristics across crop types and regions.

Irrigated agriculture is a dynamic and challenging land use class for regional-scale remote sensing due to wide climate gradients, increasingly diverse crop types, and evolving management practices over time. While many of these challenges are shared with mapping croplands, irrigation classification from satellite observations is relatively underdeveloped [*Ozdogan et al.*, 2010a], particularly for multi-year monitoring [*Abuzar et al.*, 2015]. Irrigated fields are distinguishable from rainfed fields via a variety of spectral characteristics and indices, most prominently greenness, wetness, and thermal properties. Most commonly, studies have used differences in greenness as measured by vegetation indices during the peak growing season as the key differentiator between irrigated and rainfed crops [*Ozdogan et al.*, 2010a], but these differences diminish in humid environments [*Pervez and Brown*, 2010]. Peak greenness is also crop specific, with greenness thresholds indicative of irrigation often overlapping among crops,

particularly as crop diversity increases as study areas expand to regional scales. Similarly, the timing of peak greenness varies by crop type, and many regions have more than one crop season that needs to be captured. Finally, innovations in crop choice and management can cause apparent greenness to change over time, including new crop varieties, expansion of deep green crops like soybeans [*Lark et al.*, 2015], and evolving irrigation technologies. Finally, limited historical ground-truth data are available to characterize such dynamic greenness issues within classification algorithms.

To capture differences in greenness due to irrigation status, cloud-free satellite imagery needs to be obtained during peak growth for all major crops in a region. Not controlling for optimal timing of satellite observations can lead to false negatives, since the apparent peak greenness would be artificially suppressed. Studies focusing on the very recent past have the advantage of leveraging virtual constellations of simultaneously operating moderate to high resolution systems such as Sentinel and Landsat [e.g., *Pastick et al.*, 2018], increasing the likelihood of capturing key growth stages. Other studies have combined lower resolution, higher frequency MODIS observations with Landsat [e.g., *Peña-Arancibia et al.*, 2014], thus limiting focus to years following MODIS's launch in 2000.

Long-term studies, however, are limited due to fewer operational systems and less consistent historical data acquisition. Landsat is the only continuous, relatively high resolution (30 m) dataset longer than 30 years [*Wulder et al.*, 2008], but high quality satellite observations are sparse in the early Landsat record. The Landsat constellation provides imagery on an 8-16 day revisit cycle (the 16-day single-satellite cycle is offset when two satellites orbit simultaneously, thus providing an 8-day revisit period), but available imagery has considerable temporal gaps due to cloud cover, poorly registered scenes, and inconsistent data acquisition and

storage prior to 1999 [*Arvidson et al.*, 2001; *Gutman et al.*, 2013]. This presents challenges for consistently obtaining annual satellite observations during critical windows of the crop season, which are needed to accurately distinguish between irrigated and rainfed crops.

Recent studies leveraging the full Landsat archive for other applications address these data gaps in several ways. Robinson et al. [2017] derived a 30 m, 16-day NDVI time series from 1984-2016 for the conterminous United States by filling missing data with median NDVI observed in previous years and subsequent smoothing. Hermosilla et al. [2015] filled gaps in annual composites from 1998-2012 with synthetic proxy values by combining noise detection and removal, values from adjacent years, and breakpoint analysis. These methods, however, require relatively static land cover during the analysis periods to fill missing data [*Robinson et al.*, 2017]. Image filling approaches, therefore, are not well suited for irrigation mapping, where drastic differences in greenness due to different crop rotations, fallowing, annual climate, and irrigation decisions that are common among adjacent years. Other gap filling approaches that may be more suitable for irrigation mapping include harmonics, Fourier analysis, or linear interpolation to fill missing data within each year's phenological curve, but these perform poorly without adequate observations throughout the year [*Hermosilla et al.*, 2015]. Particularly in the early (pre-1999) Landsat record, pixels missing data during peak growing seasons tend to have sparse data for the full year.

Here, we addressed these challenges to produce a 34-year record (1984-2017) of irrigation across the entire High Plains Aquifer (HPA) in the central United States (Figure 2.1). The HPA covers more than 450,000 km^2, spanning a wide East-West precipitation gradient and a latitudinal gradient that result in diverse agricultural practices across the region. It is also one of the largest and most stressed aquifer systems in the world [*Gleeson et al.*, 2012b]. The aquifer

supports a $20 billion agricultural economy [*Ashworth*, 2006], but water-level declines threaten the continued viability of irrigated agriculture dependent upon the aquifer [*Scanlon et al.*, 2012; *Haacker et al.*, 2016]. Despite this, it contains some of the most dynamic irrigated area in the country, accounting for 97% of US irrigation expansion between 2002 and 2007 [*Brown and Pervez*, 2014]. Like many areas worldwide [*Wada et al.*, 2016], this increasingly scarce but valuable resource needs to be more efficiently managed to preserve water storage, which is critical for local economies and global food supply. In the past several decades, a diverse matrix of management areas, policy interventions, and court litigation have emerged to slow aquifer depletion rates [*Smidt et al.*, 2016]. Knowledge of when and where irrigation occurs in the HPA is vital to better understand past water use, evaluate management, and improve regional crop, hydrology, and climate models to support future planning.

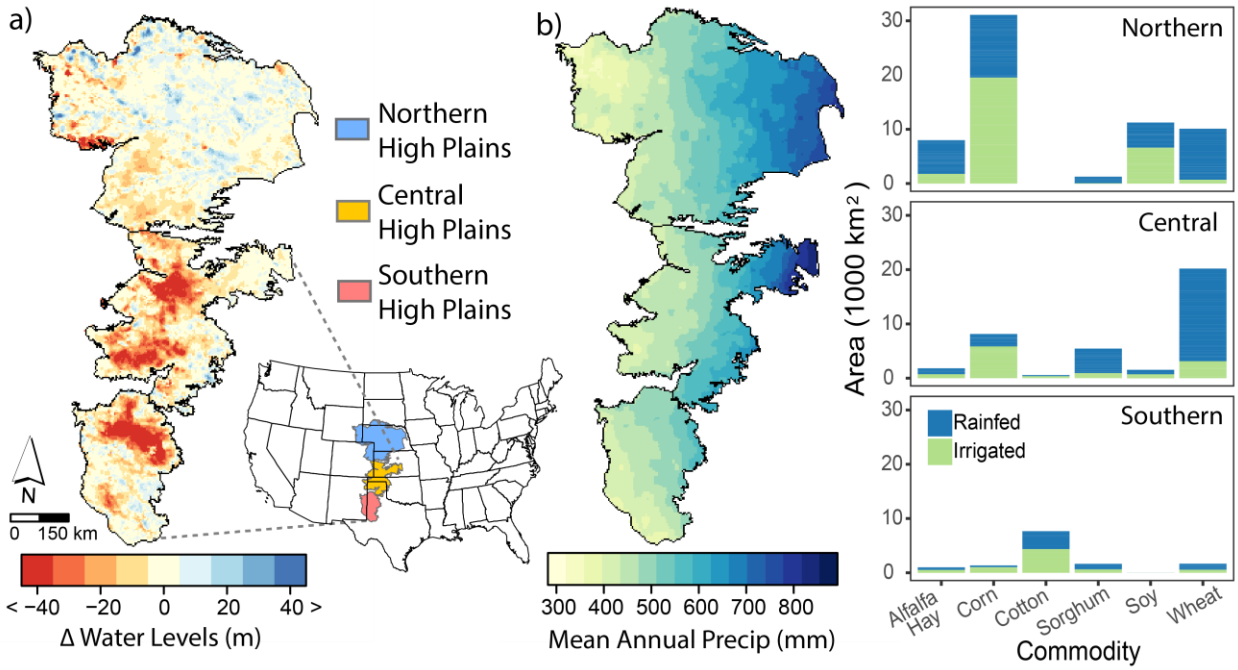

Figure 2.1 Study area: the High Plains Aquifer (HPA). (a) Study region with major aquifer regions delineated, and change in groundwater levels from predevelopment to 2016 based on [*Haacker et al.*, 2016]. (b) Variation in precipitation and crop types across the HPA. Mean annual precipitation from 1984-2017 derived from [*Abatzoglou*, 2013]. Regional summaries of crop-specific irrigated and rainfed crop area for 2012 derived from national statistics [*NASS*, 2017].

Building upon past efforts that focused on a single river basin in the northern HPA for 1999-2016 [*Deines et al.*, 2017], we use 34 years of Landsat imagery from the entire HPA to address the following challenges: 1) classifying irrigation across a dynamic climate gradient in temperature and precipitation; 2) incorporating changes in plant phenology across multiple crop types; 3) creating annual maps despite gaps in imagery during critical growing periods; and 4) incorporating a diverse set of training and validation points spanning multiple decades and jurisdictions. Specifically, we applied kernel-based filtering to quantify neighborhood greenness contrast; identified critical crop-specific observation windows based on phenology and growing-degree days to minimize false negatives in irrigation classification; and quantified the uncertainty introduced by post-classification filling of missing years on a per-pixel basis. Changes in annual maps were then analyzed to better understand irrigation dynamics in this highly stressed aquifer system. The resulting map dataset, hereafter termed the Annual Irrigation Maps – High Plains Aquifer (AIM-HPA), is to date the longest, most spatially-refined record of where and when irrigation occurs in the world. AIM-HPA is freely available for stakeholders, managers, and researchers to inform future policies and management, as well as for use in hydrology, agronomy, and climate models.

2. Methods

2.1. Study area

The HPA underlies 450,660 km^2 of eight states in the central United States (Figure 2.1). By convention, it is often subdivided into three main regions based on physical properties and logistical legacies: the Northern High Plains (NHP), Central High Plains, (CHP), and Southern High Plains (SHP; Figure 2.1) [*Weeks et al.*, 1988; *Stanton et al.*, 2011]. Aquifer development for irrigation began in the Texas SHP in the 1930s and subsequently spread northward [*Luckey et*

al., 1981]. As high-capacity wells became widespread in the 1950s [*Luckey and Becker*, 1999], groundwater extraction exceeded recharge over much of the CHP and SHP, resulting in substantial declines (Figure 2.1a). At current rates of depletion, portions of the aquifer will become unusable in the coming decades [*Haacker et al.*, 2016].

The HPA spans a wide climatological gradient, ranging from Koeppen-Geiger class Bsk (cool arid steppe) in the south and west to classes Cfa (humid subtropical) and Dfa (humid continental) in the east [*Peel et al.*, 2007]. Mean annual precipitation varies by nearly a factor of three along a longitudinal gradient, ranging from 315 mm in the west to 858 mm in the east averaged over the 1984-2017 study period (Figure 2.1b, summarized in GEE from GRIDMET gridded climate data [*Abatzoglou*, 2013]). Agriculture in this region is dominated by annual crops, approximately 30% of which are irrigated [*Scanlon et al.*, 2012]. Across the HPA, the major irrigated crops in decreasing order of total irrigated area are corn, soybeans, cotton, wheat, alfalfa/hay, and sorghum (Figure 2.1b) [*NASS*, 2017]. Soybeans are mainly grown in the more humid northeastern portions of the region, particularly in the NHP. Cotton, which is the dominant crop in the SHP, requires less water and is well suited to the more arid southerly climate [*West et al.*, 2018].

Groundwater declines have reduced groundwater well yields, increased pumping costs, and led to interstate conflict over shared water resources. Water management and rights doctrines vary substantially across the eight states and over time [*Smidt et al.*, 2016]. A diverse set of groundwater management approaches have emerged across the aquifer to slow depletion, making the region a hotbed for innovation in stakeholder-regulator partnerships [e.g, Chapters 3 & 4, this volume]. Irrigation technology has also changed considerably during the study period, with a large transition from flood irrigation to center pivots and then to low-pressure spray systems.

Uncertainty over the location of irrigated areas and their interannual variability limit the ability to fully quantify, evaluate, model, and manage water use in this region. In the United States, county-level statistics updated in five year cycles [*USGS*, 2015; *NASS*, 2017] provide large-scale trends in irrigated area, but these lack the spatial precision and temporal resolution needed for detailed analysis and modeling. The only existing multi-year map product covering the full HPA (MIrAD-US) distributes NASS county statistics in space with consistent methodology across multiple years [*Pervez and Brown*, 2010; *Brown and Pervez*, 2014]. However, reliance on MODIS satellite data limit these map products to a relatively coarse 250 m resolution for years 2002, 2007, and 2012; these products have estimated accuracies between 75-89% for the Great Plains region in 2002 and 2007 [*Pervez and Brown*, 2010; *Brown and Pervez*, 2014]. In the HPA, Landsat-based studies have produced higher resolution (30 m) data sets for localized regions with reasonable accuracy [e.g, *Dappen et al.*, 2007; *Deines et al.*, 2017] or single years across the HPA [*Qi et al.*, 2002] as early as 1982 [*Dappen and Merchant*, 2004]. This suggests that the Landsat archive provides a promising avenue to both recover historical irrigated locations across the region and develop a workflow for ongoing monitoring.

Here, our study area encompasses the full HPA. Image processing and classification was conducted on an expanded region (Figure A 2.1). The full region includes a 15 km buffer around the HPA along with the entirety of the Republican River Basin and portions of the Platte and Arkansas River basins in Colorado. These additions allow us to leverage Colorado's historical irrigation data sets for training and validation data (Text A 2.1). The irrigation classification was developed and evaluated on this full 608,260 km^2 region. Here, results are presented for the non-buffered HPA region, although the full buffered region datasets are available for download at http://hydroshare.org and through Google Earth Engine [DOI and links upon publication].

2.2. Annual image composites

Our method followed a "best available pixel" approach [*White et al.*, 2014] to generate annual composites for the study region based on available imagery. We then applied filters to exclude pixels from the annual composite that lacked observations during key crop periods (section 2.3). In this manner, the same image composites and method could be used across regions and crop types based on customized date filters, decreasing processing time and increasing the flexibility of the method.

The buffered study area underlies 50 Landsat scenes (Figure A 2.1). Working in Google Earth Engine's (GEE) cloud computing platform [*Gorelick et al.*, 2017], we used all available Landsat Collection 1 Tier 1 Surface Reflectance products [*USGS*, 2017a] between January 1, 1984 and October 15, 2017 to build yearly collections that target the crops harvested during the nominal year. These include images from Landsat 4 TM, 5 TM, 7 ETM+, and 8 OLI. Surface Reflectance products have undergone terrain, radiometric, and atmospheric corrections [*USGS*, 2017b, 2017c], and vegetation metrics derived from the three sensors have been found to be comparable without modification [*Vogelmann et al.*, 2015]. In total, 34,194 scenes were available with these specifications. As expected, we saw a marked decrease in scene availability prior to 1999 (Figure 2.2). It's worth noting that prior to opening the Landsat archive for free use, access to this number of scenes would have cost an estimated $67,829,100 in constant dollars based on year and sensor-specific costs [*Wulder et al.*, 2012; *Gartner*, 2018], which clearly would be prohibitive for such a study. Furthermore, without Google Earth Engine, an image acquisition, storage, and processing effort of this magnitude would require months of processing on many-core compute clusters, along with 100+ TB of high-speed data storage, and the expertise and tools necessary to handle data streams of this magnitude. GEE provides an end-

to-end platform that is freely accessible for research, education, and non-profit uses, greatly lowering the barrier to entry for remote sensing applications.

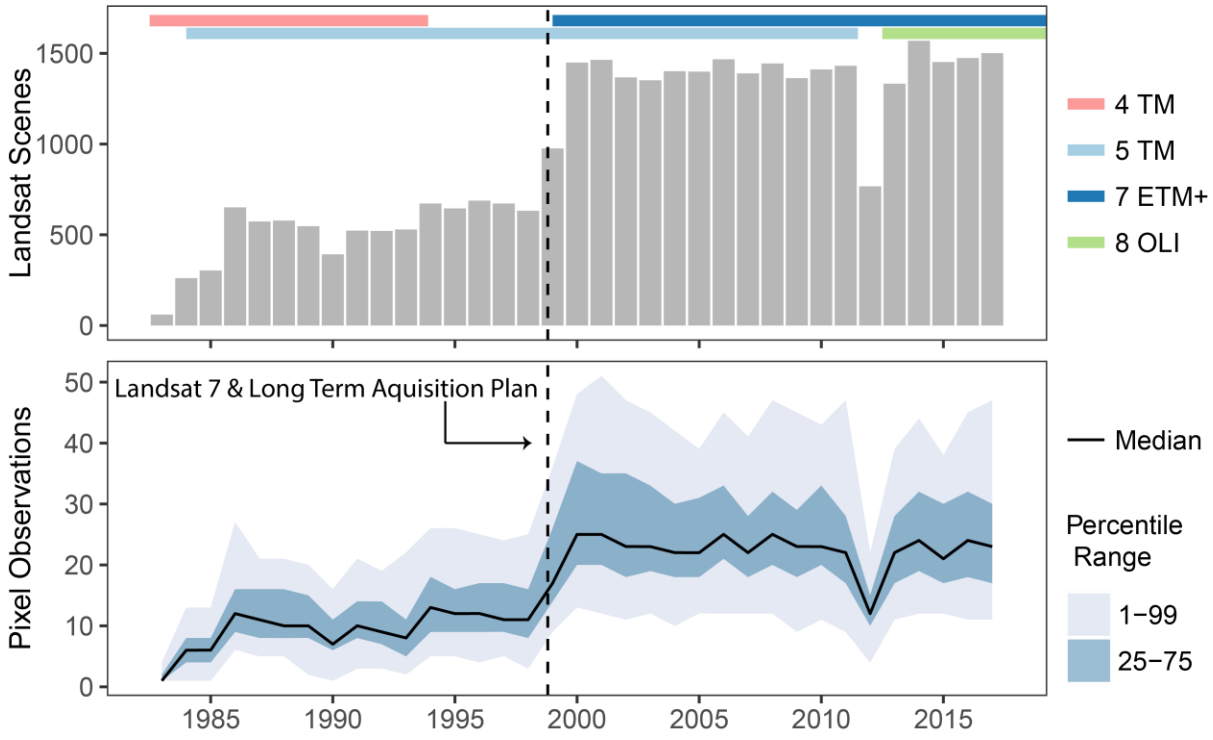

Figure 2.2 Limited Landsat availability prior to 1999 in the US High Plains Aquifer makes annual classification of irrigated area challenging. Top: the total number of Landsat scenes available in the study region by year. Bottom: summary statistics for the number of valid Landsat observations per 30 m pixel by year in the study region. Less frequent observations decrease the chances of capturing baseline and peak greenness in crops. The dotted line indicates the launch time for Landsat 7 and associated implementation of the Long Term Acquisition Plan [*Arvidson et al.*, 2001].

For each Landsat image, we masked clouds, cloud shadows, and snowy pixels based on the CFMASK-based quality band included with each scene. We then filtered out pixel observations with surface reflectance values outside of the valid range (0-10,000). We applied a negative 2 km buffer to all Landsat 5 TM scenes prior to mosaicking to remove bad pixels along scene edges. Figure 2.2 provides summary statistics for the resulting number of valid Landsat observations per 30 m pixel each year in the buffered study region. We then calculated four commonly used vegetation indices (enhanced vegetation index (EVI), green chlorophyll vegetation index (GCVI, [*Gitelson et al.*, 2005]), normalized difference vegetation index

(NDVI), and normalized difference water index (NDWI, [*Gao*, 1996]) – for calculations see Text A.1.3) and found the maximum, minimum, and annual range for each. We also calculated the water-adjusted green index (WGI) as the product of maximum GCVI and NDWI, a composite index effective at delineating irrigation status [*Deines et al.*, 2017]. In addition to these 13 Landsat-derived bands, composites also included day of year (DOY) for maximum observed GCVI along with year-to-date precipitation and growing degree days (GDD) for this DOY extracted from GRIDMET [*Abatzoglou*, 2013]. GDD, or heat accumulation units, were calculated using the simple sine method with corn-based upper and lower temperature thresholds of 10° and 30° Celsius, respectively [*Baskerville and Emin*, 1969]. Wheat-based GDD was also calculated, but patterns were similar to corn-based GDD, so we continued with the corn definition of GDD for parsimony. Table 2.1 gives a summary of these Landsat-derived variables, along with additional derived indices and ancillary variables included in the classification algorithm (sections 2.4 and 2.5).

2.3. Defining key crop windows

To separate irrigated and rainfed crops and effectively map irrigated areas based on crop condition at the peak of the growing season, one must have valid satellite observations during the peak greenness period for all major crop seasons in each year. Including pixels without observations during peak periods leads to potential false negatives as lower greenness in those pixels might fail to meet classification thresholds for irrigation. To avoid this, we generated data-informed filters to mask pixels from each annual composite that lacked observations in key periods. Timing of peak greenness in crops is a function of crop species, cultivar, temperature, day length, and management decisions. To derive crop-specific time windows of peak greenness across the HPA, we used crop type maps (Cropland Data Layers (CDL)) produced by the USDA

National Agricultural Statistics Service (NASS) [*USDA-NASS*, 2017] to randomly sample the six main crops grown in the HPA from 2010-2017. We then extracted the DOY and GDD at maximum greenness from the Landsat-derived image composites (section 2.2).

Table 2.1 Summary of variables generated for random forest classification of satellite imagery. All inputs were accessed through the data catalog in Google Earth Engine (GEE) with the exception of USDA SSURGO soil data, which was manually uploaded to GEE. Note that derived products are italicized in the source column.

Variable	Short Name	Time Period	Res (m^2)	Source
Maximum annual EVI	EVI max	Jan. 1 - Oct. 15	30	Landsat
Maximum annual GCVI	GCVI max	Jan. 1 - Oct. 15	30	Landsat
Maximum annual NDVI	NDVI max	Jan. 1 - Oct. 15	30	Landsat
Maximum annual NDWI	NDWI max	Jan. 1 - Oct. 15	30	Landsat
Minimum annual EVI	EVI min	Jan. 1 - Oct. 15	30	Landsat
Minimum annual GCVI	GCVI min	Jan. 1 - Oct. 15	30	Landsat
Minimum annual NDVI	NDVI min	Jan. 1 - Oct. 15	30	Landsat
Minimum annual NDWI	NDWI min	Jan. 1 - Oct. 15	30	Landsat
Annual range in EVI	EVI range	Jan. 1 - Oct. 15	30	Landsat
Annual range in GCVI	GCVI range	Jan. 1 - Oct. 15	30	Landsat
Annual range in NDVI	NDVI range	Jan. 1 - Oct. 15	30	Landsat
Annual range in NDWI	NDWI range	Jan. 1 - Oct. 15	30	Landsat
Water-adj. green index	WGI	Jan. 1 - Oct. 15	30	Landsat
DOY at max GCVI	DOY	Jan. 1 - Oct. 15	30	Landsat
GDD at max GCVI	GDD	Jan. 1 - Oct. 15	30	GRIDMET
Precip. at max GCVI	Ppt-YTD	Jan. 1 - Oct. 15	30	*Landsat/GRIDMET*
Annual precipitation	Ppt-Ann	Dec. 1 - Oct. 15	4000	*Landsat/GRIDMET*
Growing season precip.	Ppt-Grow	May 1 - Oct. 15	4000	GRIDMET
Early season precip.	Ppt-Early	Dec. 1 - Apr. 30	4000	GRIDMET
Annual PDSI	PDSI-Ann	Dec. 1 - Oct. 15	4000	GRIDMET
Growing season PDSI	PDSI-Grow	May 1 - Oct. 15	4000	GRIDMET
Terrain slope	Slope	static	30	USGS NED
Soil Avail. Water Content	AWC	static	30	SSURGO
Precip:Plant Avail. Water	Ppt-PAW	Dec. 1 - Apr. 30	30	*ppt-early / paw-cm*
Soil Normalized GCVI	GCVI-AWC	Jan. 1 - Oct. 15	30	*max GCVI /AWC*
Aridity	Aridity	May 1 - Oct. 15	4000	*Ppt / ET_o*
Aridity-normalized green index	AGI	Jan. 1 - Oct. 15	30	*GI max / aridity*
Soil Normalized AGI	AGI-AWC	May 1 - Oct. 15	30	*AGI / AWC*
Neighborhood Green Index	NGI	Jan. 1 - Oct. 15	30	*GCVI / GCVI_ngb_15p*
Latitude	Latitude	static	30	Generated in GEE
Longitude	Longitude	static	30	Generated in GEE
Year	Year	annual	30	Generated in GEE

To compare phenological timing across crop type and space, we examined the distributions by HPA sub-region (Figure 2.3). Two distinct groups are apparent: 1) a main season group consisting of corn, sorghum, soybeans, and cotton, and 2) an early season group consisting of winter wheat and alfalfa. In the first group, peak greenness is relatively invariant between DOY 183 (July 2) and 273 (Sept. 30) and across sub-regions, while regional distributions in GDD space showed more separation (Figure 2.3). The consistency in DOY of peak greenness for main season crops is likely due to factors such as region-specific cultivars (corn), photoperiod dependence (soy), and limited range (cotton). In the early season group, the main crop of concern is wheat, which has large regional variations in DOY of maximum greenness. This variation was greatly reduced in GDD space (Figure 2.3), likely due to similarity of wheat cultivars and dependence on temperature to reach maturity. Alfalfa tends to have peak greenness in this early season similar to wheat, but goes through several peak greenness cycles due to multiple cuttings through the growing season.

Based on this analysis, we defined two key crop windows to reliably capture irrigated agriculture across the HPA. First, we defined the main season window based on the 10^{th} - 90^{th} percentiles in peak greenness DOY across the four main season crops (DOY 196, July 15 – DOY 245, Sept. 2). This window was assumed to be invariant across years. Second, we defined the wheat-window by the 5^{th} to 95^{th} percentiles for GDD. This was implemented for each year by calculating pixel-based GDD using GRIDMET and extracting the DOY range associated with these thresholds. We assumed that alfalfa is likely captured when both the main season and wheat-based early season filters are met.

We then applied these filters to the annual composites, masking pixels that lacked qualifying observations during both of these key windows. The resulting number of years with

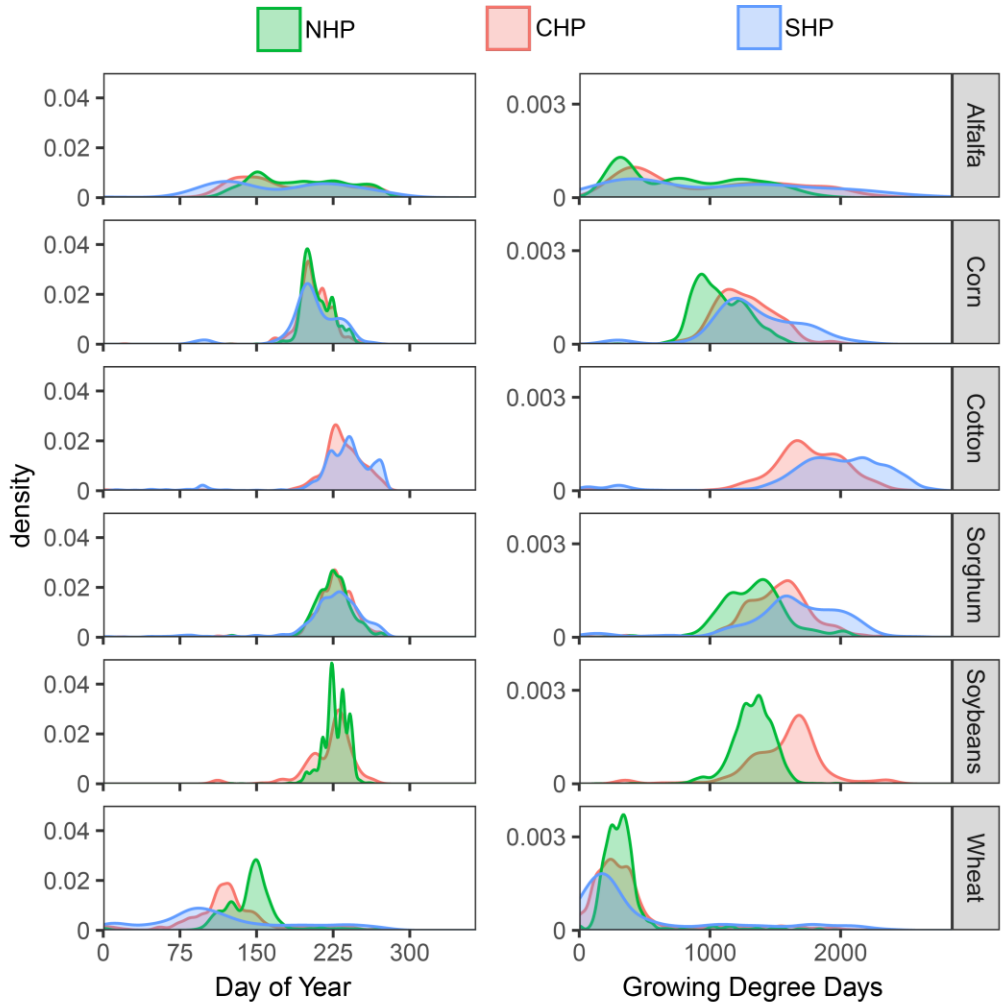

Figure 2.3. Crop-specific timing of annual maximum greenness for major High Plains Aquifer regions. Corresponding day of year (left) and growing degree days (right) on the day of maximum observed greenness based on GCVI derived from Landsat imagery randomly sampled across crop types for 2010-2017. NHP = Northern High Plains; CHP = Central High Plains; SHP = Southern High Plains.

data per pixel for each window, and both windows combined, is shown in Figure 2.4. The combined window defines the number of years with valid data for classification. Pixels lacking observations during the combined crop window for any given year were masked as no data and were subsequently assigned irrigation status during post-classification processing based on surrounding years (section 2.8). We also assumed that for unmasked pixels, at least one cloud-free pixel occurred during the non-growing season to capture annual greenness minimums, thus providing annual range values for the vegetation indices.

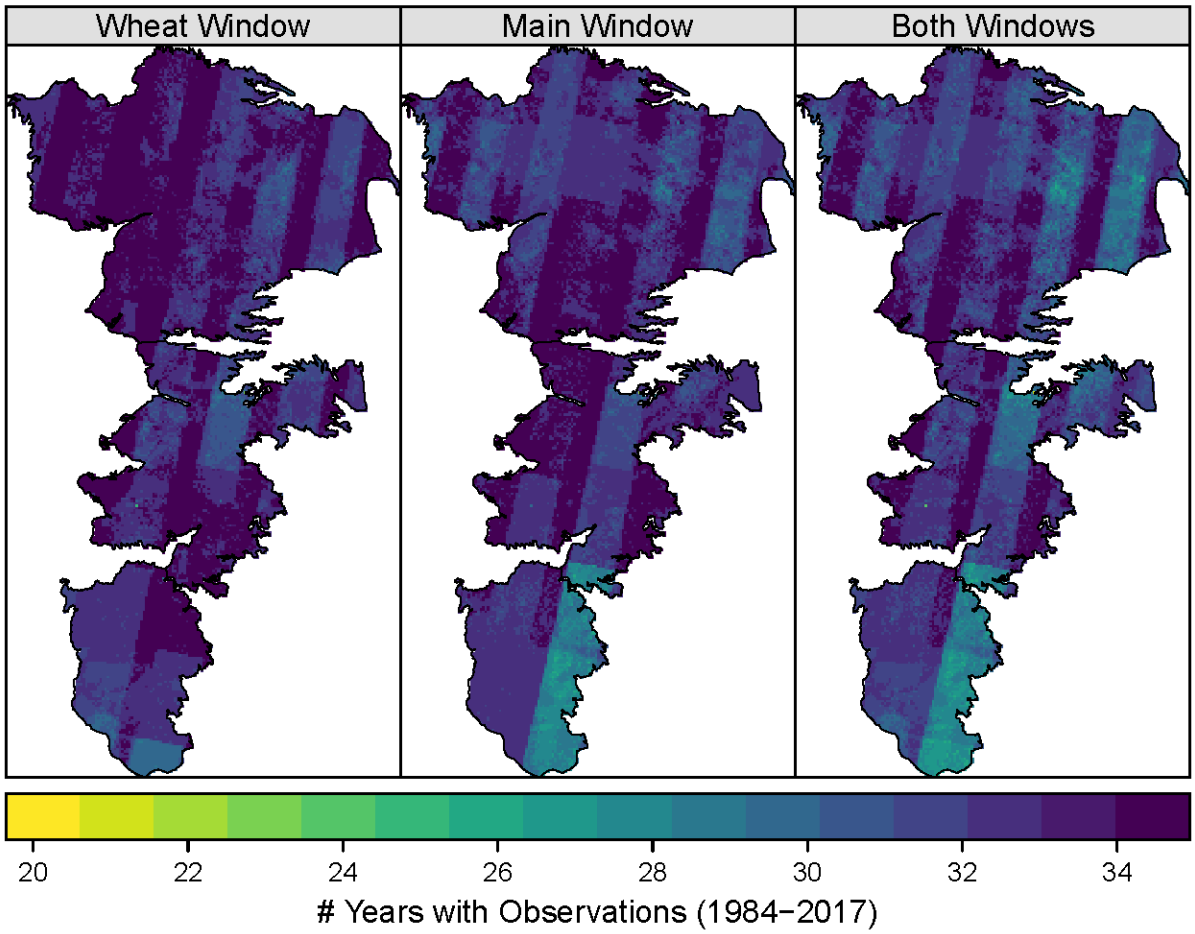

Figure 2.4 Number of years with satellite observations during crucial crop windows. Left: Number of years where Landsat data met the wheat window criteria, dynamically defined based on growing degree days. Center: Number of years where Landsat data met the main season window criteria, defined as imagery between July 15 and Sept. 2. Right: Number of years that met both wheat and main season criteria, thus depicting the number of years Landsat provided sufficient data for classification. Missing pixels for each year were later filled in to generate annual maps with complete spatial coverage across years.

2.4. Neighborhood greenness

Crop greenness as measured by GCVI is a particularly strong delineator of irrigation status [*Ozdogan et al.*, 2010a; *Deines et al.*, 2017], but GCVI both within and outside of irrigated areas varies considerably over the range of crop types and climate conditions in the study area (Figure 2.5). To increase the generalizability of our classifier, we created a normalized greenness index by converting GCVI to an index of neighborhood greenness contrast (Neighborhood

Greenness Index, NGI).

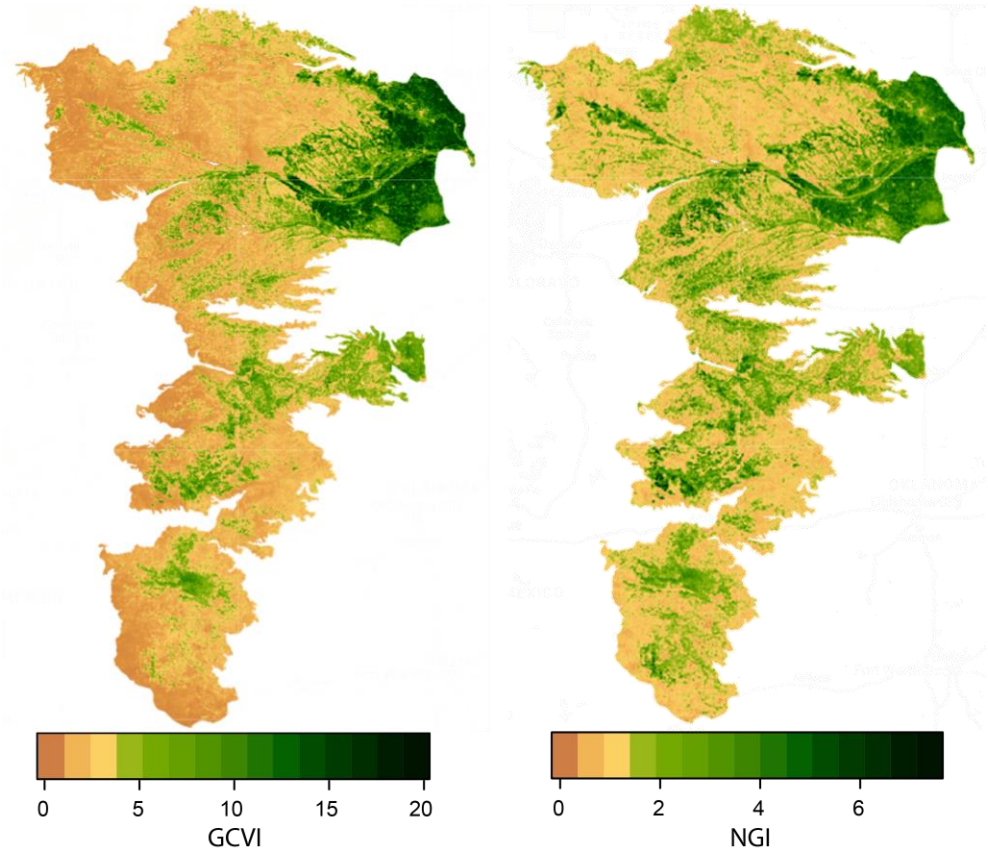

Figure 2.5 Neighborhood Greenness Index (NGI) demonstration, 2015. Left: Green Chlorophyll Vegetation Index (GCVI). Right: NGI. By normalizing greenness based on surrounding areas, the resulting index value for irrigated crops becomes more consistent across the study region.

Here, NGI is defined as the 30 m annual maximum GCVI value divided by the 15th percentile of a 50 km radius circular kernel. The moving kernel was calculated at a resampled 1000 m resolution to incorporate a large neighborhood while staying within GEE's 256-pixel kernel size limitation. Because urban areas and open water have low GCVI values and therefore depress neighborhood greenness percentiles, we masked urban and water pixels prior to running the neighborhood kernel using the most recent land cover product available from the US National Land Cover Database (NLCD) [*Fry et al.*, 2011]. For the purposes of calculating the NGI, masked areas were then filled with the median value of a moving 50 km kernel to maintain a continuous data input layer. Finally, we set a maximum value on the neighborhood greenness

percentile layers based on a dynamic annual threshold set by the median region-wide maximum GCVI value for that year. This ensured that regions with extensive green crop cover were not overly muted due to normalization (such as eastern Nebraska), and the dynamic annual threshold accounted for changes in overall greenness due to annual climate variability as well as a general observed greening trend over the study period. An example of the NGI-15 layer compared to the original GCVI layer for 2015 is shown in Figure 2.5. By normalizing greenness based on surrounding areas, the resulting index value for irrigated crops becomes more consistent across the study region.

2.5. Additional ancillary variables

Classification of irrigated areas is improved when ancillary data are included to characterize climate, soil, and slope, thus providing context for vegetation greenness [*Deines et al.*, 2017]. In addition to the 13 Landsat-based vegetation indices, NGI, and attributes at peak greenness described above, we assembled a suite of environmental co-variables to include in classification (Table 2.1). Broadly, this included seasonal precipitation and aridity time series derived from GRIDMET, slope calculated from a DEM [*USGS*, 2012], and soil water holding capacity extracted from the US SSURGO soil database [*NRCS*, 2016]. We also included pixel centroid latitude and longitude as a proxy for spatial trends, and year to capture any evolving changes in crop management and cultivars. Finally, we included several composite variables combining Landsat-derived GCVI with environmental covariables, including the aridity-normalized green index (AGI) [*Deines et al.*, 2017] and new soil-normalized inputs for GCVI and AGI. Table 2.1 summarizes the full set of 32 input variables. All datasets were obtained through GEE's data catalog and are thus readily integrated into the classification workflow with the exception of SSURGO soil metrics, which we manually uploaded to our GEE assets.

2.6. Ground truth data, accuracy assessment, and variable importance

Due to this study's historical scope over a large region, ground truth data were limited to existing data sources. We made a large effort to discover and incorporate existing data from multiple sources, including state agency irrigation and well databases, previously published maps, and physical ground truth surveys undertaken by past researchers. Notably, the state of Texas in the SHP lack abundant available data for ground truth. Since extensive efforts and positive communication with local agencies and researchers failed to yield any available ground truth data, we supplemented the acquired data with manual, expert-selected training points by visually interpreting imagery using methods similar to those in Deines et al. [2017]. We refer to this collective dataset as "ground truth" data, although in most cases it's derived from secondary sources. Ground truth locations by year and data source are shown in Figure 2.6. Text A 2.1 provides detailed ground truth data sources and processing methods, which are summarized in Table 2.2.

The point ground truth data for this study was split into three groups, with 40% going to classification training and 30% for validation used to assess accuracy during algorithm development, with the remaining 30% reserved as test data to be used as a final assessment of map accuracy. Accuracy for test data is not presented here to preserve its value for assessing final classification accuracy after improvements have been finalized. For an additional independent accuracy assessment, we used two sets of national county statistics for ten years (1997, 2002, 2007, 2012: NASS Agricultural Census (NASS, 2017); 1985, 1990, 1995, 2000, 2005, 2010: USGS water use data (USGS, 2015)) to compare total irrigated area for the 157 counties fully contained in the buffered study region. NASS county data are generated through farmer-reported irrigated areas as part of the semi-decadal census, and USGS estimates of

county-specific irrigated area are produced through state-specific statistical modeling. Agreement between AIM-HPA and county statistics was assessed with r^2 metrics from simple linear regression.

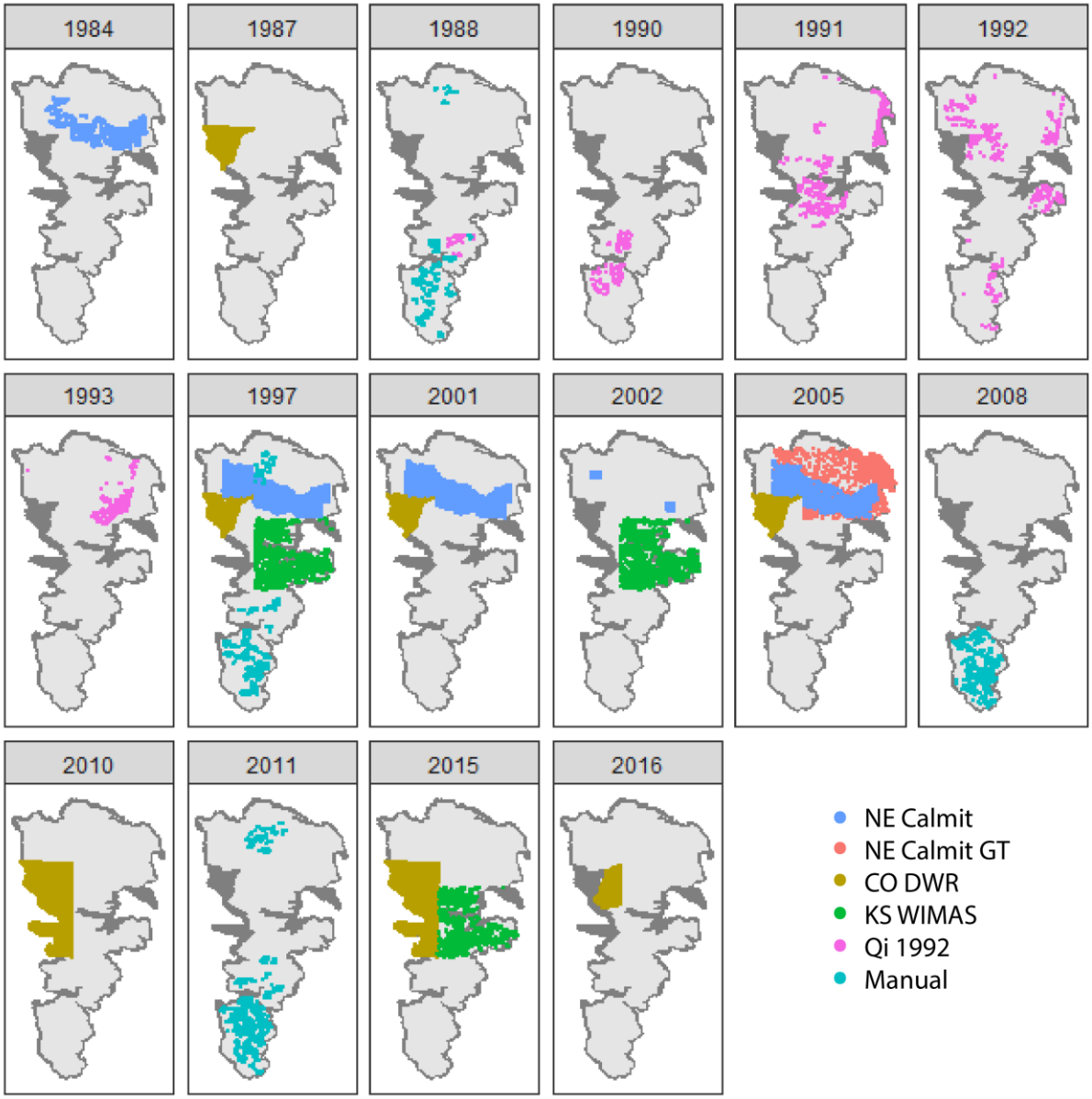

Figure 2.6 Ground truth data location by year and data source. Point ground truth data was acquired from a variety of sources across the study period, including previous Landsat-derived maps (NE Calmit), well data (WIMAS), and ground truth data used in previous studies (NE Calmit GT, Qi et al. 1992). Table 2.2 breaks down number by year and source. The buffered study area is shown in dark gray.

Table 2.2 Ground truth data summary. Detailed descriptions of data sources and processing are in Text A 2.1. Figure 2.6 provides spatial locations by year.

Region	Source	Year	Type	Sampling Method	Irrigated	Rainfed	Other
CO South Platte	CO DWR	1987	Curated Polygons	Stratified random	800	800	800
CO South Platte	CO DWR	1997	Curated Polygons	Stratified random	500	500	500
CO South Platte	CO DWR	2001	Curated Polygons	Stratified random	500	500	500
CO South Platte	CO DWR	2005	Curated Polygons	Stratified random	500	500	500
CO South Platte, Republican, and Arkansas	CO DWR	2010	Curated Polygons	Stratified random	1500	1500	1500
CO South Platte, Republican, and Arkansas	CO DWR	2015	Curated Polygons	Stratified random	1500	1500	1500
CO Republican	CO DWR	2016	Curated Polygons	Stratified random	500	500	500
NE Platte - CALMIT	Dappen & Merchant 2004	1982 (1984)	Landsat MSS 3	Manual, 5% of polygons	675	329	108
NE Platte - CALMIT	Dappen & Tooze 2001	1997	Landsat TM	Stratified random	1000	1000	1000
NE Platte - CALMIT	Dappen and Merchant 2003	2001	Landsat ETM+	Stratified random	1000	1000	1000
NE 2 Counties - CALMIT	Dappen 2003	2002	Landsat ETM+	Stratified random	300	300	300
NE Platte - CALMIT	Dappen et al. 2007	2005	Landsat TM	Stratified random	1000	1000	1000
NE State - CALMIT	USDA FSA	2005	Point locations	Used full dataset	1673	1045	NA
KS HPA	WIMAS	1997	well data	Manual, 5% of active wells	933	517	304
KS HPA	WIMAS	2002	well data	Manual, 5% of active wells	908	680	432
KS HPA	WIMAS	2015	well data	Manual, 5% of active wells	854	489	228
HPA: TX, OK	Qi et al. 2002; FSA	1988	Hand-drawn polygons	One point per polygon	69	97	NA
HPA: NM, OK, and TX	Qi et al. 2002; FSA	1990	Hand-drawn polygons	One point per polygon	322	333	NA
HPA: CO, KS, NE, OK, SD	Qi et al. 2002; FSA	1991	Hand-drawn polygons	One point per polygon	1134	1836	NA
HPA: all states but OK	Qi et al. 2002; FSA	1992	Hand-drawn polygons	One point per polygon	901	1924	NA
HPA: KS, NE, WY	Qi et al. 2002; FSA	1993	Hand-drawn polygons	One point per polygon	578	1020	NA
Manual: SHP, NE Sandhills	Landsat	1988	Visual interpretation	Targeted gaps	185	221	NA
Manual: SHP, NE Sandhills	Landsat	1997	Visual interpretation	Targeted gaps	198	266	NA
Manual: SHP	Landsat	2008	Visual interpretation	Targeted gaps	295	296	NA
Manual: SHP, NE Sandhills	Landsat	2011	Visual interpretation	Targeted gaps	295	296	NA

To understand the relative contribution of input variables (Table 2.1) to classification accuracy, we ran permutation tests and GINI index metrics in R [*R Core Team*, 2014], since GEE does not yet offer variable importance measures (GEE accessed January 2018 - July 2018). The GINI Index quantifies the decrease in node impurities resulting from classification tree splits on each variable. The permutation test measures variable importance by iteratively randomizing each variable's values and quantifying the resulting decrease in accuracy for predictions made on the out-of-bag samples. Variables with greater contributions to overall accuracy would therefore see larger decreases in accuracy following permutation. Conversely, randomly rearranging values for unimportant variables would have little effect on prediction accuracy. We used the randomForest package [*Liaw and Wiener*, 2002] to develop a proxy random forest classification with the same variables, training data, and parameters used in GEE. To account for randomization in node optimization and bagging in random forest algorithms, we created 20 classifiers from different random seeds and averaged importance scores for each variable.

2.7. Classification and post-classification cleaning

We used all ground truth data reserved for training in a random forest classification with 300 trees within GEE. Because the training data included water, urban, and natural vegetation classes, we did not use a land use mask prior to classification to preserve flexibility to detect changing land use compared to the wide intervals between existing land cover products like NLCD, particularly given demonstrated ongoing conversion of uncultivated land in the study area [*Lark et al.*, 2015]. The generated output is considered the raw classification output.

Multi-year land cover classifications can benefit greatly from the ability to leverage a time series of repeat classifications to improve the accuracy of the final dataset [*Cardille and Fortin*, 2016; *Wulder et al.*, 2018]. Here, we apply a series of three cleaning steps based on

external land cover datasets, year-specific irrigated area properties, and information gained from the time series of AIM-HPA.

First, we masked confounding land covers and reduced speckle in each raw annual classified map based on the assumption that contiguous plots were managed uniformly. We used the NLCD to mask wetland and forest land covers that can be difficult to distinguish from irrigated lands. NLCD maps are available for 1992, 2001, 2006, and 2011. We used the most recent NLCD product for each year to mask woody and herbaceous wetlands; deciduous, evergreen, and mixed forests; and open water classes; the 1992 product was used for earlier years. Urban and grassland areas were accounted for in our training and validation datasets under the "not irrigated" category, allowing them to be dynamically classified over the full time period. To reduce speckle in classified maps, we counted connected pixels of each output class (irrigated and non-irrigated) based on connectivity with cardinal-direction neighbors. For all patches smaller than 25 connected pixels at 30 m resolution (2.25 ha, compared to a typical field size of ~60 ha), we updated the classification output based on a circular kernel-based majority filter with a radius of 120 m. The overall effect of this filter is to convert isolated pixels classified as irrigated to non-irrigated, and to fill in small gaps in fields otherwise classified as irrigated. Restricting the majority filter update to small patches identified by the connected pixel count has the advantage of maintaining the location of field edges from the original classification. This first cleaning is referred to as a "despeckler".

Second, we leveraged the multi-year characteristics of the dataset to impose logical restrictions on irrigable area. Rainfed crops occasionally can overlap irrigated fields in the spectral signatures used in classification (i.e. Chapter 1, Figure 2). This is most likely to occur in years with high precipitation and can lead to spurious classifications over the 34 year study

period. To address this, we restricted irrigable area to pixels classified as irrigated at least twice during twelve specific years during the study period, since single-year irrigation is unlikely due to the cost of infrastructure required to irrigate. These twelve years included eight in the bottom quartile of precipitation to sample dry years across the study period (1989, 1991, 1994, 2000, 2002, 2006, 2011, and 2012). Because no years near the beginning or end of the study fell in the bottom quartile, we also included the two first years and two last years of the study (1984-1985 and 2016-2017) to avoid excluding fields that may have been deactivated shortly after the start of the study or recently activated. This second cleaning is referred to as "multiyear cleaning."

2.8. Addressing data gaps

Masked pixels that lacked valid observations during the defined crop windows were then filled in based on data present in surrounding years, with an accompanying flag to specify filled pixels. Pixel-filling used the following rules, applied sequentially:

1. Pixels that were never irrigated in available years were assigned as "non-irrigated
2. Pixels that were always irrigated in available years were assigned as "irrigated"
3. Pixels with an irrigated classification in adjacent years were assigned as "irrigated"
4. Pixels with a non-irrigated classification in adjacent years were assigned as "non-irrigated"
5. Pixels with an irrigated classification in the adjacent 2 years were assigned as "irrigated"
6. Pixels with a non-irrigated classification in the adjacent 2 years were assigned as "non-irrigated"
7. All remaining gaps were assigned as "non-irrigated."

In this way, we were able to generate spatially-complete annual irrigation maps for 1984-2017, while maintaining information on which pixels were filled each year for downstream users. We are working to implement a more sophisticated Bayesian updating methodology [*Cardille and Fortin*, 2016] based on the 34-year time series for each pixel, incorporating pixel-wise probabilities and producing associated uncertainties with the final map product.

3. Results and Discussion

3.1 Accuracy assessment

Our random forest workflow and subsequent gap-filling approach produced 34 annual irrigation maps from 1984-2017 across the entire High Plains Aquifer (AIM-HPA). AIM-HPA had an overall accuracy of 91.4% based on a diverse validation dataset that integrated point datasets from numerous state, federal, and imagery-derived sources across multiple years (Table 2.2). Post-classification cleaning steps reduced irrigation commission errors (false positives) from 10.5% to 6.9% while minimally increasing irrigation omission errors (false negatives) from 14.2% to 15.7%. Table 2.3 provides a breakdown of overall accuracy at each classification and post-processing step, including overall omission (producer's) and commission (consumer's) errors for each map class.

This point accuracy assessment represented reasonably wide coverage across years and regions given historic data limitations and can be considered a robust estimate of dataset accuracy. Table 2.4 and Table 2.5 break down accuracy by region and year, but these should be interpreted with caution due to uneven sampling across units typical for historical accuracy assessment [*Wulder et al.*, 2018]. Errors were fairly consistent across HPA regions (Table 2.4). The NHP did have higher commission errors than the SHP and CHP, likely due to its relatively humid climate especially in the eastern regions that support robust rainfed crops in wet years. The approach also had good performance across the 34 year period, with overall accuracies from 84.6% - 97.7% (Table 2.5). Five years (1987, 1991 – 1993, 2010) had irrigation omission errors higher than 20%, indicating there is some scope for further classification improvement.

Table 2.3. Overall map accuracy by cleaning step. Accuracy percentages are provided along with the qualifying number of points and total sample size for each category. Note that point accuracy assessment specific to gap-filled regions has yet to be implemented; county-level assessment serves as validation for gap-filled estimates (Figure 2.7).

Map Version	Overall Accuracy	Class	Omission Errors	Commission Errors
Raw Output	90.5% (13433/14845)	Irrigated	14.2% (826/5802)	10.5% (586/5562)
		Non-Irrigated	6.5% (586/9043)	8.9% (826/9283)
De-speckled	91.0% (13507/14845)	Irrigated	14.8% (856/5802)	8.9% (482/5428)
		Non-Irrigated	5.3% (482/9043)	9.1% (856/9417)
Multiyear clean	91.4% (13572/14845)	Irrigated	15.7% (910/5802)	6.9% (363/5255)
		Non-Irrigated	4.0% (363/9043)	9.5% (910/9590)
Gap-filled	91.4% (13572/14845)	Irrigated	15.7% (910/5802)	6.9% (363/5255)
		Non-Irrigated	4.0% (363/9043)	9.5% (910/9590)

Table 2.4. Point accuracy by region. NHP = Northern High Plains, CHP = Central High Plains, SHP = Southern High Plains.

		Omission Errors		Commission Errors			
Region	Overall Accuracy	Irrigated	Non-Irrigated	Irrigated	Non-Irrigated	Total Points	Irrigated Points
NHP	90.7%	15.6%	5.5%	10.1%	8.8%	7981	1177
CHP	91.1%	15.5%	2.4%	2.9%	13.3%	2402	2951
SHP	92.8%	12.8%	1.9%	2.2%	11.1%	1055	515

Comparisons with county-level national statistics for 157 counties fully contained within the buffered study area provide an additional assessment of map accuracy, providing even coverage across the full study area for a wide range of years and climate conditions (Figure 2.7). Overall county-level agreement across all years was good, with AIM-HPA slightly underestimating irrigated area compared to county statistics (simple linear regression, $r^2 = 0.81$, m=0.83). Agreement by year was strongest NHP counties (r^2 from 0.80 to 0.98) followed by those in the CHP (r^2 from 0.78 to 0.96). SHP county comparisons indicated that AIM-HPA tended to underestimate irrigated area in the south (r^2 from 0.47 to 0.94), performing particularly poorly during the early Landsat record. This is likely due to a combination of deficit-irrigated

cotton's low greenness, limited training data, and low Landsat data availability in this region (Figure 2.4), with gaps largely concentrated in early years. For the post-1999 period, agreement with county statistic was similar across high (e.g., 2007) and low (e.g., 2002) precipitation years. In general, AIM-HPA showed better agreement with NASS county statistics compared with the USGS county estimates (NASS overall: $r^2 = 0.90$; USGS: $r^2 = 0.793$). It's difficult to distinguish if this is a sampling effect due to availability in different years, or an artefact of the underlying county datasets themselves. All three datasets are imperfect. USGS irrigated area estimates are produced from state-specific statistical models and thus are not consistent in methodology across the study region. The NASS semi-decadal agricultural census is a robust effort by the USDA, but relies on self-reported irrigated area from producers who may have incentives for both under and over reporting due to water rights mechanisms. AIM-HPA provides a completely independent estimate of spatially-explicit irrigated area with misclassifications typical of remote sensing products. With the exception of the SHP in early years, which we hope to improve with more stringent image filtering algorithms, the county-level agreement among the three datasets is generally quite good.

Table 2.5 Point accuracy metrics by year. Note that data source and spatial distribution varies across years due to historical data availability (Figure 2.6, Table 2.2, Text A 2.1). Mean annual precipitation for the buffered study area derived from GRIDMET.

			Omission Errors		Commission Errors			
Year	Precip. (mm)	Overall Accuracy	Irrigated	Non-Irrigated	Irrigated	Non-Irrigated	Total Points	Irrigated Points
1984	531	97.7	1.1	4.1	2.7	1.7	306	184
1987	576	90.3	24.8	1.5	3.5	12.0	721	254
1988	455	92.5	13.7	1.6	1.8	11.8	253	124
1990	496	91.4	13.3	3.7	3.9	12.6	221	113
1991	524	86.1	25.9	6.6	12.6	14.5	883	336
1992	563	84.6	30.7	7.7	18.1	14.3	820	274
1993	618	87.0	21.4	8.2	15.4	11.8	460	168
1997	571	93.3	10.8	4.0	6.6	6.7	2369	923
2001	493	91.8	14.4	5.2	11.2	6.8	1308	424
2002	384	93.9	10.7	2.9	4.7	6.9	1140	456
2005	521	90.7	13.6	5.7	7.3	10.7	2156	978
2008	556	96.8	5.8	0.7	0.8	5.2	285	138
2010	583	91.9	23.4	1.2	3.3	9.7	1357	423
2011	442	90.7	14.2	5.1	6.3	11.6	332	155
2015	688	92.5	15.8	2.0	3.5	9.6	1786	710
2016	547	93.1	15.5	2.9	7.0	6.9	448	142

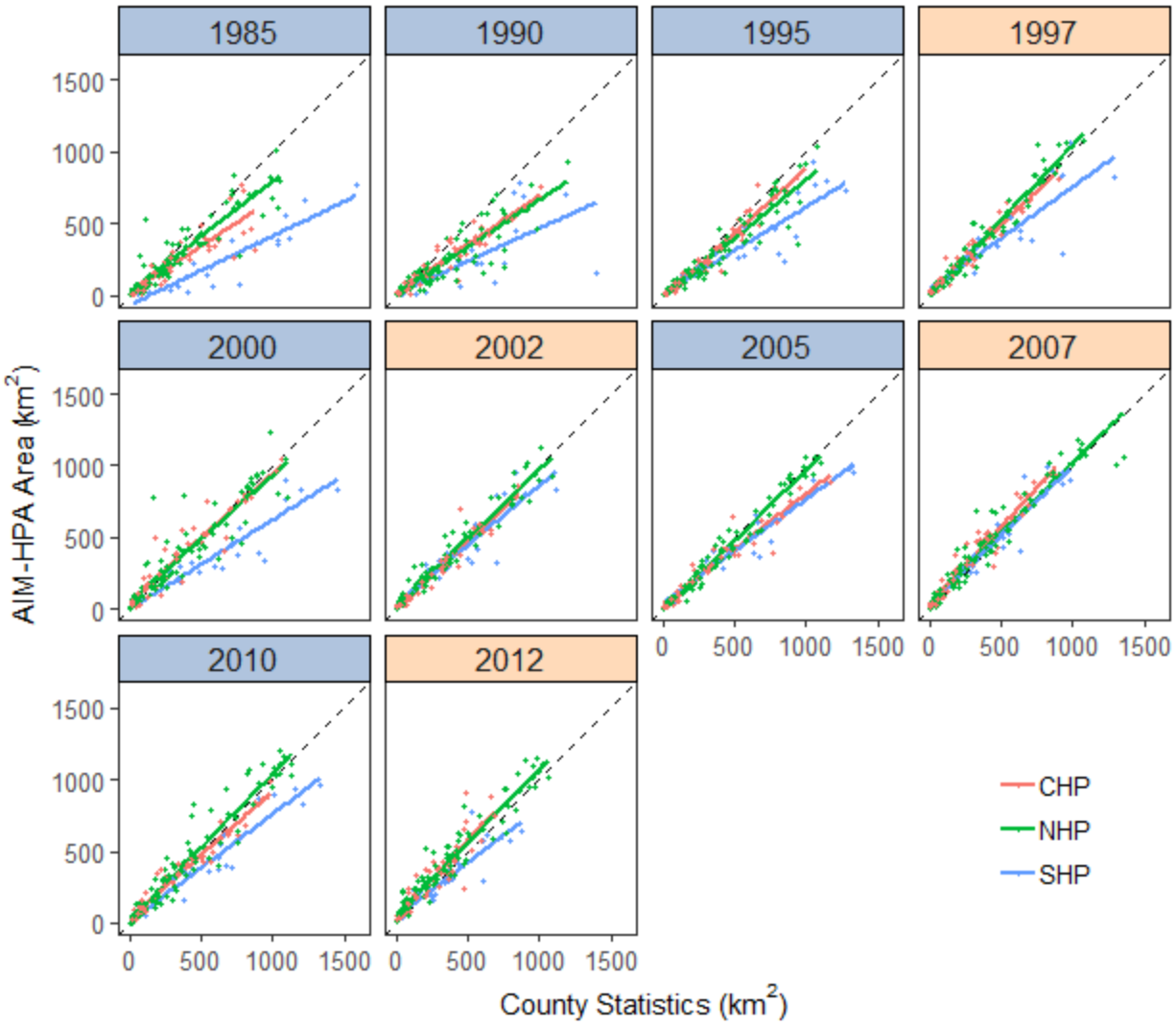

Figure 2.7. County-level accuracy assessment by region. Irrigated area from county-level statistics compared with irrigated area from AIM classified maps against a 1:1 line (dashed). Solid lines indicate trendlines by major aquifer region based on simple linear regressions. County statistics for blue panels are from the USGS [USGS, 2015], while statistics for tan panels are from the USDA National Agricultural Statistics Service [NASS, 2017] and include 157 counties fully contained within the buffered study region.

3.2. Variable importance

Variable importance metrics were separately in R to provide insight into key variables that contribute to classification accuracy. The two chosen accuracy metrics provide complementary assessment of variable performance (Figure 2.8). It is not surprising that over such a large region, important variables identified by the permutation test are those that help

group pixels by global attributes, essentially "localizing" subsequent nodes within the random forest trees. Latitude and longitude both scored in the top three, likely allowing the random forest to partition classification trees by major climate and crop type gradients. Similarly, neighborhood slope and annual minimum NDWI likely help separate non-crop or non-irrigable pixels, GDD and DOY at peak greenness likely provided a proxy for crop type based on phenology, and year allowed thresholds to vary through time. Without these large grouping variables, differences between irrigated and non-irrigated classes across such a large region would be less distinct.

GINI scores (Figure 2.9), on the other hand, are dominated by variables that most effectively distinguish irrigated and non-irrigated classes within the "localized" portion of the decision trees. The novel NGI developed in this study scored highest, demonstrating the utility of normalizing pixel greenness to the surrounding neighborhood when working across large regions. The recently developed WGI (and AGI to a lesser extent) [*Deines et al.*, 2017] scored highest after NGI, again emphasizing the improvements gained by incorporating moisture indicators with typical greenness indices. This is further emphasized by NDWI's ranking at number three, since NDWI was developed to monitor leaf water content [*Gao*, 1996]. GCVI, a primary component in both WGI and AGI, also scored high on the GINI index for annual range, maximum, and maximum normalized by soil available water content. This corroborates conclusions that GCVI is more effective at distinguishing irrigation status than more conventional vegetation indices such as NDVI and EVI [*Ozdogan and Gutman*, 2008; *Ozdogan et al.*, 2010a; *Deines et al.*, 2017].

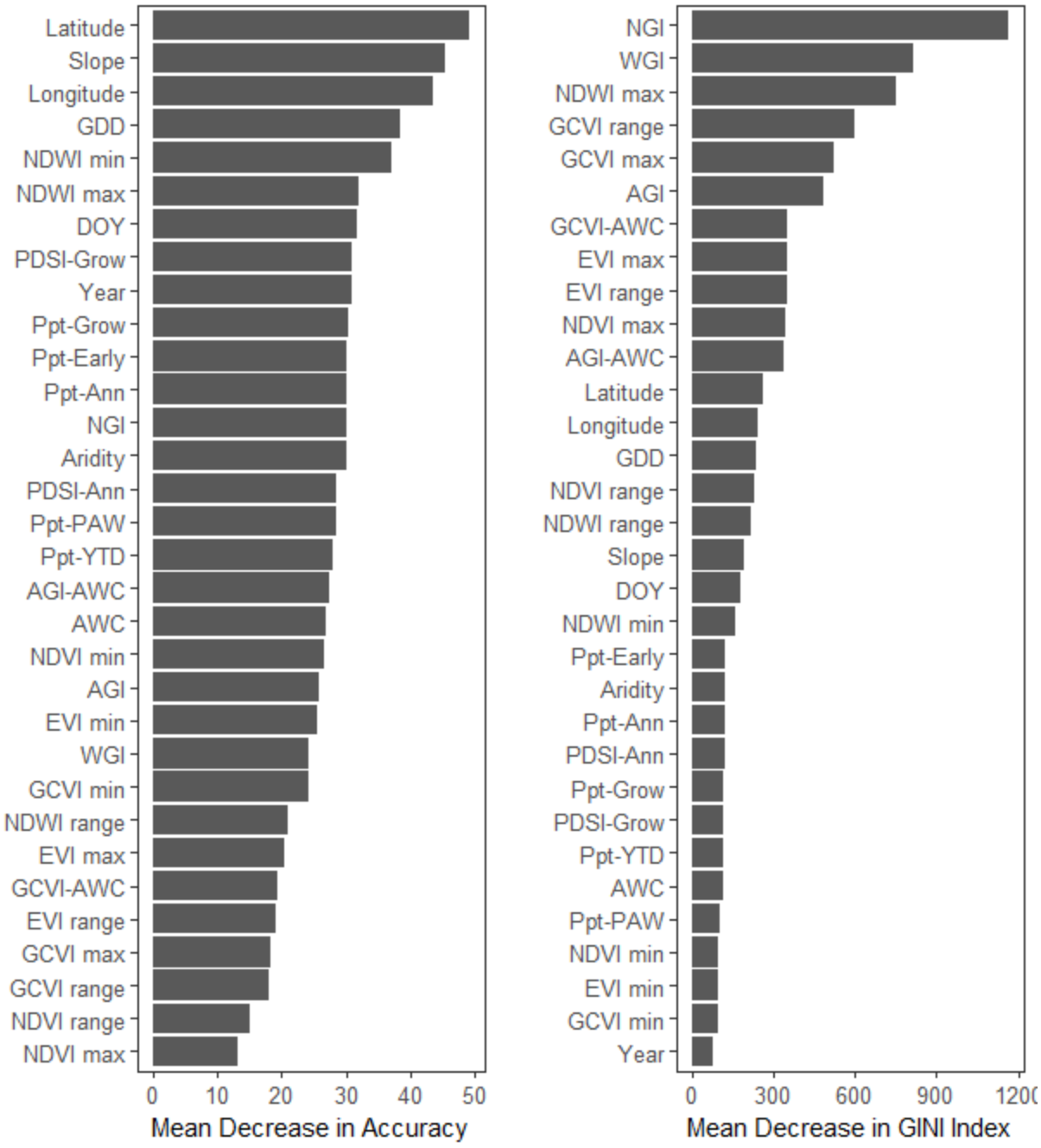

Figure 2.8. Variable importance metrics for the random forest classification. Input variables for the classification algorithm are identified by their shortname; Table 2.1 summarizes variable definitions and source data. *Left*: Permutation test rankings, which ranks variable importance by assessing decreases in classification accuracy resulting from randomization of that variable. *Right*: GINI Index rankings, which measures the mean decrease in node impurity resulting from splits on each variable.

3.3. Irrigation trends

The AIM-HPA product details dynamic irrigation patterns that can inform water

management in this stressed aquifer system. Overall, we found that 23.2% of the HPA was irrigated during the study period, but only 2.2% of this was irrigated all 34 years due to characteristic rotations of crop type and irrigation status, periodic field fallowing, and irrigated area added or deactivated over the study period. AIM-HPA is able to track these changes annually with 30 m resolution. Figure 2.9a depicts the location and frequency of irrigated area across the study region. As expected, heavily irrigated areas closely align with regions of extensive aquifer depletion, particularly in the CHP and SHP (Figure 2.1a).

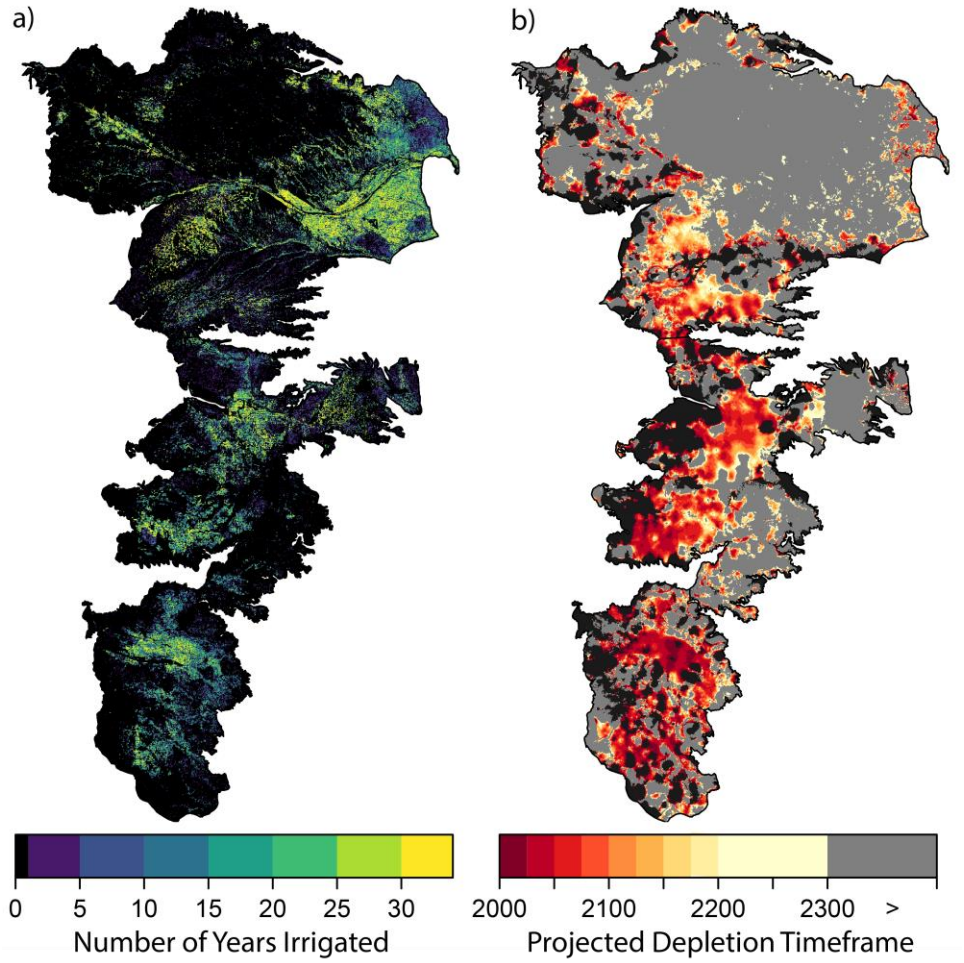

Figure 2.9 Irrigation frequency and aquifer depletion in the High Plains Aquifer. (a) Per pixel irrigation frequency between 1984-2017 based on the number of years irrigated in the classified map product. (b) Projected depletion timeline from Haacker et al. [2016], estimated assuming a continued linear trend in water level decline. The aquifer is considered depleted when saturated thickness falls below the 9 m needed for most high volume well operations. Black regions depict areas with <9 m saturated thickness at aquifer development.

Summarizing total irrigated area by sub-region revealed that the NHP has seen substantial expansion in irrigated area (Figure 2.10a). Comparison against county statistics indicates AIM-HPA may underestimate irrigated area prior to 1997, likely due to decreased image frequency to capture true peak greenness, and overestimate in the latter half of the study period (Figure 2.7, Figure 2.10a). Soybeans, which can exceed irrigated corn greenness even for rainfed conditions [*Deines et al.*, 2017], have been expanding in eastern Nebraska, possibly increasing commission errors. However, the relative rate of irrigated area increase between 1997 and 2007 is similar for AIM-HPA and NASS (Figure 2.10a). Other studies have reported increases in groundwater wells in this region exceeding 1200 per year between 2002-2005 [*Pervez and Brown*, 2010]. Given that the final NASS data point in 2012 represents a severe drought year that necessitated heavier irrigation on reduced area to meet crop water needs [*Deines et al.*, 2017], the temporal resolution of NASS inhibits inference on recent irrigated area trends. The rate of expansion found in AIM-HPA seems plausible due to wider trends for increased irrigation in humid areas to mitigate precipitation variability, relatively favorable aquifer conditions in the eastern NHP (Figure 2.1a and Figure 2.9b), biofuel expansion (17 new biofuel plants between 2002-2008 in Nebraska, [*Pervez and Brown*, 2010]), and high corn prices during this period.

AIM-HPA trends in the CHP and SHP showed a lower rate of increase in irrigated area during the 1990s. Comparisons with county datasets strongly suggest this increasing trend may be largely due to the higher omission rates in the first half of the study period. Since 2000, the AIM-HPA trend in irrigated area is relatively stable for these regions with some interannual fluctuations likely driven by precipitation and pricing [Chapter 1, this volume]. For the CHP in particular, this seems likely since Kansas is known to be fully appropriated with very few new water rights granted [*Whittemore et al.*, 2016].

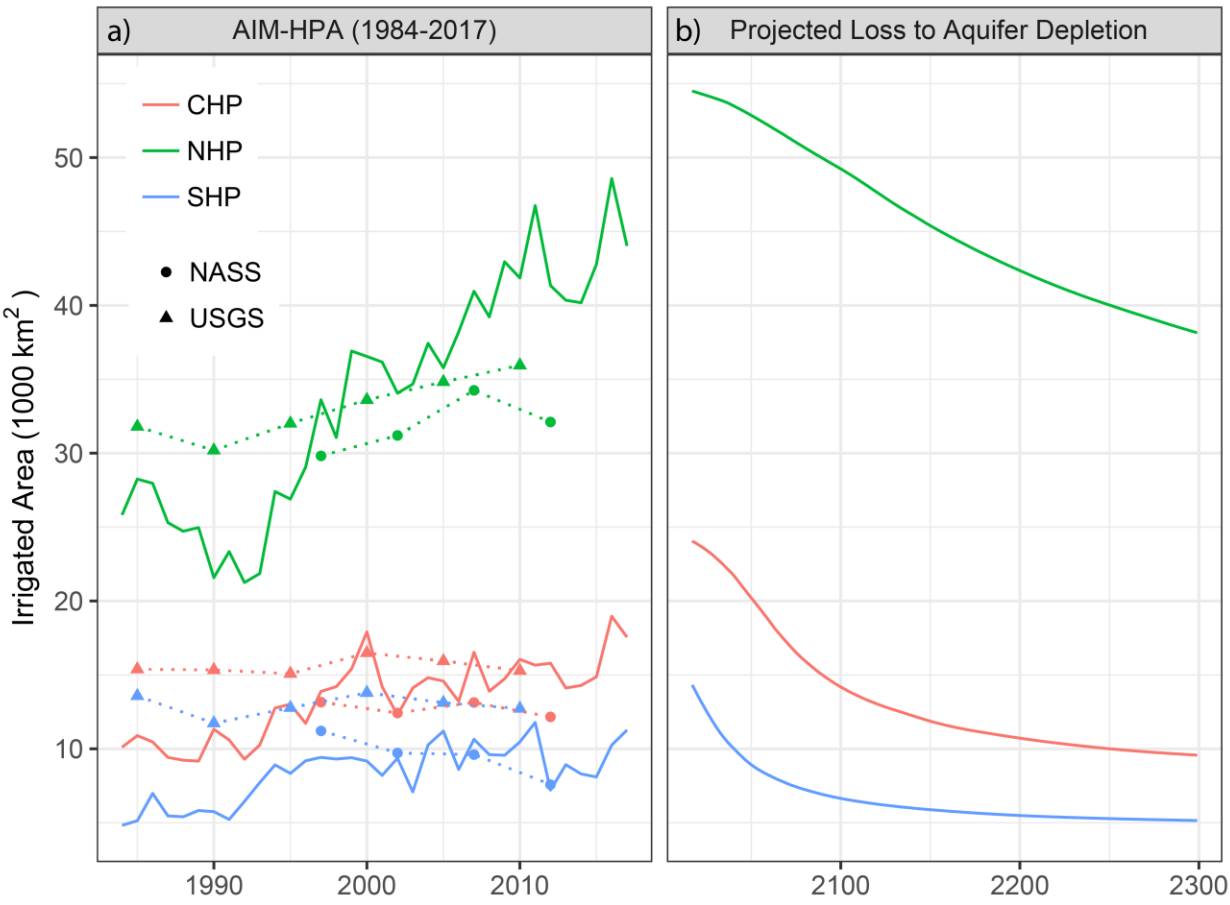

Figure 2.10. Irrigated area over time by region. (a) Total irrigated area by major aquifer subregions based on the Landsat-derived annual irrigation maps (AIM-HPA). Totals from county statistics by data source are plotted for years available. (b) Projected cumulative area lost over time through 2300 relative to combined irrigated area from 2015-2017 based in AIM-HPA. Irrigated area was considered lost when the underlying aquifer was depleted beyond viable use based on linear extrapolation of past decline (Figure 2.9b) based on [*Haacker et al.*, 2016].

We also examined fine-scale spatial trends in irrigated area over time by aggregating the 30 m AIM-HPA dataset to a uniform 8 km^2 grid and running a linear model on irrigated area per grid cell (Figure 2.11). Overall, the majority of the HPA displayed positive trends in irrigated area during the study period. Notable areas of strong increase were found in the eastern portion of Nebraska, as well as hot spots in the CHP. As discussed above, eastern Nebraska has seen a large expansion in irrigated area since the 1990s. Relatively high rates of natural recharge as well as considerable surface water resources have largely accommodated irrigation in this region

[*Scanlon et al.*, 2012; *Breña-Naranjo et al.*, 2014], but conditions should continue to be monitored going forward due to the rapid observed increases.

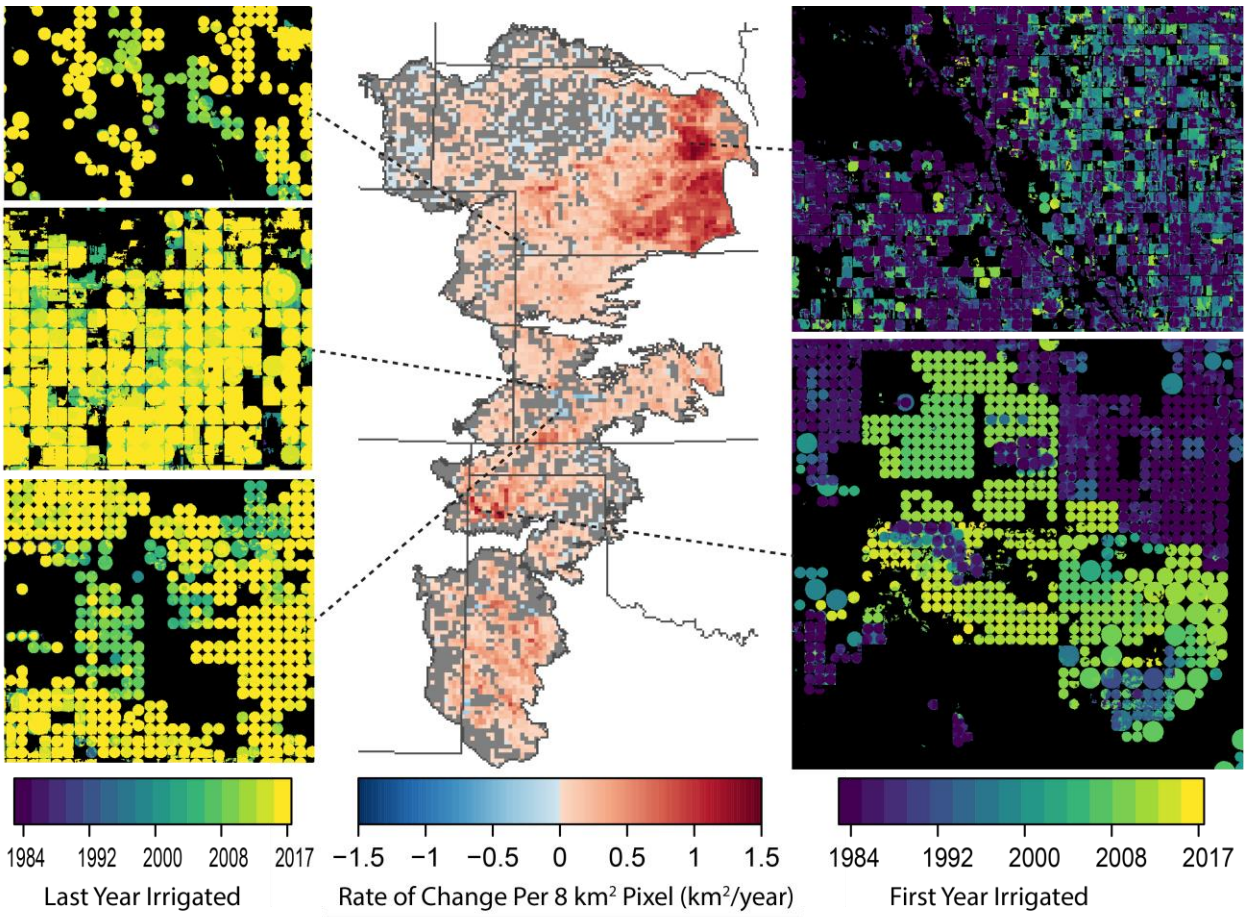

Figure 2.11 Spatially explicit trends in irrigated area. *Center*: Rate of change over time from linear regression based on a uniform 8 km^2 aggregated grid. Cells with nonsignificant trends (alpha >= 0.05) are in gray. *Left*: Zoomed insets of regions with decreasing irrigated area over time. *Right*: Zoomed insets of regions with increasing irrigation over time.

Increases in annual irrigated area in largely appropriated regions like Kansas may be due to improvements in irrigation technology, allowing producers to actively irrigate larger portions of their permitted area with their existing water allocation. For example, conversion to low-pressure irrigation systems from ~1995-2005 allowed producers in western Kansas to irrigate more area [*Pfeiffer and Lin*, 2014a; *Wang et al.*, 2015]. Dramatic increases in irrigated area were observed in the Texas portion of the CHP (Figure 2.11). The highlighted region of recent

expansion overlies an already highly stressed portion of the aquifer, suggesting this may be indicative of groundwater mining. Although Groundwater Conservation Districts are gaining momentum, Texas largely still follows an absolute ownership doctrine for water rights, granting land owners free reasonable use of any groundwater on their property and limiting regulation [*Smidt et al.*, 2016].

Regions of decreased irrigated area are not as prevalent but do occur in isolated areas across the HPA. Reasons for reductions vary across the aquifer. In Wyoming, for instance, programs such as the Agricultural Water Enhancement Program helped retire several thousand acres of irrigated cropland by purchasing water rights from producers willing to transition to dryland farming or pasture [*WSEO*, 2018]. The southwest corner of Nebraska (top left inset, Figure 2.11) has seen reduced area due to increasing regulation on water use, both by local management districts and as a result of lawsuits among Kansas, Nebraska, and Colorado settled in 2002 and 2015 by the Supreme Court. Irrigation technology can also drive changes in irrigated area, as indicated by the termination of irrigation in field corners due to the increased adoption of center pivot systems and discontinuation of the use of high-pressure end guns during the study period (Figure 2.11, center left). Other irrigated areas are retired due to aquifer depletion [*Marsalis et al.*, 2018] or conservation programs such as the Conservation Reserve Program, which pays producers to enroll their land and return to grassland for a specified period generally 10 years per contract [*Hellerstein*, 2017].

3.4. The future of the High Plains Aquifer

It's widely acknowledged that in many parts of the HPA, current rates of groundwater use cannot be sustained [*Scanlon et al.*, 2012; *Butler et al.*, 2016; *Haacker et al.*, 2016; *Whittemore et al.*, 2016]. Management goals reflect this in several jurisdictions, treating

groundwater as a nonrenewable resource that should be used wisely under "managed depletion" strategies [*Waskom et al.,* 2006; *Peck,* 2007; *Smidt et al.,* 2016]. Haacker et al. [2016] projected depletion timeframes across the HPA based on a linear extrapolation from historic 1993-2012 trends (Figure 2.9b). Starting from a baseline reference of 2015-2017 combined irrigated area from AIM-HPA, we translated these depletion timeframes into annual irrigated area lost by HPA region through 2300 (Figure 2.10b). By this estimate, 54%, 41%, and 10% of currently irrigated area would no longer be viable by 2100 in the SHP, CHP, and NHP, respectively, together accounting for 24% (22,821 km^2) of 2015-2017 irrigated area.

Given this stark outlook, strategies have emerged across the aquifer to slow depletion, making the region a hotbed for innovation in irrigation technology, crop selection, management interventions, and stakeholder-regulator partnerships [e.g, Chapters 3 & 4, this volume]. Robust evaluation and comparison of programs within the context of historical irrigation trends can help identify successful strategies. For example, recent work has highlighted how improved irrigation efficiency through technological advances alone doesn't translate to water conservation. In a case study in the Kansas portion of the HPA, groundwater extraction actually increased as producers used the saved water over greater areas by reducing fallowing frequency and switched to more water intensive and profitable crops [*Pfeiffer and Lin*, 2014a]. This study was enabled by Kansas's extensive, publicly available well database detailing annual water use [*KDA DWR*, 2017]. Well metering is increasingly common and mandatory across HPA jurisdictions. Combining AIM-HPA with water use data can provide unique insights into farmer adaptation strategies [e.g., Chapter 3, this volume].

Producers in the SHP are transitioning away from water demanding crops like corn to those with less water demand, including sorghum and cotton, as well as testing new crops like

winter canola [*Marsalis et al.*, 2018; *West et al.*, 2018]. Field management to enhance yields without additional water is also spreading, including cover crops and reduced tillage to increase soil water storage [*Marsalis et al.*, 2018]. The AIM-HPA product can help assess these strategies by providing a consistent dataset across multiple state and local jurisdictions with field-scale precision. This allows annual assessments at spatial aggregations relevant to each intervention, providing provide a timely assessment of trends not influenced by sampling effects inherent in the 5-year intervals of existing data or aggregations to larger, rigid spatial units such as counties.

4. Conclusions

Effective management of irrigation resources first requires that we know when and where irrigation occurs. Cloud-computing tools like Google Earth Engine and open access to the full Landsat archive enabled us to produce annual irrigation maps in the High Plains Aquifer from 1984-2017, thus quantifying the history of irrigation extent across three decades. We developed a novel indicator, the Neighborhood Greenness Index (NGI), that performed well, contributing to a single random forest classification algorithm that achieved 91.4% overall accuracy across a wide region and time period. Neighborhood normalization of greenness may thus be a promising avenue for future continental to global-scale efforts. The approach presented here developed robust methods to handle sparse data in the early Landsat record, and it can be readily applied for ongoing monitoring through the computational infrastructure developed in Google Earth Engine.

The resulting 34-year map dataset provides a rich and detailed accounting of irrigated lands across one of the key agricultural regions of the United States. Our broad-scale analyses indicated continued expansion across the full aquifer, including both areas of high water stress in the Central and Southern High Plains in addition to the less-stressed Northern High Plains. Assuming continued water use trends and management, however, we estimated that existing

irrigated area could decline by 24% over the 21st century due to groundwater depletion. Finding solutions to extend aquifer life that simultaneously sustain agricultural economies and groundwater resources is a pressing challenge. The classification workflow described here translates available satellite imagery into information that can be used to quantify, evaluate, model, and manage agricultural water use.

Acknowledgements

This chapter was co-authored by Anthony Kendall, Jeremy Rapp, and David Hyndman. We thank Jesslyn Brown, Patti Dappen, Bethany Kurz, Shahriar Pervez, Santhosh Seelan, and Sharon Qi for dusting off old external hard drives and CDs to recover and share ground truth datasets. This work benefitted enormously from their past work and generosity. Funding was provided by NSF grant WSC 1039180, USDA NIFA grant 2015-68007-23133 and USDA-NIFA/NSF INFEWS program grant 2018-67003-27406. Jillian Deines was partially supported by NASA Headquarters under the NASA Earth and Space Science Fellowship Program grant 14-EARTH14F-198. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of NSF, USDA NIFA, or NASA.

APPENDIX

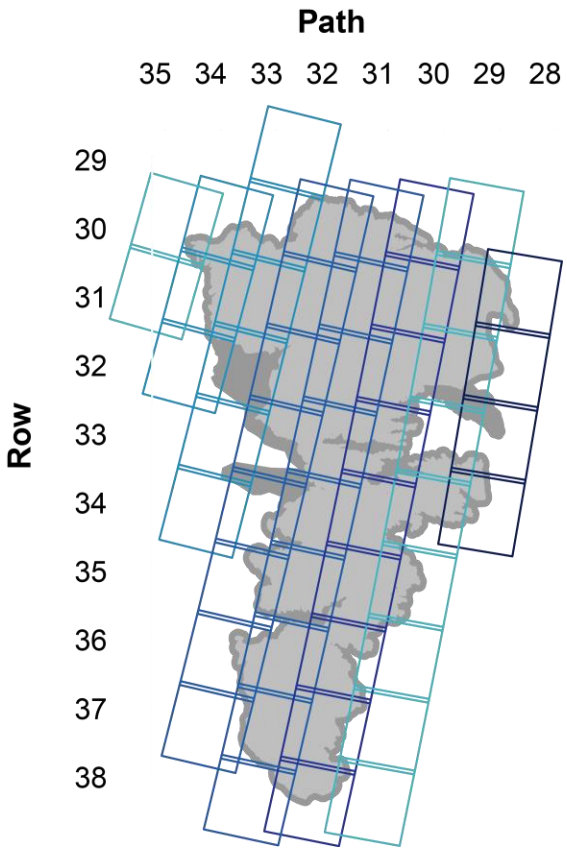

Figure A 2.1 Landsat scenes covering the buffered study area. All available Landsat Collection 1 Surface Reflectance imagery for Landsat 4 TM, 5 TM, 7 ETM+, and 8 OLI from 50 scenes between 1984-2017 were used.

Text A 2.1 Expanded ground truth data description

In Colorado, the Division of Water Resources has mapped irrigated area in the HPA region intermittently since 1956. The state is divided up into seven hydrologic divisions, two of which (the South Platte and Arkansas) overlie or are directly adjacent to the HPA. Some of the Arkansas division is outside the High Plains Aquifer but is climatically similar to the HPA, so we included these as training points as well. The polygon dataset is produced using available well permits to delineate allowable irrigated areas and available data from Landsat, NASS, CDL, and NAIP (Chris Brown, CO-DWR, Nov. 2017, personal communication). For each year polygons were available (Table 2.2), we used the polygons and a crop mask (CDL or NLCD prior to CDL availability) to generate a continuous, 30 m categorical raster for irrigated, rainfed,

and non-crop pixels within the division boundary. We then used a stratified random approach to create 500 points in each class per available region, restricting locations to unmasked pixels in the annual composite. Due to limited data in earlier years, we made 800 points per class in 1987. We then split these points into training, validation, and test points.

In Nebraska, the Center for Advanced Land Management Information Technologies (CALMIT) at the University of Nebraska – Lincoln has generated a series of remotely-sensed irrigation maps covering 28,800 square miles of the Central Platte River Basin in support of hydrological modeling efforts by the ongoing Platte River Cooperative Hydrology Study [*Dappen and Tooze*, 2001; *Dappen*, 2003; *Dappen and Merchant*, 2003, 2004; *Dappen et al.*, 2007]. Vector-format GIS maps are available for 1982, 1997, 2001, and 2005. Additional maps are available in 2002 for Scotts Bluff and Kearney Counties only. For 1997, 2001, 2002, and 2005, we used the same approach described for CO DWR data to randomly sample irrigated, rainfed crop, and non-crop classes. Sample size was reduced for 2002, as it was restricted to two counties (Table 2.2). To leverage 1982 data to generate "ground truth" at the start of the study period, we randomly sampled 5% of 1982 irrigation polygons and overlaid them on 1984 Landsat composites. Points were then manually placed in fields interpreted to be actively irrigated in 1984, as well as in adjacent rainfed crop or non-crop locations. Points were then split into training, validation, and test datasets. In addition, true ground truth locations used to generate these products were obtained for 2005, providing 2,718 points for the full extent of Nebraska within the study region. This ground truth was generated from USDA Farm Service Agency certified reporting records and in-season field excursions.

In Kansas, the Division of Water Resources maintains the WIMAS well database, which contains annual records of irrigation wells in Kansas, including locations and annual pumping

volume. We used these records to guide manual placement of irrigated and non-irrigated points in the Kansas HPA in 1997, 2002, and 2015 (Table 2.2). After extracting well records located over the study area, we split the dataset into wells with non-zero pumping volume for the specified year (indicating active irrigation), and wells with zero pumping volume (a possible indication of inactive irrigation, depending on proximity of other active wells). We then randomly sampled 5% of each data set and overlaid them on the year-specified Landsat composite of maximum GCVI, including a layer for the day of year on which the maximum value was observed. Locations for irrigated crops, rainfed crops, and non-crop locations were then manually interpreted based on well locations and greenness patterns.

We also obtained ground truth data previously used to produce a nominal 1992 HPA-wide irrigation map [*Qi et al.*, 2002]. Qi et al.[2002] identified between 1 and 10 one-square-mile sections in all 154 counties across the HPA that were ≥ 50% agricultural land by area based on the 1992 NLCD. They then requested crop type and irrigation status for each Common Land Unit within the identified one-square-mile sections from the USDA FSA, and the FSA returned photocopied aerial images with fields labeled by crop type and irrigation status as self-reported by farmers to the FSA. Because their study used the best available Landsat TM imagery from 1988-1993 to completely cover the HPA with peak-season imagery, crop type and irrigation status was requested for years matching the location-specific Landsat year in this time period. Qi et al. then hand-digitized the dataset, resulting in 10,992 polygons. During the gap in time between when this data was generated for the 2002 study and our data request, attribute data on image date was no longer available. We therefore digitized the map in Figure 11 in Qi et al. [2002] denoting years for Landsat scenes selected across the study area and performed a spatial join to ascribe year for each polygon. Polygons overlying image borders were omitted, since we

could not confidently ascribe a year. As this dataset was never reviewed or published, it should be considered provisional. Because these polygons did not have a 1:1 relationship with crop management units based on visual inspection with underlying Landsat imagery, we manually placed one point per polygon in the most likely location based on the crop type, irrigation status, and visual cues in the imagery. To reduce bias introduced by manual placement, we used polygon centroids in all cases where the centroid seemed plausible. For the remaining cases, including centroids which fell outside of irregularly shaped polygons or centroids which fell in clearly visible borders between differently managed fields in the polygon, we moved the point to a pseudo-random location in the described field. Data that overlaid regions lacking imagery in our critical crop windows for the specified year were omitted, along with locations that were clearly incorrect (e.g., a fallowed field that was marked as irrigated). In some cases, it was difficult to judge if the description was correct; these cases were marked as "uncertain" and removed from the training data set, but retained in validation and test data, to avoid biasing the classifier. The final number of points per year after cleaning is giving in Table 2.2.

After initial classification based on these acquired ground truth data, qualitative evaluation of resulting maps and comparisons with county data (not shown) indicated underestimation of irrigation in the SHP and overestimation in the central Nebraska Sandhills region. Given minimal representation in the ground truth data set in these areas (Figure 2.6), we created additional ground truth points by manually locating points based on visual interpretation of Landsat imagery in GEE for four years across the study period chosen to fill existing gaps (1988, 1997, 2008, 2011). Irrigated area from the preliminary map product was overlain on annual Landsat GCVI composites to help target manual point placement in currently underrepresented cases, including irrigated fields in the SHP (particularly cotton) and wetland

swales in the NE Sandhills not fully captured by the NLCD land use mask. We created between 185-295 additional irrigated ground truth points for each year, paired with similar numbers of points placed in surrounding rainfed agricultural fields and non-crop uses (combined in one non-irrigated class, Table 2.2).

CHAPTER 3:
QUANTIFYING WATER USE AND FARMER ADAPTATION STRATEGIES IN RESPONSE TO NOVEL STAKEHOLDER-DRIVEN GROUNDWATER MANAGEMENT IN THE US HIGH PLAINS AQUIFER

Abstract

Irrigation greatly enhances agricultural yields and stabilizes farmer incomes, but overexploitation has depleted groundwater resources around the globe. Strategies to address this sustainability challenge differ widely. Socio-ecological systems research suggests management of common pool resources like groundwater would benefit from localized approaches that combine self-organization with active monitoring. In 2012, the U.S. state of Kansas established a Local Enhanced Management Area (LEMA) program, empowering farmers to work with local and state officials to develop five-year, enforceable groundwater conservation programs. Here, we assessed the efficacy of the first LEMA implemented from 2013 to 2017 using a causal impact methodology that is new to agrohydrology. Compared to control scenarios, we found that the LEMA reduced water use by 33% through 2016, slowing the decline in groundwater levels. We then combined satellite-derived irrigated areas and crop type maps with well records to partition water savings among three conservation strategies, revealing that farmers focused on adaptive cropping choices and increased efficiency while largely maintaining irrigated area. The results of this analysis demonstrate that conservation programs that are irrigator-driven with regulatory oversight can provide a path toward sustainability in stressed aquifers worldwide.

1. Introduction

Irrigated agriculture helps meet global food demand by enhancing agricultural yields and buffering crop productivity and farmer income from climate variability and change [*Lobell et al.*, 2009; *Troy et al.*, 2015; *Smidt et al.*, 2016; *Rufin et al.*, 2018]. Groundwater contributes about half of the world's irrigation water and is often the primary source in arid to semiarid regions [*Kustu et al.*, 2010; *Siebert et al.*, 2010; *Aeschbach-Hertig and Gleeson*, 2012], but overexploitation has depleted aquifers around the globe [*Gleeson et al.*, 2012b; *Wada and Heinrich*, 2013; *Rodell et al.*, 2018]. In the United States, the High Plains Aquifer (HPA) supports more than $20 billion in annual economic activity [*Ashworth*, 2006]. However, water-level declines threaten the continued viability of irrigated agriculture over much of the aquifer [*Scanlon et al.*, 2012; *Haacker et al.*, 2016; *Cotterman et al.*, 2018].

Policy and management institutions developed to address this sustainability challenge differ widely across the HPA and beyond. Aquifer depletion can be costly, since the value of irrigation water should increase over time considering expected future higher yielding varieties and irrigation's ability to mitigate droughts, which are likely to become more frequent and severe with climate change [*Zipper et al.*, 2016; *Foster et al.*, 2017; *Quintana Ashwell et al.*, 2018]. At the same time, improved management could boost crop water productivity around the world [*Brauman et al.*, 2013], indicating producers might obtain similar yields using less water and thus slow the rate of aquifer depletion. Top-down approaches to management are typically met with resistance by farmers who are understandably concerned with near-term profit [*Wang et al.*, 2015]. Since groundwater can be considered a common pool resource [*Hardin*, 1968; *Ostrom et al.*, 1994], approaches that operate on local scales, allow self-organization, and include active monitoring and enforcement are more likely to achieve sustainability [*Ostrom*, 2009b].

A management framework with these characteristics has emerged in Kansas, where HPA water levels are rapidly declining and pumping reductions appear to be the only viable option for reducing decline rates [*Butler et al.*, 2016; *Whittemore et al.*, 2016]. Legislation in 2012 allowed stakeholder groups to establish Local Enhanced Management Areas (LEMAs) and work with local (groundwater management districts or GMDs) and state officials to develop enforceable and monitored water use reduction programs that operate over five year cycles [*K.S.A. 82a-1041*, 2012]. The pioneering LEMA began in 2013, following a vote by irrigators within a 256 km^2 highly stressed region in northwestern Kansas referred to as Sheridan 6 (hereafter SD-6, Figure 3.1) [*KDA*, 2013]. The group sought to reduce the total groundwater pumping over the five year (2013-2017) LEMA period by 20% relative to 2002-2012 levels [*NW KS GMD 4*, 2016]. Allocations were reduced to a total of 55 inches (139.7 cm) per irrigated acre over the five year period, with acreages varying by existing water rights; up to one year of unused water (11 inches) can be carried over to subsequent LEMA cycles. In 2017, stakeholders voted to renew the SD-6 LEMA for 2018-2022. In the spring of 2018, a second LEMA was approved for most of the surrounding district (GMD4), and additional LEMAs are being discussed in parts of three other Kansas GMDs.

Understanding the effectiveness and impact of the SD-6 LEMA is vital as the LEMA program expands, and opportunities for stakeholder-driven management spread across Kansas, the United States (e.g., California's recent Sustainable Groundwater Management Act [*Babbitt et al.*, 2018]), and the world [e.g, *Tringali et al.*, 2017]. Here, we analyzed the effects of this first LEMA on groundwater pumping, water levels, and irrigated crop dynamics to address two main questions: 1) How did the observed pumping volumes following LEMA establishment differ from the pumping that would have occurred in its absence, controlling for climate and evolving

management trends?; and 2) What adaptation strategies did producers use to meet required pumping reductions?

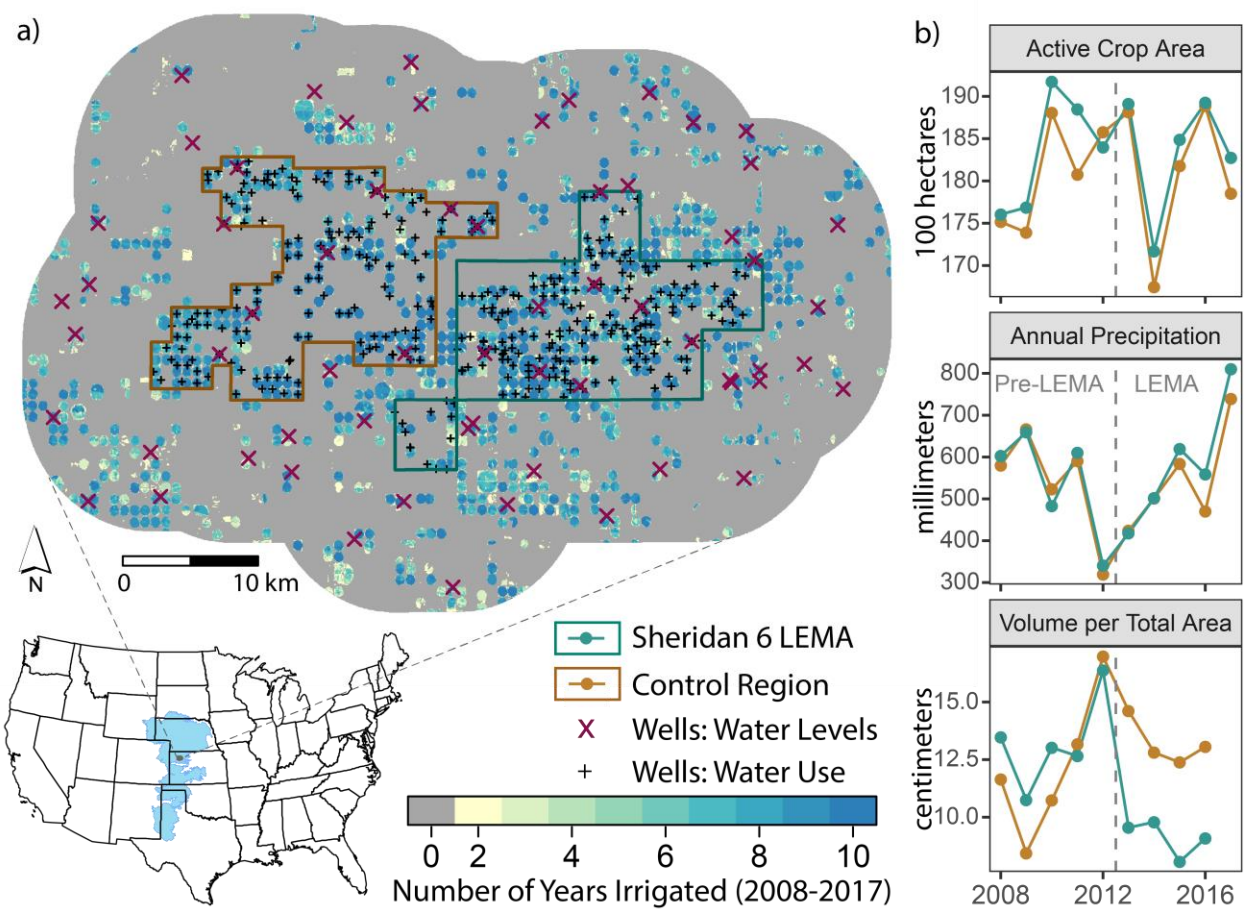

Figure 3.1. Study area map and regional characteristics. a) Locations of the Sheridan 6 Local Enhanced Management Area (LEMA), paired control region, and a combined 10 km buffer within the High Plains Aquifer. Wells used to generate annual aquifer level maps are depicted with X's, and wells with irrigation extraction volumes are shown with +'s. Irrigated areas are colored by irrigation frequency between 2008-2017 from Deines et al. [2017]. b) Variables used to select control region boundaries from 2008-2017. A vertical dashed line divides the pre-LEMA and LEMA periods. Top: annual active crop area by region [*USDA-NASS*, 2017]. Center: Annual precipitation by region [*Abatzoglou*, 2013]. Bottom: Annual total pumping volume divided by total area.

To account for climate fluctuations and wider trends in management and/or technology, we employed two complementary controls in the absence of a randomized experimental control. First, we established a paired *control region* that matched characteristics of the SD-6 region. Second, we generated a statistical control to estimate a business-as-usual scenario in the absence of the LEMA program (hereafter *BAU scenario*). To calculate the BAU scenario, we used causal

impact analysis, which is an emerging Bayesian structural time-series method [*Brodersen et al.*, 2015] new to agrohydrology. We then combined detailed well records, our remotely sensed annual irrigation maps dataset [*Deines et al.*, 2017], and annual national crop maps to quantify how pumping reductions were achieved to understand land use impacts and farmer adaptation strategies.

2. Methods

2.1. Control region design and data processing

We established the control region by manually demarcating an area analogous to SD-6 during the five years prior to the LEMA (2008-2012, Figure 3.1). We targeted adjacent areas (at least 1.5 km away to reduce direct well effects [*Fileccia*, 2016]) with similar well density and irrigation frequency based on annual irrigation maps (AIM) [*Deines et al.*, 2017]. Working in Google Earth Engine [*Gorelick et al.*, 2017], we iteratively adjusted control region boundaries until the 2008-2012 mean control region statistics were within 10% of SD-6 for total area (ultimately a 0.12% difference), crop area based on the USDA Cropland Data Layers (CDL; 1.38%) [*USDA-NASS*, 2017], annual precipitation derived from GRIDMET 4 km gridded daily climate data (0.01%) [*Abatzoglou*, 2013], and total pumped volume over total area based on WIMAS well data (7.1%, described below).

The state of Kansas maintains high quality, publicly available groundwater level and well-specific annual pumping data. The WIZARD well database contains water depth measurements that have been curated by the Kansas Geological Survey since 1996 [*KGS*, 2018]. To translate these irregularly located wells into geostatistically robust groundwater levels, we used R software [*R Core Team*, 2014] to extract 1996 – 2017 well measurements within a 10 km buffer around the study regions, filtered for observations recorded between December 10 and

February 28 (~50 annual observations from 64 wells, Figure 3.1a). These winter measurements provided consistent timing for water tables to partially recover following the active pumping season, which typically ends by mid-September. We kriged these measurements with the gstat R package [*Pebesma*, 2004; *Gräler and Pebesma*, 2016] to produce annual water table elevation maps at 250 m resolution. Due to high longitudinal anisotropy in groundwater levels, we used universal kriging with an easterly trend. With this approach, the longitudinal trend was first modeled using a first-order polynomial. Residuals from the linear trend model were then kriged and combined with the trend surface to produce the estimated water table surface. Annual Gaussian variograms for model residuals were automatically fit with gstat, with mean variogram parameters of 0.83, 32.4, and 12,535 m for nugget, partial sill, and range respectively. The mean water level for each region was calculated and attributed to the year of the preceding active growing season.

Annual groundwater pumping for each region was calculated based on the WIMAS water use database maintained by the Division of Water Resources (DWR) of the Kansas Department of Agriculture [*KDA DWR*, 2017], which documents annual pumping for 203 and 162 wells within SD-6 and the control region during 1996-2016, respectively (Figure 3.1a). The WIMAS data also reports well-specific crop mixes and irrigated areas. The 2017 WIMAS data were still undergoing QA/QC protocols and were thus unavailable for this analysis.

To track land use changes in crop type and irrigation status, we used a novel fusion of satellite-derived annual crop type (CDL) and annual irrigation maps (AIM) at 30 m resolution from 2008 to 2017 to capture the 5-year periods before and after LEMA establishment. To our knowledge, no other data set for this region is able to track crop-specific irrigated area at this spatial and temporal resolution. Because the previously-published AIM dataset ends in 2016, we

used the method and classifier described in Deines et al. [2017] to extend the irrigation map product to 2017. To minimize misclassification in the satellite-derived maps, the AIM product was filtered by removing irrigated pixels outside allowable place-of-use tracts maintained by the Kansas DWR [*KS DWR*, 2017]. Both map datasets were accessed and processed through Google Earth Engine.

2.2. Business As Usual (BAU) scenario and causal impact analysis on pumping and water levels

We generated the BAU scenario using causal impact analysis implemented via the CausalImpact R package [Brodersen et al., 2015]. This approach originated in marketing and website analytics to provide robust analysis of time series data to assess market interventions when appropriate control groups are unavailable. It has since been applied widely, including to assess aviation fuel tax impact on aircraft emissions [González and Hosoda, 2016] and population-level vaccine effects [Bruhn et al., 2017], but has not to our knowledge been applied in the agriculture or hydrology literature. Causal impact analysis implements a Bayesian structural time series model, which uses supplied covariates to construct a BAU estimate with uncertainty bounds to enable causal attribution in the absence of a randomized experiment [Brodersen et al., 2015]. This state-space model approach is preferred over often-used Ordinary Least Squares regression or difference-in-differences methods because it addresses autocorrelations in time series data, incorporates changes in external conditions that can affect the response variable, flexibly allows regression coefficients to vary over time while avoiding overfitting, and provides inference about the temporal evolution of the response rather than simply comparing before and after conditions [*Bertrand et al.*, 2002; *Brodersen et al.*, 2015].

To evaluate how the LEMA affected groundwater use and water levels in SD-6, we used

the CausalImpact package with default priors to separately model the BAU scenario for two response variables: (1) total pumping volume from 1996 to 2016 based on WIMAS well data, and (2) mean water levels from 1996 to 2017 based on the kriged annual water levels. For covariates, we used the following annual time series: 1) GRIDMET-derived annual precipitation, growing season (May through August) precipitation, and pre-season through harvest precipitation (January through August); 2) seasonal aridity, defined as accumulated potential evapotranspiration / precipitation for May through August; 3) corn prices as a proxy for all commodity prices [NASS, 2017]; and 4) year. These are suitable covariates since they correlate with the response variables but are not themselves affected by the LEMA program [Brodersen et al., 2015]. The model then uses the response variable's observed time series behavior, the relationships among the response and covariate time series variables from 1996 through 2012, and the covariate time series during the LEMA period to construct the posterior distribution of the response variable's BAU behavior. If observed values fall outside of this estimate and the 95% confidence interval at alpha = 0.05, it can be concluded that the LEMA program had a significant impact on the response variable. We then used the same approach to generate a BAU scenario in the control region. If no differences between observed responses and BAU scenarios are found in the control region, then any significant changes in SD-6 are considered due to the LEMA program and not external regional-scale drivers such as altered management, technology adoption, and cropping trends unrelated to LEMA establishment.

2.3. Evaluating relative contributions of water saving strategies

Farmers can decrease water use in three primary ways: 1) reduce irrigated area, 2) reduce irrigation volume per area (hereafter, irrigation depth) applied to existing crops, and/or 3) switch to crops with lower irrigation demand [Hendricks and Peterson, 2012]. To partition water

savings among these three conservation strategies, we used the fused AIM-CDL annual maps of crop type and irrigation status along with WIMAS data that specifies well-specific pumping volume, crop mix, and area irrigated (section 2.1).

First, we assessed changes in total irrigated area within SD-6 and the control regions to compare the five-year LEMA period (2013-2017) against the preceding five years (2008-2012, hereafter the pre-LEMA period), using both AIM and WIMAS as complementary lines of evidence. WIMAS is a well-curated data source, but irrigated area is self-reported. It is unclear how producers may vary in reporting year-specific active irrigated area compared to allowable irrigable area, or if reports include or omit area that received some irrigation but was then abandoned due to drought-induced water constraints. On the other hand, AIM is satellite-derived and is thus an independent data source, but it may not detect subtle differences between some rainfed and irrigated areas [Deines et al., 2017]. To overcome these potential issues, we chose to report statistics from both datasets for comparison.

Second, we evaluated changes in irrigation depths for SD-6 and the control region by calculating annual depth applied by crop type from WIMAS, focusing on the four dominant crops (corn, soybeans, sorghum, and winter wheat). Because WIMAS does not explicitly break down irrigated area among crops for reported mixed-crop fields, we restricted this analysis to single-cropped fields from 1996-2016. We again applied causal impact analysis with the same covariates described in section 2.2 for each of four response variables (crop-specific irrigation depths) to estimate changes due to the LEMA, thus controlling for external climate conditions. We also compared the pre-LEMA and LEMA periods to describe overall changes in irrigation depths. Third, we used AIM-CDL to evaluate changes in crop-specific irrigated area by region between both 5-year periods.

Finally, we calculated the contribution of each of these three conservation strategies to overall water use reductions in the SD-6 LEMA based on differences between the pre-LEMA and LEMA periods. Water savings from reductions in total irrigated area (change in pumping, ΔP_{Area}) were estimated for both WIMAS- and AIM-specified areas based on the following equation:

$$\Delta P_{Area} = 5 \times \bar{d}_{LEMA}(\bar{A}_0 - \bar{A}_{LEMA}) \quad (1)$$

where $\bar{d}_{LEMA}$ is mean annual irrigation depth in the 5-year LEMA period based on annual pumping volume and annual irrigated area for 2013-2016, $\bar{A}_0$ is pre-LEMA mean irrigated area, and $\bar{A}_{LEMA}$ is LEMA mean irrigated area. We used average applied irrigation depth during the LEMA period ($\bar{d}_{LEMA}$) in Eq. (1) to avoid double counting savings from change in both total irrigated area and change in irrigations depths. Mean annual water savings are then multiplied by 5 to estimate ΔP_1 for the full LEMA period. We then averaged estimates for WIMAS and AIM to obtain a final estimate.

Water savings due to reduced irrigation depths on existing crops (ΔP_{Depth}) were calculated based on annual crop-specific irrigated area obtained from fused AIM-CDL maps for 2013-2017 and irrigation depth reductions found via causal impact analysis:

$$\Delta P_{Depth} = \sum_i^{years} \sum_j^{crop\ types} a_{LEMA,ij} \times \epsilon_{LEMA,j} \quad (2)$$

where $a_{LEMA,ij}$ is the year-specific irrigated area for each crop type and $\epsilon_{LEMA,j}$ is the crop-specific LEMA effect on pumping depths based on causal impact models (see ϵ estimates in Table 1). Sorghum was not included here because there was no significant reduction in sorghum irrigation depths (Table 1), thus Eq. (2) applied to corn, soy, and wheat. Because results showed that wheat area increased in SD-6 during the LEMA period, we used pre-LEMA wheat area in Eq. (2) to avoid double counting water savings with changes in crop choice (below in Eq. 3).

To quantify water saved by changes in crop choice (ΔP_{Crop}), we compared water use for the mean crop mix in the pre-LEMA and LEMA periods based on irrigation depths during the LEMA periods. In other words, we first estimated the hypothetical irrigation volume had relative crop areas not changed based on LEMA irrigation depth reductions. This hypothetical volume was then compared with the LEMA period as follows:

$$\Delta P_{Crop} = 5 \times \sum_{i}^{croptypes} (\bar{a}_{0,i} - \bar{a}_{LEMA,i}) \times \bar{d}_{LEMA,i} \quad (3)$$

where $\bar{a}_{0,i}$ is crop-specific mean area in the pre-LEMA period, $\bar{a}_{LEMA,i}$ is crop-specific mean area in the LEMA period, and $\bar{d}_{LEMA,i}$ is crop-specific mean irrigation depth during the LEMA period. Water savings were then compared among management responses.

All raw data used in this study are publicly available online. Derived data along with Earth Engine and R processing scripts can be found at https://zenodo.org/ [full DOI available upon publication].

3. Results and Discussion

3.1. LEMA Impacts on Groundwater Use and Water Table Elevations

We found that irrigators in the SD-6 LEMA significantly decreased groundwater use. Although some reduction was expected given the program's targeted 20% pumping reduction from 2002-2012 levels, analysis of WIMAS pumping data indicated mean annual pumping declined by 36.1%, from 36.4 million m^3 to 23.3 million m^3. However, mean growing season precipitation derived from GRIDMET was 26.7% higher during the LEMA period, suggesting that at least part of the decreased pumping may be related to reduced water deficits.

The BAU scenario generated through causal impact analysis allowed us to quantify the LEMA's impact while accounting for changes in external conditions that can affect irrigation

demand, such as increased precipitation. We found that pumping following establishment of the LEMA decreased 33% compared to BAU estimates (p = 0.001, 95% confidence interval, CI = [23%, 42%], Figure 3.2a). Over the four year period for which pumping data are available (2013-2016), this amounts to a cumulative reduction of 45.2 million m^3 (CI = [32, 58]) or 11.3 million m^3 per year, which is substantial relative to pre-LEMA mean annual pumping volumes (36.4 million m^3). Moreover, we found no significant effect in the control region, where observed pumping volumes closely tracked BAU predictions (p = 0.48, Figure 3.2). This indicates that the changes observed were unique to SD-6 and were not caused by other regional factors.

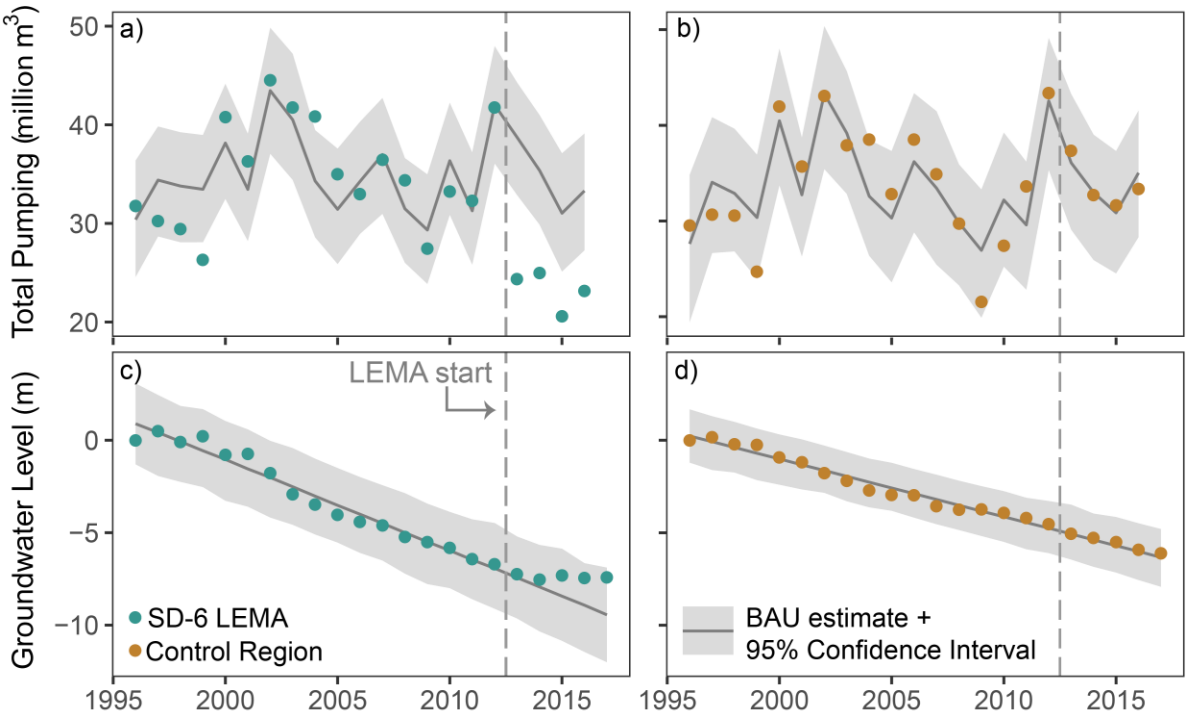

Figure 3.2. Causal impact analyses on groundwater pumping and water table elevations in Sheridan 6 (SD-6) compared to the control region. a) The Local Enhanced Management Area (LEMA) significantly reduced groundwater pumping in SD-6 compared to the modeled business-as-usual (BAU) scenario. b) No pumping change occurred in the control region, where data observations tracked modeled estimates. c) The LEMA resulted in a non-significant increase in groundwater levels compared to BAU expectations. d) Groundwater levels continued to decline in line with BAU estimates in the control region. Groundwater levels in (c) and (d) are relative to water-table position in 1996.

For groundwater levels, we found a non-significant 2.0 m increase in SD-6 relative to BAU expectations through 2017 (p = 0.08, CI = [-0.52, 4.6], Figure 3.2c). In the control region,

we found no evidence of changes in water level trajectory ($p = 0.39$, Figure 3.2d). Given the positive trend evident in SD-6 groundwater levels following LEMA implementation in Figure 3.2c, more time may be needed to be able to fully define the effects on groundwater levels. In particular, as there are relatively few well observations within SD-6 to adequately influence the kriging surface (Figure 3.1), the effects on water levels are likely understated here. We thus hypothesize that this presents a conservative estimate of groundwater level changes. Butler et al. [*In Press*] used a lumped water balance approach to estimate a 67% reduction in the rate of water level decline based on the reduced pumping through 2016, further indicating a positive effect on groundwater levels.

3.2. Land Use Impacts and Farmer Adaptation

3.2.1. Changes in total irrigated area

Analysis of annual, satellite-derived land use (CDL & AIM) and reported irrigated area statistics (WIMAS) suggested that farmers made only minor changes in total irrigated area to meet water reduction targets. AIM irrigated area estimates indicated non-significant -1.8% (T-test, $p = 0.84$) and +3.9% (T-test, $p = 0.62$) changes in irrigated area for SD-6 and the control region, respectively (Figure 3.3a). WIMAS self-reported irrigated area showed the same directions of change, with a statistically significant 4.1% decrease in SD-6 (T-test, $p = 0.003$), and a non-significant 0.4% increase in the control region (T-test, $p = 0.67$, Figure 3.3a). Irrigated area estimates between the two sources generally agreed, although AIM displayed higher variability and tended to underestimate area compared to WIMAS. Overall, our results indicated SD-6 largely was able to sustain nearly the same irrigated cropping area following LEMA establishment. Although irrigated area apparently decreased in SD-6, this 2-4% reduction is modest given the large reduction in irrigation pumping volumes.

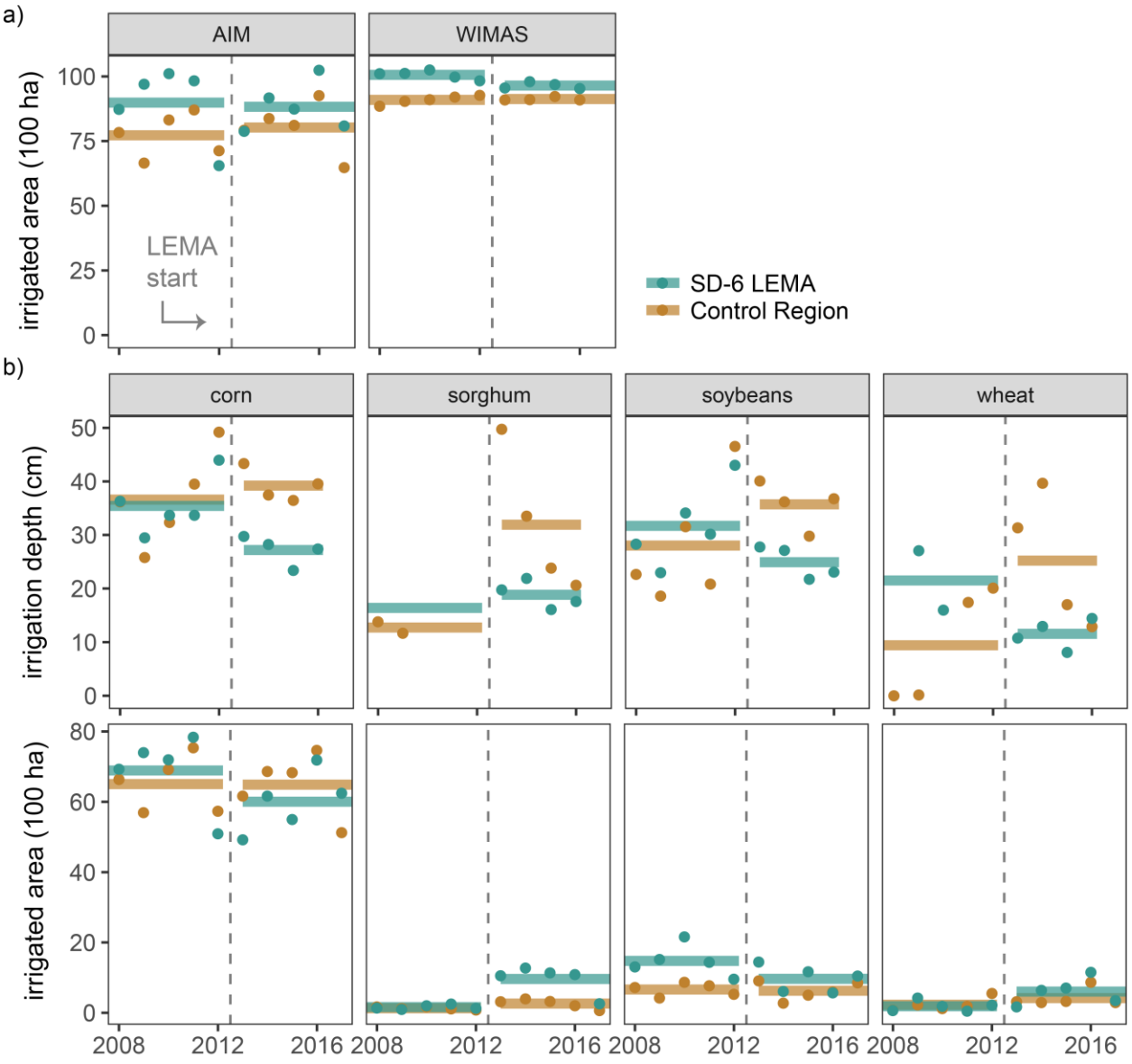

Figure 3.3. Farmer adaptation to water restrictions. a) Changes in total irrigated area based on remotely sensed annual irrigation maps (AIM) and WIMAS irrigator-reported data for the Sheridan-6 (SD-6) Local Enhanced Management Area (LEMA) and the control region. b) Changes in crop-specific irrigation depths derived from WIMAS, and crop-specific irrigated area derived from fusion of USDA Cropland Data Layers and AIM. Colored bars indicate five-year means (2013-2016 means for WIMAS-based irrigation depth). Sorghum irrigation depth mean for SD-6 represents 2002-2012 mean due to lack of sorghum fields in the 2008-2012 period.

3.2.2 Changes in crop-specific irrigation depths

Farmers did show considerable adaptation in terms of water use and crop choices. Based on causal impact analysis of the 1996-2016 WIMAS data, we found the SD-6 LEMA produced

significant decreases in irrigation depths relative to the BAU scenario of 26%, 22%, and 44% for corn, soybeans, and wheat, respectively (Table 3.1). In contrast, within the control region we found no significant changes in irrigation depths for wheat or corn, and a significant increase of 19% for soybeans. We found no significant changes in sorghum irrigation depth for either region, although high uncertainty limited inference (Table 3.1), since there were few single-cropped sorghum fields prior to LEMA establishment in the WIMAS data set for either SD-6 or the control region.

Table 3.1. Causal Impact of the Sheridan-6 LEMA Program on Irrigation Depth by Crop.
* = significant at the alpha = 0.05 level.

Crop	Region	Effect (ϵ, cm)	[95% CI]	Relative effect	p value
Corn	LEMA	-9.4	[-13, -5.9]	-26%	*0.001
	Control	0.2	[-3.6, 4]	0.57%	0.46
Soybeans	LEMA	-7.2	[-14, -1.5]	-22%	*0.005
	Control	5.7	[0.97, 9.8]	19%	*0.006
Wheat	LEMA	-8.9	[-25, 1.9]	-44%	*0.049
	Control	5.6	[-3.5, 15]	29%	0.11
Sorghum	LEMA	3.2	[-36, 40]	21%	0.40
	Control	1.2	[-8.8, 14]	3.8%	0.45

In addition to this causal impact analysis, we also visualized changes in irrigation depths for the pre-LEMA and LEMA periods (Figure 3.3b). Several features likely enabled the substantial reduction in irrigation depths within SD-6. First, structural changes incorporated in the LEMA framework lowered barriers for deficit irrigation practices, which can generate similar yields while using less water [*Chai et al.*, 2016]. For example, it removed the "use it or lose it" system that traditionally could void a water right for non-use [*Streeter et al.*, 2018]. Similarly, it resulted in the development of a limited-irrigation crop insurance product that irrigators could use to avoid needing to meet irrigation depth mandates for full irrigated crop insurance [*Manning et al.*, 2018]. However, few producers in SD-6 took advantage of this change due to the more involved enrollment process and an incomplete understanding of the program (R. Rockel, Kansas

Water Office, personal communication, 27 June 2018).

Beyond lowering structural barriers, the LEMA framework induced SD-6 producers to emphasize net profits, part of a shifting mindset from targeting "highest yield" to "highest return on investment" [*Waskom*, 2017]. For example, reduced water use requires less energy to operate groundwater pumps. Energy supplies traditionally accounted for almost 10% of corn growing costs in western Kansas [*Pfeiffer and Lin*, 2014b]. Similarly, some producers used a lower seed density in irrigated fields as a strategy to maintain a fully irrigated crop under water constraints (R. Luhman, Groundwater Management District #4, personal communication, 27 June 2018). Preliminary analyses comparing production in SD-6 with irrigated fields just outside the LEMA boundary indicate that despite small yield decreases, the majority of LEMA producers reported higher net profit [*Golden and Liebsch*, 2017]. For corn, a 1.2% decrease in yield corresponded to 4.3% higher net profits when comparing 20 fields within SD-6 with 11 neighboring fields outside the LEMA [*Golden and Liebsch*, 2017]. Observations for other crops were low (<5 per class) but suggest that LEMA producers improved water productivity and overall net profit for corn, sorghum and wheat, but not soybeans [*Golden and Liebsch*, 2017].

Finally, water resource use became more efficient through increased awareness and new tools, particularly surrounding irrigation scheduling and soil moisture monitoring [*Lauer and Sanderson*, 2017; *NW KS GMD 4*, 2017]. This allows producers to better take advantage of precipitation events and target irrigation during periods of crop need. Precision agriculture practices can help optimize management by specifying needed water, fertilizer, and other inputs in space and time, reducing waste and increasing net profits [*Basso et al.*, 2013].

3.2.3 Changes in crop choice

SD-6 irrigators also reduced water use by switching to crops with lower irrigation demand, namely planting sorghum and wheat rather than corn and soybeans (Figure 3.3b). When comparing the pre-LEMA and LEMA periods, we found that mean irrigated corn and soybean area decreased 12.9% and 34.5%, respectively, in SD-6. In comparison, the control region had decreases of 0.2% and 5.2% for irrigated corn and soybeans. Both SD-6 and the control region had increases in irrigated sorghum and wheat area, but increases within SD-6 were considerably higher for both crops (sorghum: 493% vs 101%; wheat: 224% vs 82.2%; Figure 3.3b). The annual evolution of this pattern indicated that this could be a flexible strategy to manage the 5-year water allocation cycle of the LEMA program. Basso et al. [2013] suggest there is opportunity across the aquifer to improve sustainability by choosing crops with water requirements that match local availability.

3.2.4 Relative contributions of water conservation strategies

Based on these changes in irrigated area, irrigation depths, and crop types, we estimated the relative contribution of each management response to overall water reductions in SD-6 over the LEMA period using equations (1)-(3). Reductions in irrigation depths accounted for 72.9% of total water savings; reductions in corn irrigation depths accounted for 7/8 of total water saved through this strategy due to irrigated corn's dominance on the landscape (approximately 2/3 of irrigated area in SD-6 during the LEMA period, Figure 3.3b). Changes in crop choice further contributed 19.0% of water reductions, based on the difference between mean crop areas from 2008-2012 versus 2013-2017 using mean crop-specific irrigation depths during the LEMA program. These additional gains are largely due to lower irrigation water requirements for sorghum [*Araya et al.*, 2018] and wheat, which gained area previously used for more water

intensive corn and soybeans (Figure 3.3b). Reductions in total irrigated area accounted for the remaining 8.1% of water reductions.

4. Conclusions

The combined causal impact and control region approach allowed us to quantify the effects of stakeholder-driven groundwater management while accounting for changes in external conditions that can affect irrigation demand. By leveraging rich publicly available datasets, we found that this pioneering LEMA in the High Plains Aquifer in northwest Kansas surpassed goals for reduced water use, leaving enough water in the aquifer in its first four years to provide over a year's worth of historic water needs. Farmers made only minor adjustments to total irrigated area to meet water reductions, instead relying on more efficient water management and less water intensive crops. Preliminary economic analyses suggested that farmers are maintaining net profit despite lower yields due to reduced input costs; the recent stakeholder-voted renewal for another five-year cycle corroborates the economic feasibility of the SD-6 LEMA [*Golden and Liebsch*, 2017].

There remains a need to robustly quantify trade-offs in crop yield as well as impacts to the full water budget, accounting for complexity in the physical system through coupled crop-hydrology models. Because the SD-6 LEMA has unique elements hypothesized to promote self-organization [*Ostrom*, 2009b], the generalizability remains to be tested on larger scales, such as the recently approved LEMA over most of the Groundwater Management District that includes SD-6. As aquifer depletion threatens crop production in many parts of the world, the successful water reduction pathways detailed here can serve as a road map for economically and hydrologically sustainable management.

Acknowledgements

This chapter was co-authored by Anthony Kendall, James J. Butler, Jr., and David Hyndman. We thank Bruno Basso for helpful comments on the manuscript. Funding was provided by NSF grant WSC 1039180, USDA NIFA grant 2015-68007-23133 and USDA-NIFA/NSF INFEWS program grant 2018-67003-27406. Jillian Deines was partially supported by NASA Headquarters under the NASA Earth and Space Science Fellowship Program grant 14-EARTH14F-198. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of NSF, USDA NIFA, or NASA.

CHAPTER 4:
EVALUATION OF STAKEHOLDER-DRIVEN GROUNDWATER MANAGEMENT IN THE US HIGH PLAINS AQUIFER THROUGH CROP MODELING AND REMOTE SENSING

Abstract

Non-renewable groundwater sources contribute approximately 20% of global irrigation water. As a result, key agricultural regions around the world are on unsustainable water trajectories due to aquifer depletion, threatening food production and local economies. With increasing resource scarcity in the central High Plains Aquifer in the United States, an innovative stakeholder-driven groundwater management program has emerged in northwest Kansas referred to as the Local Enhanced Management Area (LEMA) program. Here, we assessed the efficacy of the first LEMA in moving the region towards system sustainability with a process-based crop model driven by remotely sensed annual agricultural land use. We found that groundwater extraction volumes decreased by ~25% as a result of increased irrigation efficiency under the LEMA program, but only 11.8% of this reduction translated to aquifer savings due to a corresponding 27.3% reduction in irrigation return flow. Based on simulated crop-specific yields at sub-field resolution, commodity prices, and energy saved from reduced groundwater pumping, however, we estimated that cost savings from pumping reductions were ~4 times greater than income loss from minor yield penalties, suggesting the program promotes both economic and water sustainability. As aquifer depletion threatens crop production in many parts of the world, approaches that integrate models with in-situ and remotely sensed data can help inform economically and hydrologically sustainable management strategies.

1. Introduction

Over the latter half of the twentieth century, the use of non-renewable groundwater resources for irrigated agriculture more than tripled to become the source for approximately 1/5 of global irrigation water [*Wada et al.,* 2012]. As a result, key agricultural regions around the world are on unsustainable trajectories due to aquifer depletion, including California's Central Valley [*Scanlon et al.*, 2012; *Faunt et al.*, 2016], the North China Plain [*Cao et al.*, 2013], and northern India [*Rodell et al.*, 2009; *Tiwari et al.*, 2009]. In the central United States, the High Plains Aquifer (HPA) provides water for over 6 million hectares of irrigated land [Chapter 2, this volume], accounting for approximately 30% of US groundwater irrigation [*Dennehy*, 2000; *Scanlon et al.*, 2005] and enabling a significant portion of the region's $7.5 billion agricultural net income [*Waskom et al.*, 2006; *NASS*, 2017]. However, water use exceeds natural recharge over much of the aquifer, particularly in its central and southern regions [*Scanlon et al.*, 2012; *Breña-Naranjo et al.*, 2014; *Haacker et al.*, 2016]. Under current use, approximately 24% of irrigated area could be lost by the end of century due to falling groundwater levels [Chapter 2, this volume].

With limited water resources defining the 21st century [*Rodell et al.*, 2018], finding ways to maximize water use and operate within system boundaries is crucial. Traditional water management systems in the western United States can exacerbate the problem. For example, water rights historically could be lost for non-use, providing perverse incentives for farmers to use their full water allocation to maintain their right even in years where this did not benefit crop yield [*Peck*, 2003]. However, systems are changing across the HPA. States are gradually establishing local groundwater management areas that have more power to restrict existing allocations, and court litigation has forced increased regulation in some areas [*Smidt et al.*,

2016]. Most efforts seek to work collaboratively with local producers as stakeholders of water resources to define mutually desirable conservation goals and associated regulations to preserve aquifer lifespan [*Hutchinson*, 2010; *Gleeson et al.*, 2012a]. Despite diverse perspectives, there is growing consensus that better aquifer management is beneficial for both water and economic sustainability in the region.

In Kansas, 2012 legislation created the Local Enhanced Management Area (LEMA) program, which established a framework for local stakeholder groups to work with local and state officials to create enhanced management zones [*K.S.A. 82a-1041*, 2012]. Notably, the LEMA program combines key factors identified for successful management of common resources, including self-organization, local focus, and active monitoring and enforcement [*Ostrom*, 2009a]. LEMAs operate on 5-year cycles and must pass a vote by the local irrigators to go into effect. The first LEMA was operational from 2013-2017 within a 256 km^2 highly stressed region in northwest Kansas known as Sheridan-6 (hereafter SD-6, Figure 4.1a) [*KDA*, 2013]. The SD-6 LEMA included restrictions to reduce total groundwater pumping by 20% compared to 2002-2012 levels [*NW KS GMD 4*, 2016]. This was implemented by reducing existing water rights to a 5-year allocation of 55 inches (139.7 cm) per irrigated acre, with the flexibility to apply up to 11 inches of unused water to subsequent LEMA cycles.

Initial assessments of SD-6 LEMA effectiveness show promising results. Farmers exceeded water reduction targets for 2013-2016 [Chapter 3, this volume], resulting in an estimated 67% reduction in the rate of water table decline [*Butler et al.*, 2018]. Water savings have largely been accomplished without reduction in irrigated area through improved management, including increased water use efficiency and use of crops with lower irrigation demand such as sorghum [Chapter 3, this volume]. Moreover, preliminary analyses indicate that

despite slight yield loses (~1.2%), net profits have been stable or increased due to reduced energy and input costs [*Golden and Liebsch*, 2017]. Stakeholders voted to renew the LEMA for 2018-2022, indicating the program's continued support among producers. There remains a need, however, to robustly quantify the effects on the full water budget and trade-offs in crop yield, particularly as a second LEMA was approved in the spring of 2018 for most of the surrounding groundwater management district (GMD) (Figure 4.1a). Discussions for additional LEMAs in parts of three other Kansas GMD's are ongoing. Understanding the effectiveness of the SD-6 LEMA can improve understanding and help inform economically and hydrologically sustainable management strategies across the HPA and elsewhere.

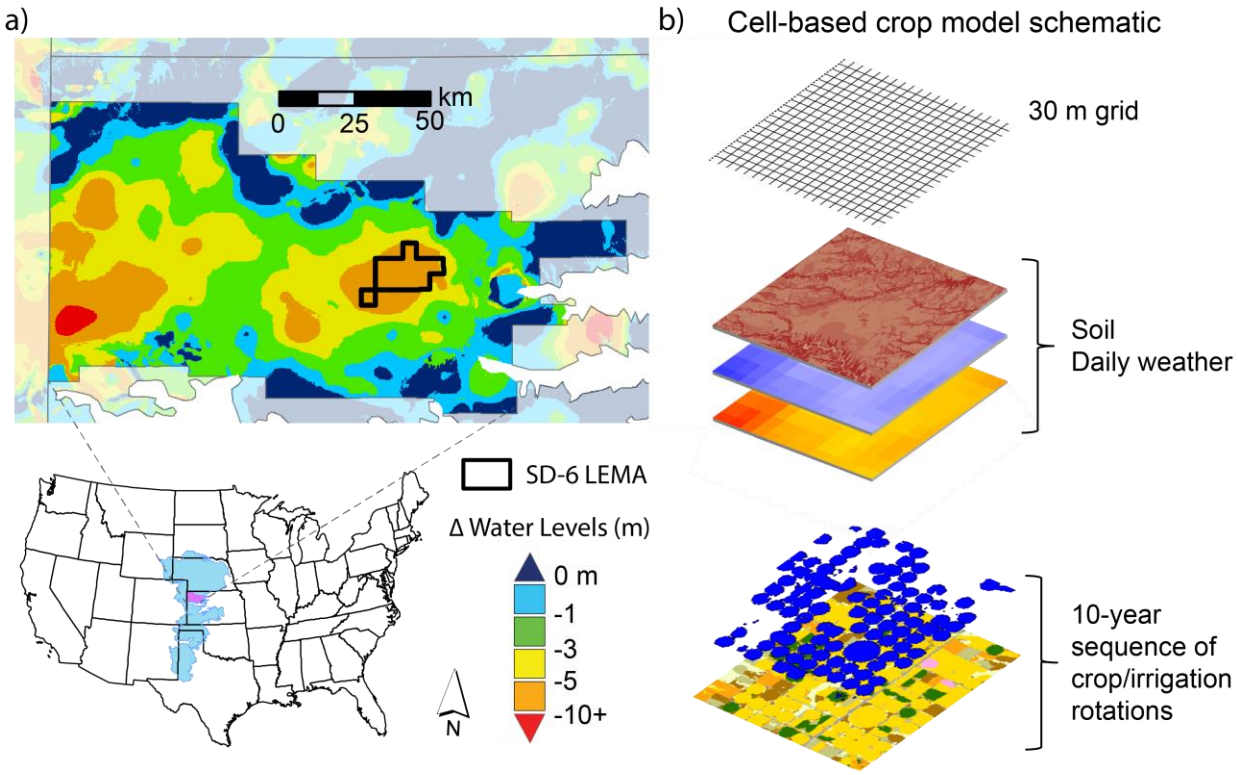

Figure 4.1. Study area and modeling approach. (a) Location of the Sheridan-6 (SD-6) Local Enhanced Management Area (LEMA) within Groundwater Management District 4 (bold colors) in the United States High Plain Aquifer (inset, blue), including change in groundwater level between 1996-2012 prior to the start of the LEMA in 2013. (b) Modeling approach using the SALUS crop model. Models were run for each 30 x 30 m grid cell based on cell-specific soil, climate, and 10-year crop rotation sequences. Results were then aggregated for SD-6.

Although the state of Kansas is relatively data rich, existing datasets are inadequate for full evaluation of the SD-6 LEMA [Chapter 3, this volume]. A reliably curated water well dataset is publicly available to track annual groundwater pumping volume [*KDA DWR*, 2017], but this is not adequate on its own to fully quantify how reduced irrigation applications affect aquifer recharge via irrigation return flow. Because the region is relatively flat with few perennial streams, excess irrigation water generally infiltrates into the ground and returns to the aquifer as enhanced natural recharge. Effective aquifer conservation cannot be determined without factoring in this component. Similarly, remote sensing products specifying annual agricultural land use from 2006 to present [*Deines et al.*, 2017; *USDA-NASS*, 2017] provide information on annually varying crop proportions and active irrigation locations but cannot be translated to changes in crop income without simultaneously measuring changes in crop yields. Currently, yield data representative of the SD-6 region is not available. This, combined with limited literature or empirical data that relates relationships between reduced water use and producer net profits [*Golden and Liebsch*, 2017], limits our ability to quantify the economic feasibility of programs such as this LEMA.

Process-based crop models provide an opportunity to simulate difficult to measure quantities such as irrigation return flow, evaporative water losses, and crop-specific yield responses to different irrigation regimes. By simulating regional yields, water reductions can be translated into estimated income loss from yield penalties due to reduced irrigation application. Similarly, model estimates of business-as-usual water use can estimate water savings, which can be translated to energy cost reductions for groundwater pumping. These components are vital to assess agricultural sustainability, which must consider economic local socio-economic conditions in addition to environmental protections [*Häni et al.*, 2003].

Here, we applied the System Approach to Land Use Sustainability (SALUS) crop model [*Basso and Ritchie*, 2012] to simulate historic cropping and water use in SD-6 and evaluate the sustainability implications of the LEMA program. To accurately quantify changes in crop yields, water use, and irrigation return flow, we developed a new methodology to drive the SALUS model with spatially explicit annual crop and irrigation based on remote sensing data. We used this modeling framework to ask the following questions: 1) What was the effect of the LEMA program on net aquifer change, considering both changes in groundwater extraction as well as irrigation return flows? and 2) How did the LEMA program affect crop yields and cash flow, incorporating yield penalties and energy saved associated with reduced pumping?

The approach presented here simulated water savings due to changes in irrigation depth and frequency while assuming that the spatial distribution and annual sequence of crop types and irrigated area remained the same. Reducing irrigation depth and frequency is one of three key strategies available for farmers to reduce water use; in this case we estimated that this accounts for ~73% of reduced groundwater pumping during 2013-2017 [Chapter 3, this volume].

2. Methods

2.1. SALUS crop model

Here, we used the System Approach to Land Use Sustainability (SALUS) model [*Basso et al.*, 2006; *Basso and Ritchie*, 2012] to simulate crop yield and crop water budget components, including evapotranspiration (ET), irrigation return flow, and irrigation water applied. SALUS is a crop bioenergetics model that simulates daily plant growth and soil-water-nutrient conditions under a range of potential specified management conditions. It improves upon the widely validated CERES crop model [*Ritchie and Otter*, 1985] and has been well validated [*Hoang et al.*, 2014; *Dzotsi et al.*, 2015], including for the central HPA [*Cotterman et al.*, 2018].

SALUS is a point-based model that runs continuously (vs being annually reinitialized), accumulating year-to-year changes to properly track components across years [*Basso et al.*, 2015] and simulating both the actively-cropped and dormant periods of the year. Therefore, we developed a set of real-world, spatially-explicit experiments that captured both the year-to-year sequences of crop choice and irrigation status along with the soil and climate variability across the study area. We discretized the study area into 30 m grid cells based on the resolution of the remotely sensed inputs (see section 2.2), and simulated individual experiments for each grid cell based on cell-specific soil properties, observed climate, and annual crop type and irrigation status (see subsequent sections) between 2006-2017 (Figure 4.1b). The absence of spatially-explicit crop data prior to 2006 limited an earlier start of the study period, thus 2006-2007 provided our model spin-up period. SALUS results were then summarized based on five year periods, hereafter termed Pre-LEMA (2008-2012) and LEMA (2013-2017). For this analysis, we analyzed yields for irrigated crops only to isolate the effects of LEMA water restrictions. Rainfed crops, fallowed fields, and grasslands were modeled to generate spatially complete estimates of soil drainage, thus enabling us to assess changes in groundwater recharge due to changes in irrigation regimes due to the LEMA program.

2.2. Annual crop and irrigation map data

The USDA's NASS Cropland Data Layers (CDL) [*Boryan et al.*, 2011; *USDA-NASS*, 2017] provide annual maps of crop types for 2006-2017 at 56 m resolution for 2006-2007 and 30 m resolution since 2008. This data is derived from a suite of satellite imagery and has reasonable accuracy in Kansas for the crops considered in this study, with accuracies ranging from approximately 70 – 95% for the study period depending on year and crop type [*USDA-NASS-RRD Spatial Analysis Research Section*, 2017]. To minimize spurious crop classification, we

first smoothed annual CDL images with a two-step process in Google Earth Engine (GEE) [*Gorelick et al.*, 2017], which hosts the CDL product in its cloud-based data catalog. First, we resampled all CDL years to 30 m where needed and ran a connected pixel counter that identifies the number of connected pixels by class type. We flagged patches with < 25 pixels (2.25 ha) as potentially spurious classifications given the typical field size ~60 ha in this region. Second, we updated values at these flagged patches with the most common crop class within a circular moving window with radius of 120 m. The effect of this "despeckling" process is to re-classify small patches with the most common class type within a moving local window, thus reducing the "salt and pepper" effect often produced from per-pixel classifications. We then used the 2016 TIGER road vector dataset to update underlying pixels as roads. This provided consistent location of roads across CDL years and protected them from the implemented despeckling technique, since roads are often small patches due to imagery resolution and their long, narrow shape.

Analysis of CDL data in R [*R Core Team*, 2014] indicated that from 2006-2017, corn area dominated within SD-6, followed by grassland, winter wheat, fallowed land, sorghum, soybeans, developed land, and alfalfa. Together, these classes covered 98.7% of the SD-6 study area (2006-2017 average). Of these top 8 classes, we simulated the water balance in all 7 non-developed classes. The LEMA program lead to changes in proportional crop area compared with the previous five years [Chapter 3, this volume], including a decrease in corn and soy along with a corresponding increase in sorghum and wheat areas (Figure 4.2). Developed land covered 3.8% of the study region, largely as roads, and was not included in recharge estimates. In reality, roads and other impervious surfaces can serve as points of enhanced recharge due to concentration of runoff along the edges of these surfaces. However, because we are focused on recharge

differences resulting from the LEMA program, we don't consider recharge along these impervious surfaces here simplicity. We used CDL maps to assign a series of annual crop rotations to each 30 m pixel in our study area grid for 2006-2017.

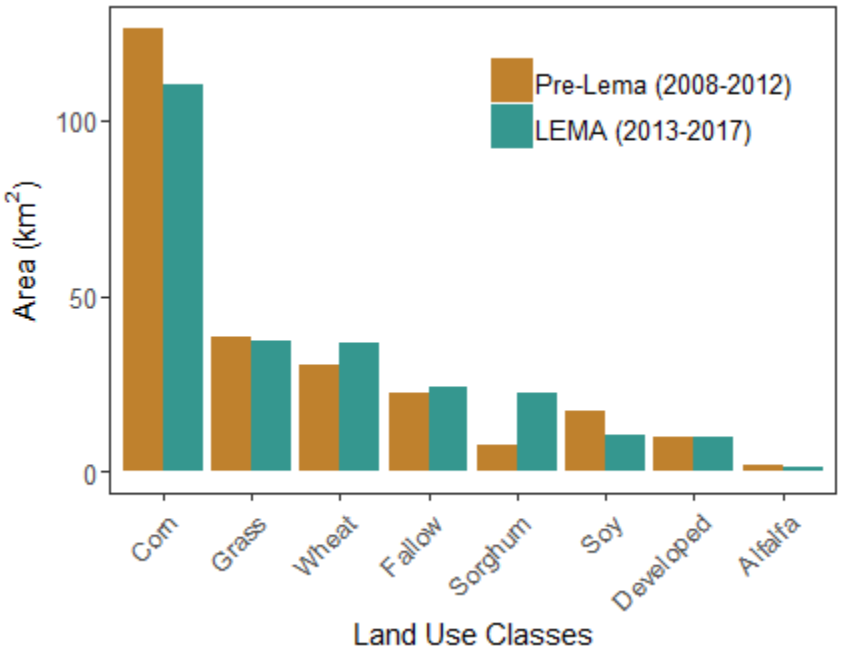

Figure 4.2. Dominant land cover in the Sheridan-6 Local Enhanced Management Area. Mean area for the eight predominant land cover classes based on NASS Cropland Data Layers [*USDA-NASS*, 2017] for the five years prior to the LEMA (2008-2012) and the initial 5-year LEMA period (2013-2017).

Irrigation status was assigned for each 30 m grid cell based on the remotely sensed Annual Irrigation Maps – High Plains Aquifer (AIM-HPA) dataset [Chapter 2, this volume]. AIM-HPA provides annual irrigation status at Landsat resolution, allowing us to specify historic irrigation occurrence during the 2006-2017 model and spin-up period. AIM-HPA is estimated to have a 91.7% overall accuracy, leading to occasional misclassified pixels. To minimize these effects, we filtered AIM-HPA with allowable place-of-use tracts maintained by the Kansas DWR [*KS DWR*, 2017], removing any irrigated pixels that occurred outside of these tracts.

2.3. Weather and soil data

Daily climate data were obtained from GRIDMET, a daily 4 km gridded surface meteorological dataset that includes precipitation, maximum daily temperature, minimum daily temperature, and shortwave radiation [*Abatzoglou*, 2013]. Data was accessed through GEE, where we combined daily observations with mean elevation per grid cell derived from the USGS National Elevation Dataset [*USGS*, 2012]. The resulting data was formatted for the SALUS model using R. Soil data was obtained from SSURGO, which provides soil classes at 30 m resolution and extensive soil properties [*NRCS*, 2016]. There are 16 SSURGO soil map units in SD-6, texturally dominated by silty loam.

2.4. SALUS experiments

We extracted all unique combinations of soil type, GRIDMET climate cell, and 12-year crop/irrigation rotations including fallow and grassland classifications. Some spurious combinations arose due to the inherent uncertainty in remote sensing products, such as irrigated grassland. In these cases, irrigation status was changed so that crop cover and irrigation status aligned. Crop rotation sequences had missing crop values in 4.0% of non-developed cells due to less common crop choices, such as sunflowers, omitted from this study. For cells with crop sequences missing two years or less due to these uncommon crop types, we assigned the alfalfa class in the missing years as a reasonable proxy to maintain model output for these locations across years. Sequences with more than two years of non-included crops were not simulated (0.4% of non-developed area). This resulted in 48,279 unique experiments from 270,814 active SALUS cells not covered by roads or other developed lands, repeated rare crops, or small water bodies, covering 95.5% of the total study area.

SALUS crop management and water use parameters were defined in a three-step process.

First, we gathered information from available agronomic resources related to crop growth in northwest Kansas to set base model parameter conditions. Management parameters including annual state median planting dates were obtained through the USDA's National Agricultural Statistics Service [*NASS*, 2017]. Typical planting densities and regional offsets from state planting dates were obtained from the Kansas Crop Planting Guide [*Shroyer et al.*, 1996]. Initial cultivar selection for corn and wheat were taken from a previously published SALUS model in the adjacent Central High Plains portion of the HPA, including large parts of Kansas [*Cotterman et al.*, 2018]; soybeans, sorghum, and alfalfa were run in "simple" mode within SALUS since "complex" mode with specific cultivars are not yet available for these crops. All crops were set to harvest at maturity as determined by the model. For this study, nitrogen was assumed to be non-limiting. We therefore ran SALUS with the nitrogen module off. All fields were set to no-till for parsimony, based on Cotterman et al. [2018]. Finally, although we planted pasture grass to mimic grassland and fallow land use types within the SALUS runs, we had difficulty simulating reasonable recharge values for perennial grasses. To get an improved overall recharge estimate while ultimately not affecting estimation of the relative impact of LEMA water reductions, we manually overrode recharge estimates for these cells using 1% of annual precipitation based available literature [*Hansen*, 1991].

Second, we calibrated these reference-informed starting parameters for cultivar type and planting density to better match state yield data from NASS. County-level annual statistics were also considered when available but were infrequent for the study area. Rainfed and irrigated crops were calibrated separately. During yield calibration, irrigation parameters were set to be non-limiting by delivering 25 mm (~1 in) of irrigation via sprinklers when soil moisture content dropped below 75% of plant available water (defined as the difference between the drained upper

limit and lower limit of the soil). Initial parameters for soybeans and sorghum performed well, so these were not modified. The medium-high yielding wheat variety and planting density from Cotterman et al. [2018] matched yield data better than other available cultivars and was thus selected for both rainfed and irrigated wheat. For corn, we found good agreement with NASS statistics by using a low yielding cultivar for rainfed fields, and a moderate yielding cultivar for irrigated fields. We did not calibrate alfalfa parameters due to its small proportional representation on the landscape. Final crop-specific cultivars and planting densities were uniform for the entire period.

Finally, we used this calibrated yield model to develop two scenario models for analysis: 1) one calibrated to pre-LEMA irrigation behavior and water use from 2008-2012, which we used to estimate business-as-usual (BAU) pumping and irrigation during the LEMA period, and (2) one which used BAU parameters for the pre-LEMA period but was then calibrated to LEMA irrigation behavior between 2013-2017. To identify SALUS parameters matching each period, we iteratively varied the soil moisture threshold, which triggers SALUS's automatic irrigation between 25 and 90% of plant available water in steps of 5% for each of four application depths: 12.7, 19.1, 25, and 31.8 mm (roughly 0.5, 0.75, 1, and 1.25 in.). These are reasonable given irrigation management in the region [*Kranz et al.*, 2008]. Fully crossing these parameters resulted in 56 model runs. To account for water extracted from the aquifer but lost in delivery through factors such as wind-drift evaporation and other efficiency factors, we then divided SALUS modeled irrigation volumes for each run by a 90% efficiency penalty based on typical values for well-maintained center pivot sprinkler irrigation systems [*Kranz et al.*, 2008]. Finally, we used the WIMAS well water use database maintained by the Kansas Geological Survey to select the irrigation parameter set that best estimated actual pumped water use during the target

periods for each scenario. To access system changes from the LEMA program, modeled yields and water budget components from the two scenarios were compared during the LEMA period.

2.5. Economic analysis

To assess the economic sustainability of the LEMA program, we estimated changes to farmer net income based on differences in regional crop yields and water use between the BAU and LEMA model scenarios. To estimate income from crop production, we obtained annual crop prices from NASS and adjusted prices to consistent 2017 dollars per kilogram of yield. We then summed simulated annual crop-specific yields for the full SD-6 region for each scenario (total kg of production for each crop), and converted this to monetary regional totals based on price data. For this analysis, we assume that other costs associated with production such as fertilizer, seed, equipment, land, and labor are fixed across scenarios to focus only on yield and price.

To estimate monetary savings from reduced pumping costs, we first quantified the pumping volumes for the BAU and LEMA models for 2013-2017. We then translated the volume of water extracted into required energy based on a uniform 3.1 megajoules per cubic meter [*McCarthy et al.*, In Prep]. A typical cost for industrial energy in the state of Kansas of 1.97 cents per megajoule (https://www.electricitylocal.com/states/kansas/) was used to convert this energy required for pumping into dollars. We then calculated differences in total pumping costs between the BAU and LEMA scenarios.

3. Results and Discussion

3.1. Model calibration

To assess the impact of the LEMA groundwater management program, we calibrated the SALUS crop model to represent historic yield and water use conditions in the SD-6 LEMA.

SALUS simulated yields generally showed good agreement with NASS state statistics (Figure 4.3). These statistics were most complete for corn, with statewide crop yield for both irrigated and rainfed corn in all years of the calibration period (2008-2017). Median simulated yields were within 14% and 34% for irrigated and rainfed corn, respectively. NASS data for Sheridan county (which encompasses most of SD-6) indicated that simulated rainfed corn yields represented local conditions better than statewide conditions in 3 out of 4 available years (Figure 4.3). Given a wide E-W precipitation gradient across Kansas, it's likely that rainfed SALUS yields estimated SD-6 yields with more accuracy than statewide yields.

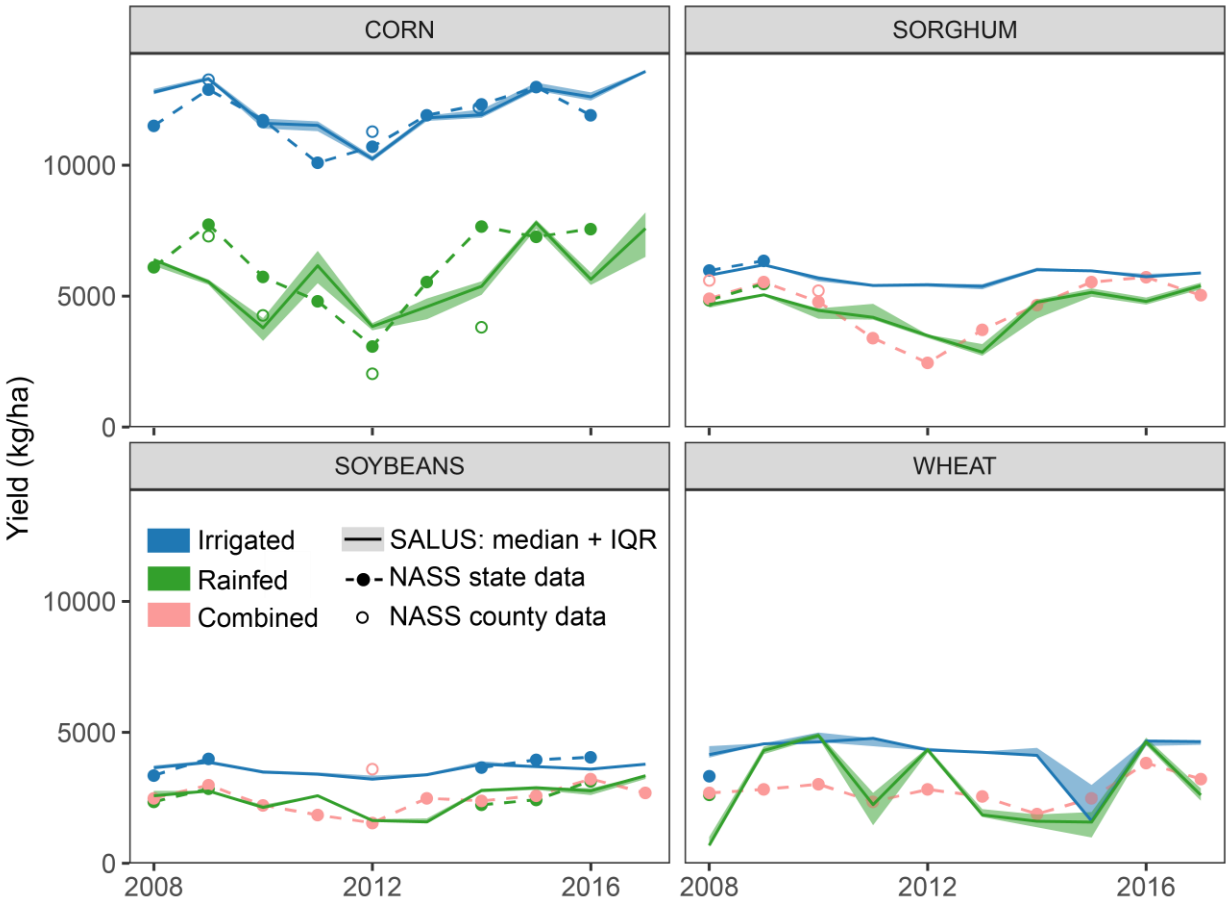

Figure 4.3. SALUS simulated yield validation, 2008-2017. SALUS calibrated yields for the four primary crops in the Sheridan-6 LEMA by irrigation status. USDA NASS annual statewide crop yield statistics as well as Sheridan county statistics are shown where available. Values for combined irrigated and rainfed fields are shown when NASS data were not available by irrigation status.

Available NASS yields were less frequently separated by irrigation status for sorghum, soybeans, and wheat. When irrigated statistics were available (Figure 4.3), simulated yields were within 3%, 11%, and 24% for sorghum, soybeans, and wheat, respectively. Combined yields for all irrigated and rainfed fields were available in other years. Although these summarize both irrigated and rainfed fields across the state, the close agreement between combined and rainfed data when both are available (e.g., sorghum in 2008-2009; soybeans in 2008-2009 and 2014-2016, Figure 4.3) suggested that combined statistics best represent rainfed conditions statewide. Reasonable agreement between SALUS simulated rainfed yields and NASS combined yields indicated that SALUS captured rainfed crop growth sufficiently. The sole Sheridan county data point for combined soybeans was also similar to simulated irrigated yields in SD-6 since the majority of soy grown in the study area is irrigated. Simulated yields for rainfed wheat were inconsistent with available data, with rainfed yields nearing irrigated yields in 2009, 2010, and 2012. Combined state-level wheat yields from NASS did approach irrigated yields in 2016, so it is possible that conditions were particularly favorable for wheat in SD-6 for these years, but not for the wider state. Without county or region specific yield data, it is difficult to determine if this is a modeling artefact or indicative of SD-6 yields. Wheat yields therefore present an area for future investigation and improvement.

For simulated water use, we selected the best SALUS model for each scenario based on agreement with WIMAS well pumping data from among the 56 model runs used for irrigation parameter selection (Figure 4.4b). For the BAU scenario, the model that best captured 2008-2012 groundwater use delivered 31.8 mm (1.25 in) irrigation applications when soil moisture dropped below 85% of plant available water. For the LEMA scenario, the model that best captured 2013-2017 groundwater use delivered 25 mm (1 in) irrigation applications when soil moisture dropped

below 80% of maximum capacity. Interestingly, the LEMA scenario model displays opposite year-to-year trends than WIMAS data in 2 out of 4 years, whereas the BAU scenario matches year-to-year trends during 2008-2012. It's plausible that the observed pumping quantities don't follow model simulations during LEMA because producers are now operating under a 5-year water budget, which likely influenced their cropping and irrigation decisions.

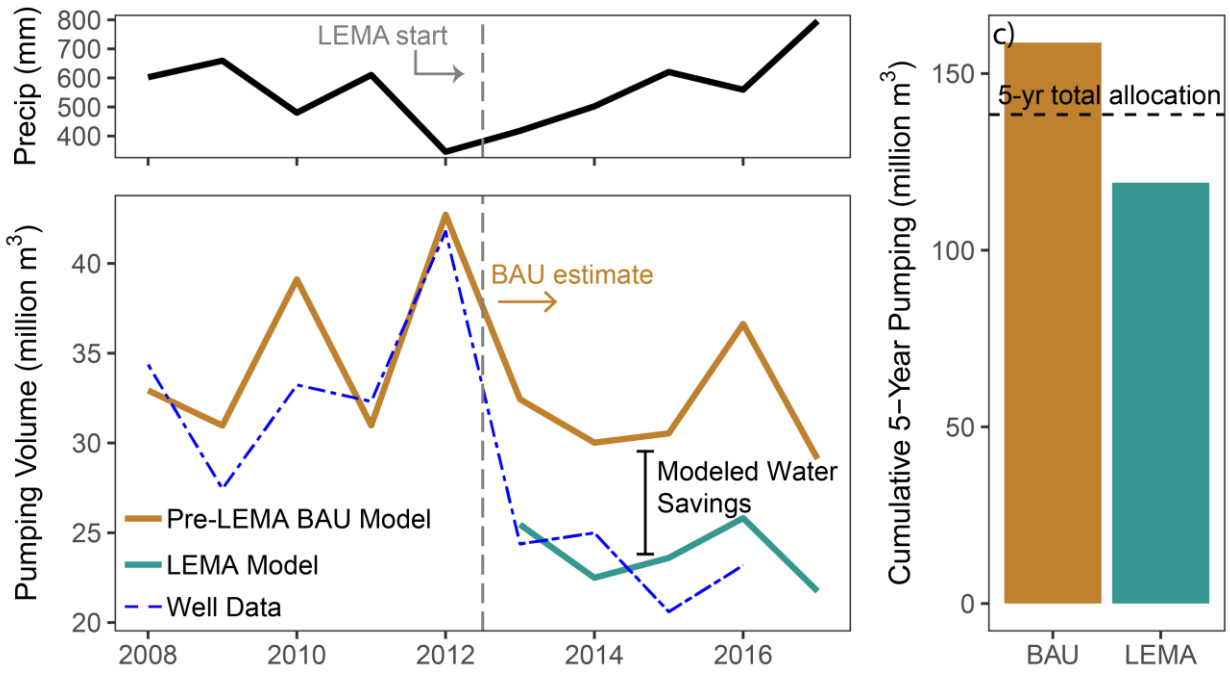

Figure 4.4. SALUS crop model water use for the business-as-usual (BAU) and Local Enhanced Management Area (LEMA) scenarios. (a) Mean annual precipitation for the Sheridan-6 (SD-6) study region. (b) Total pumping volume estimated for SD-6 via SALUS for the BAU scenario based on pre-LEMA groundwater use (2008-2012, brown) and the LEMA scenario, based on LEMA groundwater use (2013-2017, green). Actual irrigation pumping volumes extracted from WIMAS well data [*KDA DWR*, 2017] is indicated with the blue dashed line. Following the start of the LEMA program in 2013, the pre-LEMA model served as a business-as-usual (BAU) estimate of water use had the LEMA not been implemented. Differences between the BAU and LEMA scenario models from 2013-2017 represent reductions in pumping volumes due to the LEMA program. (c) Total 2013-2017 pumping volume based on the BAU and LEMA scenarios. Dashed line shows the total water allocation for the five year LEMA period, based on 20% of 2002-2012 water use.

Based on these selected irrigation parameters, producers on aggregate could achieve the observed water savings by irrigating less often (here, when the soil water threshold drops below 80%, instead of 85% in the BAU scenario) and at reduced application irrigation depths (25 mm

instead of 31.8 mm). These irrigation application depths fall within observed practices in the region, where medium and fine-textured soils provide adequate field capacity for 19.1 – 33 mm (0.75 – 1.3 in) water applications for corn [*Kranz et al.*, 2008], the dominant crop in SD-6. Typical ranges for soil moisture thresholds triggering irrigation for this region were not available to corroborate parameter estimates. If the assumed 90% irrigation efficiency was too high, it is possible that soil moisture thresholds are artificially inflated; lower efficiencies would require more water extracted from the aquifer to meet simulated water demand by SALUS, resulting in models parameterized with lower soil moisture thresholds better matching the WIMAS well data. Regardless of this uncertainty around the absolute values, the selected model scenarios revealed a decrease in soil moisture threshold for LEMA compared to BAU. Combined with the lower application depth, these parameters reflect on-the-ground observations that SD-6 farmers are becoming better groundwater managers through increased awareness of irrigation scheduling and soil moisture monitoring [*Lauer and Sanderson*, 2017; *NW KS GMD 4*, 2017].

3.2. LEMA program impacts on the regional water budget and crop yields

We then compared model scenarios to quantify LEMA-induced changes in water use, aquifer net balance, and crop yields. Based on modeled water use, we found that the LEMA program reduced total 5-year groundwater use by 25% to 119 million m^3 compared to BAU estimates of 159 million m^3 (Figure 4.4c). This translates to average annual savings of 7.9 million m^3. The resulting 39.6 million m^3 saved over the 5-year LEMA was greater than BAU estimates for mean annual pumping (31.8 km^3). The region surpassed the 20% water reduction goal, which would not have been met under BAU irrigation behavior (Figure 4.4c).

The approach presented here modeled water savings from changes in irrigation depth and frequency assuming that the spatial distribution of crop types and irrigated area remained static.

Previous work based on statistical modeling found that such changes in overall irrigation depths accounted for 72.9% of water savings in the SD-6 LEMA [Chapter 3, this volume]. Thus, modeled estimates from SALUS were similar to previous findings based on statistical analysis in Chapter 3 while providing further insight into how water savings translate to yield and recharge changes. The causal impact analysis based on available 2013-2016 data estimated that in total, the SD-6 LEMA reduced water use by 33% at an average rate of 11.3 million m^3 per year across adaptation strategies. Based on the 72.9% contribution, reductions in irrigation depth would contribute 24.1% of overall water savings at a rate of 8.2 million m^3 per year. The close agreement from these two complementary approaches based on statistical and process-based modeling indicates that the SALUS scenarios developed here reasonably captured observed changes due to the LEMA program, lending confidence to using the SALUS model for full water budget evaluation, yield assessment, and economic analysis.

Once irrigation water is applied, SALUS simulated whether the water transpired through the plant, evaporated from the soil, or infiltrated the ground past the root zone as irrigation return flow, thus recharging the aquifer. We found a small mean annual decrease of 0.73% (standard deviation, s.d. 0.57), 0.62% (s.d. 0.13), and 1.6% (s.d. 3.9) in the LEMA scenario for total SD-6 plant transpiration, soil evaporation, and run off, respectively, totaling a net water savings of 4.62 million m^3. The soil evaporation ratio between BAU and LEMA was nearly constant across years, whereas differences in plant transpiration were larger in dry years such as 2013 and nearly identical in wet years such as 2017, likely accounting for the differences in yield (below). The five-year recharge for the LEMA scenario totaled 114 million m^3, a 27.3% decrease compared to the BAU scenario. Figure 4.5 shows the spatial distribution of this decreased recharge across SD-6, illustrating that the reduction in recharge is driven by irrigated fields (as denoted by the

characteristic circular shape due to the dominance of center pivot irrigation technology in the region). Figure 4.5 demonstrates how our method can capture sub-field level variability.

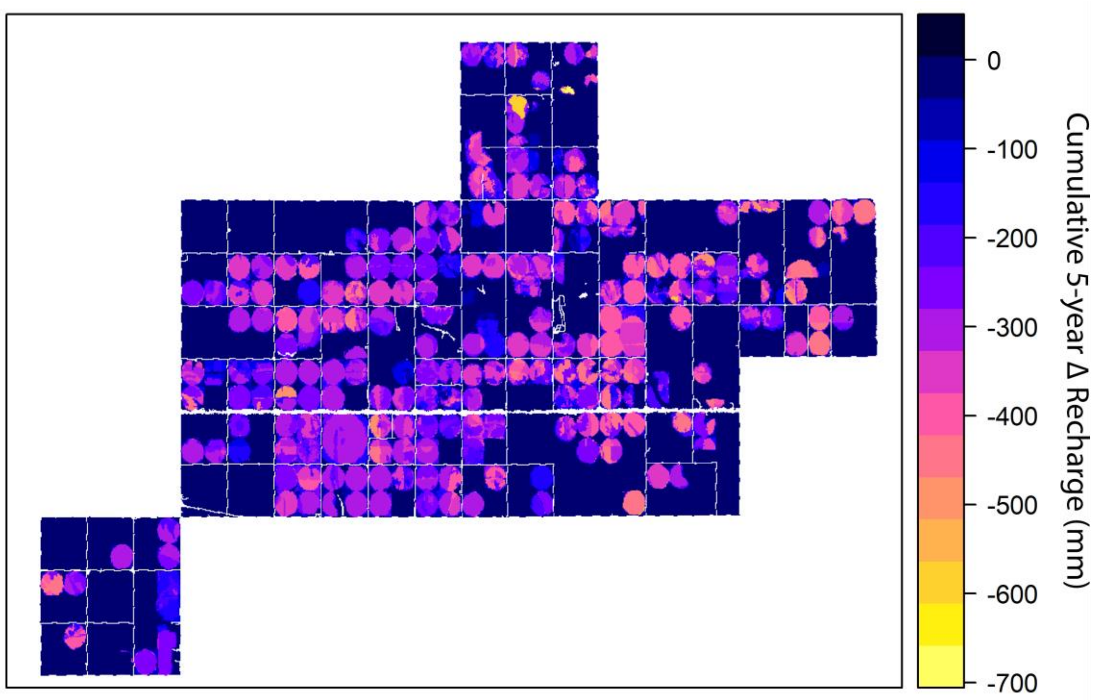

Figure 4.5. Cumulative change in recharge in Sheridan-6 from the LEMA program, 2013-2017. Spatially-explicit comparison of total recharge in Sheridan-6 (SD-6) during the Local Enhanced Management Program (LEMA) compared to business-as-usual (BAU) scenario based on SALUS crop model output. LEMA-induced reductions in irrigation applications resulted in reduced recharge across agricultural fields, leading to a 32.1% decrease compared to BAU. Circular shapes are due to center pivot irrigation technology, which dominates SD-6. White lines represent "no data" for roads, which were omitted from the model.

For net aquifer change based on the difference between water extracted and irrigation return flow, we found that the net change (extraction – return flow) was negative or near zero for all years with the exception of 2017, where high precipitation totals contributed to a positive net change (Figure 4.4a, Figure 4.6a). This corroborates studies reporting low recharge in the region, often less than the extraction rate except during very wet years which result in rare but substantial recharge [*Whittemore et al.*, 2016].

When comparing the net change between scenarios, we found that the LEMA scenario

had a more favorable net balance than BAU with respect to aquifer preservation in all years (Figure 4.6b). Over the five-year period, this amounted to 4.68 million m^3 of water gain compared to the BAU estimates (Figure 4.6c), roughly equivalent to the 4.62 million m^3 of water saved in the LEMA scenario from decreases in plant transpiration, soil evaporation, and run off. Given that SD-6 covers 256 km^2, this translates to a regional average 5-year recharge depth of 18.3 mm.

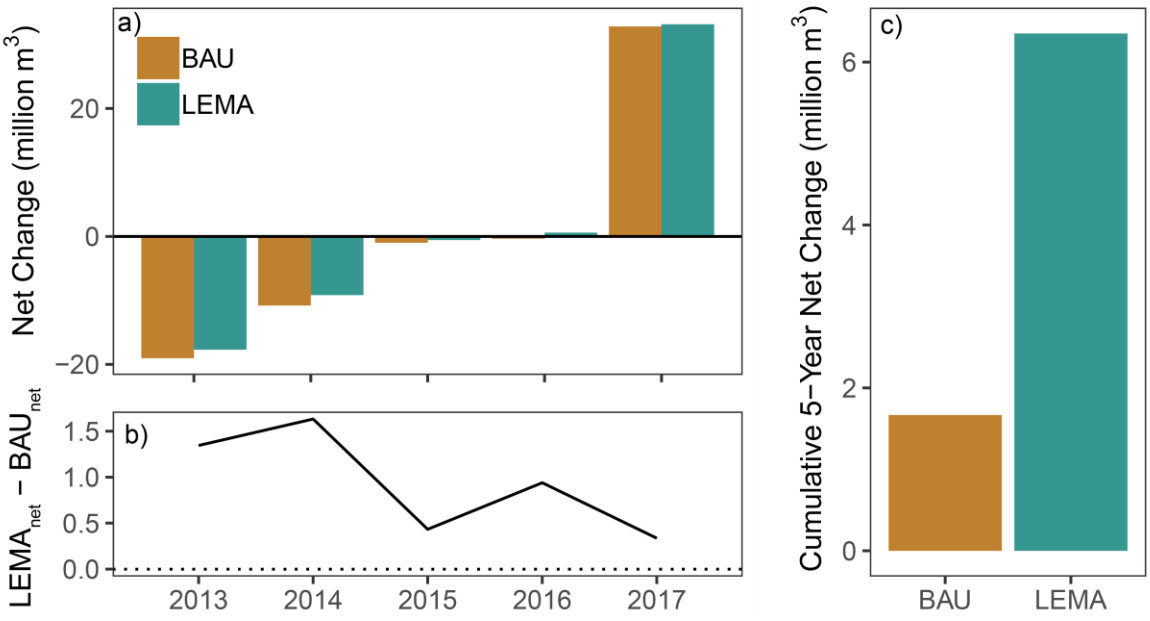

Figure 4.6. Net groundwater savings quantified by aquifer balance. Aquifer balance indicated by net change calculated from the SALUS crop model (recharge – pumping). (a) Annual net change by model scenario. (b) Annual difference between net change from LEMA and BAU model scenarios. Negative values indicate years where the BAU model conserved more groundwater as indicated by the annual net change, and positive values indicate years where the LEMA model conserved more water. (c) Cumulative net change for the 5-year LEMA management period (2013-2017).

The yield penalty from LEMA-induced irrigation reductions was small. By comparing the LEMA and BAU scenarios, we found that the 5-year mean decrease in median yield was 0.67%, 1.21%, 1.41%, and 0.07% for corn, sorghum, soybeans, and wheat, respectively (Figure 4.7). For all crops, interannual differences in median yield were larger than differences between modeled scenarios.

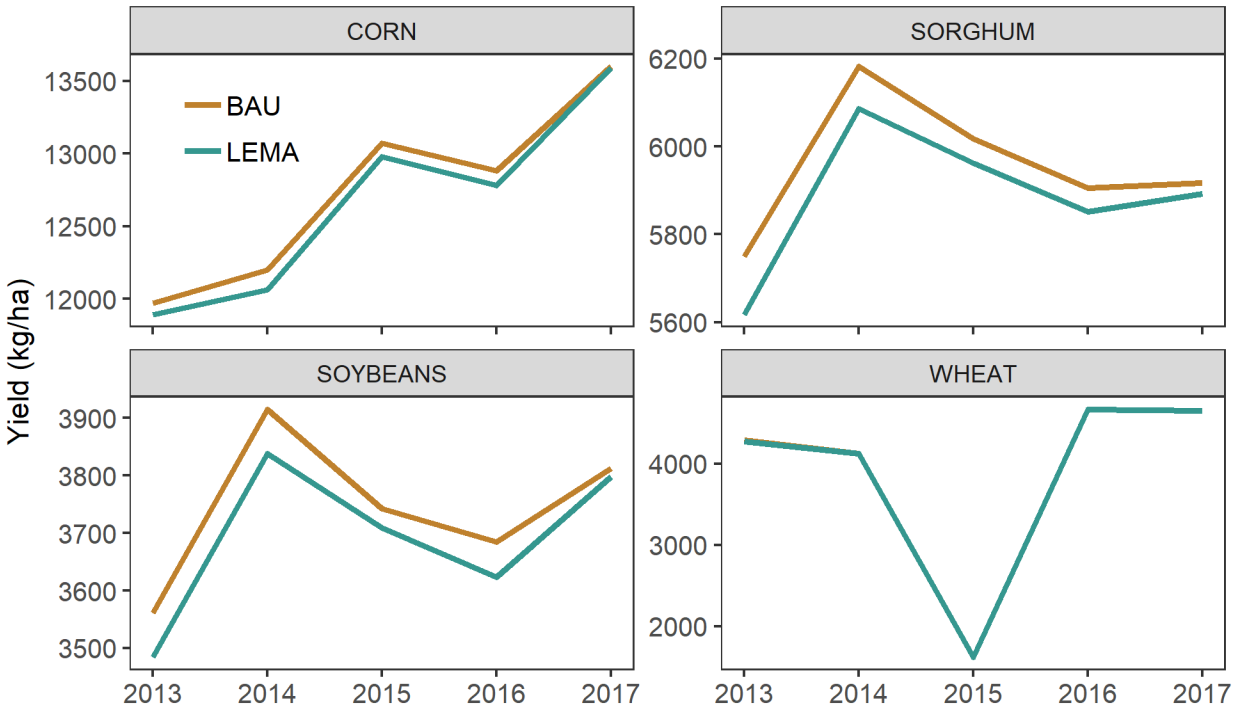

Figure 4.7. Yield penalty due to reduced irrigation water use, 2013-2017. Annual median yield by crop estimated from the SALUS crop model for the Sheridan-6 Local Enhanced Management Area (LEMA). The LEMA scenario (teal) slightly underperformed the business-as-usual (BAU) scenario (brown) for all crops except wheat, which were highly similar between the two model scenarios.

The ability to track yield effects of different irrigation regimes is a key metric needed to evaluate the LEMA program, but to date, data on crop yields for the region has been collected only from voluntary reporting on a limited number of fields and disconnected from soil and rainfall attributes, inhibiting robust statistical conclusions [*Golden and Liebsch*, 2017]. Preliminary conclusions from this data for corn, which is best represented with 20 observations within SD-6 and 11 observations in neighboring fields, suggest there was a 1.2% yield decline due to the LEMA water restrictions [*Golden and Liebsch*, 2017]. Given the uncertainties involved, this provides some support that yield declines estimated by SALUS are valid.

3.3. Economic analysis results

Based on crop yields and pumping reductions, we found that the SD-6 LEMA generated an overall 5-year net benefit of $1,815,130 across the region compared to the BAU scenario (Table 4.1). By combining annual national commodity prices with crop-specific yields, we found that total crop gross income for the LEMA ranged from $14.7 million in 2015 to $21.2 million in 2013. Over the 5-year LEMA period, the irrigation regime under the LEMA program resulted in a $604,380 loss in gross income compared to the BAU irrigation regime. This was a 0.75% reduction in crop income for irrigated fields. Differences were worse in dry years (2013: $194,144) and smallest in wet years ($18,019 in 2017, Table 4.1), likely mirroring differences in plant transpiration (see section 3.2.2). Estimated energy costs based on pumping volume, energy needs to lift water from the ground, and Kansas energy prices ranged from $1.33 million in 2017 to $1.58 million in 2016 for LEMA. Compared to the BAU scenario, the LEMA saved a total of $2.42 million in pumping costs over 2013-2017, a 25% reduction in energy costs for the 5 years, scaling linearly from the reduced water volume.

Table 4.1. Estimated gross crop income and groundwater pumping costs in the Sheridan-6 Local Enhanced Management Area (LEMA). Absolute estimated amounts are given for the LEMA scenario, as well as changes from business-as-usual. Amounts are in U.S. dollars, adjusted to 2017 dollars.

Year	LEMA Crop Income ($)	Δ Crop Income ($)	LEMA Pumping Costs ($)	Δ Pumping costs ($)
2013	21,183,697	-194,144	1,554,309	425,962
2014	15,003,266	-179,193	1,373,225	459,504
2015	14,737,243	-98,098	1,440,761	423,456
2016	14,793,577	-114,926	1,576,115	660,199
2017	15,004,331	-18,019	1,327,595	450,391
Total		**-604,380**		**2,419,511**

Combined, these calculations indicate that the LEMA program provided a substantial net profit for SD-6 farmers. Our estimate of lost crop income is conservative, since we held annual crop types and irrigation extents constant between the BAU and LEMA scenario. In reality,

under BAU conditions producers would have irrigated approximately 2-4% more crop area as well as planted more corn [Chapter 3, this volume], presumably resulting in a higher total crop income for the SD-6 region. Still, even total losses are likely fully recovered by the $2.42 million saved in energy costs.

4. Conclusions

We used the process-based SALUS crop model to assess water and economic sustainability from stakeholder driven agricultural management in the Sheridan-6 Local Enhanced Management Area (SD-6 LEMA). Our approach leveraged annual agricultural land use maps derived from remote sensing to simulate crop yields and the agricultural water budget within the SD-6 region from 2008-2017. With this approach, we estimated that groundwater extraction volumes decreased by ~25% (39.6 million m^3) due to reductions in irrigation application depths and frequency, which is consistent with previous statistical modeling efforts [Chapter 3]. Critically, however, SALUS was able to translate this reduced pumping volume into corresponding changes in simulated recharge due to diminished irrigation return flow when irrigation becomes more efficient. We estimated that the resulting decreases in irrigation volume reduced irrigation return flow by 27.3%. Combined with reduced pumping due to LEMA restrictions, the net aquifer water storage increased 4.68 million m^3 for the 2013-2017 LEMA, or 11.8% of estimated water use reductions.

Our results suggest that the SD-6 LEMA program improves both the economic and hydrologic sustainability of the region, increasing net profits while improving the aquifer water balance compared to business-as-usual conditions. SALUS's ability to model yield at sub-field resolution across SD-6 allowed us to assess how the LEMA program affected regional income. Overall, we estimated that changes in irrigation behavior in the LEMA scenario substantially

increased net profits by $1.8 million when both yield penalties and energy savings were included. Furthermore, extending aquifer lifespan can generate long-term benefits to the system given expected future higher yielding varieties and preserving the ability to mitigate drought [*Zipper et al.*, 2016; *Foster et al.*, 2017; *Quintana Ashwell et al.*, 2018].

However, it remains unclear if LEMA-era levels of irrigated agriculture in the SD-6 region are fully sustainable. Our analysis of net aquifer change indicated that the aquifer balance was positive only in abnormally wet years (Figure 4.6). The initial LEMA cycle from 2013-2017 was 26.7% wetter than the 2002-2012 period upon which reduction targets are based [Chapter 3, this volume]. This indicates that any realized benefits from the first LEMA cycle may not apply under typical to drought conditions, particularly considering the relatively small net gain to the aquifer balance during this wetter than average period. Given the region's intermediate drought frequency and expected increases in water stress due to climate change [*Dai*, 2013], LEMA effectiveness and yield implications in drought conditions need to be better understood to inform future planning efforts. Future work is planned to extend the modeling approach developed here to include drought scenarios, thus testing if the SD-6 LEMA framework is sufficient to avoid aquifer depletion under drought stress. As aquifer depletion threatens crop production in many parts of the world, approaches that integrate models with in-situ and remotely sensed data can estimate critical system components that are difficult to directly measure, thus informing economically and hydrologically sustainable management strategies.

Acknowledgements

This chapter was co-authored by Anthony Kendall, Bruno Basso, and David Hyndman. We thank Brian Baer, Lydia Rill, Alexandria Kuhl, and Kayla Cotterman for SALUS technical support. Funding for this work was provided by NSF grant WSC 1039180, USDA NIFA grant

2015-68007-23133 and USDA-NIFA/NSF INFEWS program grant 2018-67003-27406. Jillian Deines was partially supported by NASA Headquarters under the NASA Earth and Space Science Fellowship Program grant 14-EARTH14F-198. This work was also supported in part by Michigan State University through computational resources provided by the Institute for Cyber-Enabled Research. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of NSF, USDA NIFA, NASA, or MSU.

REFERENCES

REFERENCES

Abatzoglou, J. T. (2013), Development of gridded surface meteorological data for ecological applications and modelling, *Int. J. Climatol.*, *33*(1), 121–131, doi:10.1002/joc.3413.

Abdullah, K. Bin (2006), Use of water and land for food security and environmental sustainability, *Irrig. Drain.*, *55*(3), 219–222, doi: 10.1002/Ird.254.

Abuzar, M., A. McAllister, and D. Whitfield (2015), Mapping irrigated farmlands using vegetation and thermal thresholds derived from Landsat and ASTER data in an irrigation district of Australia, *Photogramm. Eng. Remote Sensing*, *81*(3), 229–238, doi:10.14358/PERS.81.3.229.

Aeschbach-Hertig, W., and T. Gleeson (2012), Regional strategies for the accelerating global problem of groundwater depletion, *Nat. Geosci.*, *5*(12), 853–861, doi:10.1038/ngeo1617.

Aleksandrova, M., J. P. A. Lamers, C. Martius, and B. Tischbein (2014), Rural vulnerability to environmental change in the irrigated lowlands of Central Asia and options for policy-makers: A review, *Environ. Sci. Policy*, *41*, 77–88, doi:10.1016/j.envsci.2014.03.001.

Araya, A., I. Kisekka, P. H. Gowda, and P. V. V. Prasad (2018), Grain sorghum production functions under different irrigation capacities, *Agric. Water Manag.*, *203*(March), 261–271, doi:10.1016/j.agwat.2018.03.010.

Arvidson, T., J. Gasch, and S. N. Goward (2001), Landsat 7's long-term acquisition plan - An innovative approach to building a global imagery archive, *Remote Sens. Environ.*, *78*(1–2), 13–26, doi:10.1016/S0034-4257(01)00263-2.

Ashworth, W. (2006), *Ogallala Blue: Water and Life on the Great Plains*, First Edit., W.W. Norton & Company.

Azzari, G., and D. B. Lobell (2017), Landsat-based classification in the cloud: An opportunity for a paradigm shift in land cover monitoring, *Remote Sens. Environ.*, *202*, 64–74, doi:10.1016/j.rse.2017.05.025.

Babbitt, C., K. Gibson, S. Sellers, N. Brozović, A. Saracino, A. Hayden, M. Hall, and S. Zellmer (2018), The future of groundwater in California: Lessons in sustainable management from across the West, Environmental Defense Fund and Daugherty Water for Food Institute, 122.

Baskerville, A. G. L., and P. Emin (1969), Rapid estimation of heat accumulation from maximum and minimum temperatures, *Ecology*, *50*(3), 514–517.

Basso, B., and J. T. Ritchie (2012), Assessing the Impact of Management Strategies on Water Use Efficiency Using Soil–Plant–Atmosphere Models, *Vadose Zo. J.*, *11*(3), 1–8, doi:10.2136/vzj2011.0173.

Basso, B., J. T. Ritchie, P. R. Grace, and L. Sartori (2006), Simulation of tillage systems impact on soil biophysical properties using the SALUS model, *Ital. J. Agron.*, *4*, 677–688, doi:10.4081/ija.2006.677.

Basso, B., A. D. Kendall, and D. W. Hyndman (2013), The future of agriculture over the Ogallala Aquifer: Solutions to grow crops more efficiently with limited water, *Earth's Futur.*, *1*(1), 39–41, doi:10.1002/2013EF000107.

Basso, B., D. W. Hyndman, A. D. Kendall, G. Robertson, and R. Grace (2015), Can impacts of climate change and agricultural adaption strategies be accurately quantified if crop models are annually reinitialized?, *PLoS One*, *10*, e0127333, doi:https://doi.org/10.1371/journal.pone.0127333.

Bertrand, M., E. Duflo, and S. Mullainathan (2002), *How much should we trust differences-in-differences estimates?*, Cambridge, MA.

Boryan, C., Z. Yang, R. Mueller, and M. Craig (2011), Monitoring US agriculture: The US department of agriculture, national agricultural statistics service, cropland data layer program, *Geocarto Int.*, *26*(5), 341–358, doi:10.1080/10106049.2011.562309.

Brauman, K. A., S. Siebert, and J. A. Foley (2013), Improvements in crop water productivity increase water sustainability and food security—a global analysis, *Environ. Res. Lett.*, *8*(2), 24030, doi:10.1088/1748-9326/8/2/024030.

Breiman, L. (2001), Random forests, *Mach. Learn.*, *45*(1), 5–32, doi:10.1023/A:1010933404324.

Breña-Naranjo, J. A., A. D. Kendall, and D. W. Hyndman (2014), Improved methods for satellite-Based groundwater storage estimates: A decade of monitoring the high plains aquifer from space and ground observations, *Geophys. Res. Lett.*, *41*(17), 6167–6173, doi:10.1002/2014GL061213.

Brodersen, K. H., F. Gallusser, J. Koehler, N. Remy, and S. L. Scott (2015), Inferring causal impact using bayesian structural time-series models, *Ann. Appl. Stat.*, *9*(1), 247–274, doi:10.1214/14-AOAS788.

Brown, J. F., and M. S. Pervez (2014), Merging remote sensing data and national agricultural statistics to model change in irrigated agriculture, *Agric. Syst.*, *127*, 28–40, doi:10.1016/j.agsy.2014.01.004.

Butler, J. J., D. O. Whittemore, B. B. Wilson, and G. C. Bohling (2016), A new approach for assessing the future of aquifers supporting irrigated agriculture, *Geophys. Res. Lett.*, *43*, 2004–2010, doi:10.1002/2016GL067879.Received.

Butler, J. J., D. O. Whittemore, G. C. Bohling, and B. B. Wilson (2018), Sustainability of aquifers supporting irrigated agriculture: A case study of the High Plains Aquifer in Kansas, *Water Int.*, *In Press*.

Cao, G., C. Zheng, B. R. Scanlon, J. Liu, and W. Li (2013), Use of flow modeling to assess

sustainability of groundwater resources in the North China Plain, *Water Resour. Res.*, *49*(1), 159–175, doi:10.1029/2012WR011899.

Cardille, J. A., and J. A. Fortin (2016), Bayesian updating of land-cover estimates in a data-rich environment, *Remote Sens. Environ.*, *186*, 234–249, doi:10.1016/j.rse.2016.08.021.

Chai, Q., Y. Gan, C. Zhao, H. L. Xu, R. M. Waskom, Y. Niu, and K. H. M. Siddique (2016), Regulated deficit irrigation for crop production under drought stress. A review, *Agron. Sustain. Dev.*, *36*(1), 1–21, doi:10.1007/s13593-015-0338-6.

Conrad, C., S. Schönbrodt-Stitt, F. Löw, D. Sorokin, and H. Paeth (2016), Cropping intensity in the Aral Sea Basin and its dependency from the runoffformation 2000-2012, *Remote Sens.*, *8*(8), doi:10.3390/rs8080630.

Cotterman, K. A., A. D. Kendall, B. Basso, and D. W. Hyndman (2018), Groundwater depletion and climate change: future prospects of crop production in the Central High Plains Aquifer, *Clim. Change*, *146*, 187–200, doi:10.1007/s10584-017-1947-7.

Dai, A. (2013), Increasing drought under global warming in observations and models, *Nat. Clim. Chang.*, *3*(1), 52–58, doi:10.1038/nclimate1633.

Dappen, P. (2003), Using satellite imagery to estimate irrigated land: A case study in Scotts Bluff and Kearney counties, summer 2002, Center for Advanced Land Management Information Technologies, Lincoln, NE, 1–38.

Dappen, P., and J. Merchant (2003), Delineation of 2001 land use patterns for the Cooperative Hydrology Study in the Central Platte River Basin, Center for Advanced Land Management Information Technologies, Lincoln, NE, 1–83.

Dappen, P., and J. Merchant (2004), Delineation of 1982 land use patterns for the Cooperative Hydrology Study in the central Platte River Basin, Center for Advanced Land Management Information Technologies, Lincoln, NE, 1–85.

Dappen, P., and M. Tooze (2001), Delineation of land use patterns for the Cooperative Hydrology Study in the Central Platte River Basin, Center for Advanced Land Management Information Technologies, Lincoln, NE, 1–73.

Dappen, P., J. Merchernt, I. Ratcliffe, and C. Robbins (2007), Delineation of 2005 land use patterns for the state of Nebraska Department of Natural Resources, Center for Advanced Land Management Information Technologies, Lincoln, NE, 1–80.

Deines, J. M., A. D. Kendall, and D. W. Hyndman (2017), Annual irrigation dynamics in the U.S. Northern High Plains derived from Landsat satellite data, *Geophys. Res. Lett.*, *44*(18), 9350–9360, doi:10.1002/2017GL074071.

Dennehy, K. F. (2000), *High Plains regional ground-water study*. USGS Fact Sheet FS-091-00.

Devineni, N., U. Lall, E. Etienne, D. Shi, and C. Xi (2015), America's water risk: Current

demand and climate variability, *Geophys. Res. Lett.*, *42*(7), 2285–2293, doi:10.1002/2015GL063487.

Doll, P. (2009), Vulnerability to the impact of climate change on renewable groundwater resources: A global-scale assessment, *Environ. Res. Lett.*, *4*(3), doi:10.1088/1748-9326/4/3/035006.

Döll, P. (2002), Impact of Climate Change and Variability on Irrigation Requirements: A Global Perspective, *Clim. Change*, *54*, 269–293, doi:10.1023/A:1016124032231.

Dzotsi, K. A., B. Basso, and J. W. Jones (2015), Parameter and uncertainty estimation for maize, peanut and cotton using the SALUS crop model, *Agric. Syst.*, *135*, 31–47, doi:10.1016/j.agsy.2014.12.003.

Elliott, J. et al. (2014), Constraints and potentials of future irrigation water availability on agricultural production under climate change, *Proc. Natl. Acad. Sci.*, *111*(9), 3239–3244, doi:10.1073/pnas.1222474110.

Faunt, C. C., M. Sneed, J. Traum, and J. T. Brandt (2016), Water availability and land subsidence in the Central Valley, California, USA, *Hydrogeol. J.*, *24*(3), 675–684, doi:10.1007/s10040-015-1339-x.

Fileccia, A. (2016), Some simple procedures for the calculation of the influence radius and well head protection areas (theoretical approach and a field case for a water table aquifer in an alluvial plain), *Acque Sotter. - Ital. J. Groundw.*, *4*(3), 7–23, doi:10.7343/as-117-15-0144.

Foster, T., N. Brozović, and A. Butler (2014), Modeling irrigation behavior in groundwater systems, *Water Resour. Res.*, 6370–6389, doi:10.1002/2014WR015620.Received.

Foster, T., N. Brozović, and A. P. Butler (2017), Effects of initial aquifer conditions on economic benefits from groundwater conservation, *Water Resour. Res.*, *53*, 744–762, doi:10.1002/2016WR019365.Received.

Fry, J. A., G. Xian, S. Jin, J. A. Dewitz, C. G. Homer, Y. LIMIN, C. A. Barnes, N. D. Herold, and J. D. Wickham (2011), Completion of the 2006 national land cover database for the conterminous United States, *Photogramm. Eng. Remote Sensing*, *77*(9), 858–864.

Gao, B. C. (1996), NDWI - A normalized difference water index for remote sensing of vegetation liquid water from space, *Remote Sens. Environ.*, *58*(3), 257–266, doi:10.1016/S0034-4257(96)00067-3.

Gartner, P. (2018), This Landsat project would cost..., accessed at https://philippgaertner.github.io/2018/06/landsat-cost-estimator/ on July 1, 2018.

Gitelson, A. A., A. Viña, V. Ciganda, D. C. Rundquist, and T. J. Arkebauer (2005), Remote estimation of canopy chlorophyll content in crops, *Geophys. Res. Lett.*, *32*(8), 1–4, doi:10.1029/2005GL022688.

Gleeson, T., W. M. Alley, D. M. Allen, M. a Sophocleous, Y. Zhou, M. Taniguchi, and J. VanderSteen (2012a), Towards sustainable groundwater use: setting long-term goals, backcasting, and managing adaptively., *Ground Water*, *50*(1), 19–26, doi:10.1111/j.1745-6584.2011.00825.x.

Gleeson, T., Y. Wada, M. F. P. Bierkens, and L. P. H. van Beek (2012b), Water balance of global aquifers revealed by groundwater footprint., *Nature*, *488*(7410), 197–200, doi:10.1038/nature11295.

Gleick, P. H. (2003), Global freshwater resources: soft-path solutions for the 21st century, *Science*, *302*(5650), 1524–1528, doi:10.1126/science.1089967.

Golden, B., and K. Liebsch (2017), *Monitoring the Impacts of Sheridan County 6 Local Enhanced Management Area: Interim Report for 2013 – 2016*, Kansas State University, Manhattan, KS.

Gorelick, N., M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore (2017), Google Earth Engine : Planetary-scale geospatial analysis for everyone, *Remote Sens. Environ.*, *202*, 18–27, doi:10.1016/j.rse.2017.06.031.

Gräler, B., and E. Pebesma (2016), Spatio-temporal geostatistics using gstat, *R J.*, *8*, 204–218, doi:10.1007/978-3-319-17885-1.

Griggs, B. W. (2017), The political cultures of irrigation and the proxy battles of interstate water litigation, *Nat. Resour. J.*, *57*(1), 1–73.

Gutman, G., C. Huang, G. Chander, P. Noojipady, and J. G. Masek (2013), Assessment of the NASA-USGS Global Land Survey (GLS) datasets, *Remote Sens. Environ.*, *134*, 249–265, doi:10.1016/j.rse.2013.02.026.

Haacker, E. M. K., A. D. Kendall, and D. W. Hyndman (2016), Water level declines in the High Plains Aquifer: Predevelopment to resource senescence, *Groundwater*, *54*(2), 231–242, doi:10.1111/gwat.12350.

Häni, F., F. Braga, A. Stämpfli, T. Keller, M. Fischer, and H. Porsche (2003), RISE, a tool for holistic sustainability assessment at the farm level, *Int. Food Agribus. Manag. Rev.*, *6*(4).

Hansen, C. V (1991), *Estimates of freshwater storage and potential natural recharge for principal aquifers in Kansas*, U.S. Geological Survey, Reston, VA.

Hansen, M. C. et al. (2013), High-resolution global maps of 21st-century forest cover change, *Science (80-.).*, *342*(6160), 850–853, doi:10.1126/science.1244693.

Hardin, G. (1968), The tragedy of the commons, *Science (80-.).*, *162*, 1243–1248.

Hellerstein, D. M. (2017), The US Conservation Reserve Program: The evolution of an enrollment mechanism, *Land use policy*, *63*, 601–610, doi:10.1016/j.landusepol.2015.07.017.

Hermosilla, T., M. A. Wulder, J. C. White, N. C. Coops, and G. W. Hobart (2015), An integrated Landsat time series protocol for change detection and generation of annual gap-free surface reflectance composites, *Remote Sens. Environ.*, *158*, 220–234, doi:10.1016/j.rse.2014.11.005.

Hoang, T. Van, T. Van Hoang, T. Y. Chou, B. Basso, M. L. Yeh, and C. Y. Chien (2014), Climate change impact on agricultural productivity and environment influence based on simulation model, *Int. J. Adv. Remote Sens. GIS*, *3*(1), 642–659.

Hothorn, T., K. Hornik, and A. Zeileis (2006), Unbiased recursive partitioning: A conditional inference framework., *J. Comput. Graph. Stat.*, *15*(3), 651–674.

Huete, A., K. Didan, T. Miura, E. P. Rodriguez, X. Gao, and L. G. Ferreira (2002), Overview of the radiometric and biophysical performance of the MODIS vegetation indices, *Remote Sens. Environ.*, *83*, 195–213, doi:10.1016/S0034-4257(02)00096-2.

Hutchinson, W. (2010), The use of groundwater availability models in Texas in the establishment of desired future conditions, in *GSA Annual Meeting*, Denver.

K.S.A. 82a-1041 (2012), *Local enhanced management areas; establishment procedures; duties of chief engineer; hearing; notice; orders; review*, Kansas Statutes Annotated.

KDA (2013), Order of designation approving the Sheridan 6 Local Enhanced Management Area within Groundwater Management District No. 4, Kansas Department of Agriculture, 60.

KDA DWR (2017), Water information management & analysis system (WIMAS), Kansas Department of Agriculture, Division of Water Resources, Accessed at http://hercules.kgs.ku.edu/geohydro/wimas/query_setup.cfm on 9/27/2017.

KGS (2018), Water Information Storage and Retrieval Database (WIZARD), Kansas Geological Survey. Accessed at http://www.kgs.ku.edu/Magellan/WaterLevels/index.html.

Kranz, W. L., S. Irmak, S. J. van Donk, C. D. Yonts, and D. L. Martin (2008), *Irrigation Management for Corn*, University of Nebraska-Lincoln Extension, Lincoln, NE.

KS DWR (2017), Kansas authorized irrigation place of use tracts, Kansas Department of Agriculture, Division of Water Resources, Accessed on 11/02/17.

Kustu, M. D., Y. Fan, and A. Robock (2010), Large-scale water cycle perturbation due to irrigation pumping in the US High Plains: A synthesis of observed streamflow changes, *J. Hydrol.*, *390*(3–4), 222–244, doi:10.1016/j.jhydrol.2010.06.045.

Kuwayama, Y., and N. Brozović (2013), The regulation of a spatially heterogeneous externality: Tradable groundwater permits to protect streams, *J. Environ. Econ. Manage.*, *66*(2), 364–382, doi:10.1016/j.jeem.2013.02.004.

Lark, T. J., J. M. Salmon, and H. K. Gibbs (2015), Cropland expansion outpaces agricultural and biofuel policies in the United States, *Environ. Res. Lett.*, *10*, 44003, doi:10.1088/1748-

9326/10/4/044003.

Lauer, S., and M. Sanderson (2017), Managing groundwater together in western Kansas, *Color. Water*, November/December 16-17.

Liaw, A., and M. Wiener (2002), Classification and regression by randomForest, *R News*, *2*(3), 18–22.

Lobell, D. B., G. Bala, and P. B. Duffy (2006), Biogeophysical impacts of cropland management changes on climate, *Geophys. Res. Lett.*, *33*(6), 4–7, doi:10.1029/2005GL025492.

Lobell, D. B., K. G. Cassman, and C. B. Field (2009), Crop Yield Gaps: Their Importance, Magnitudes, and Causes, *Annu. Rev. Environ. Resour.*, *34*(1), 179–204, doi:10.1146/annurev.environ.041008.093740.

Luckey, R. R., and M. Becker (1999), *Hydrology, water use, and simulation of flow in the High Plains Aquifer in northwesten Oklahoma, southeastern Colorado, southwestern Kansas, northeastern New Mexico, and northwestern Texas*, U.S. Geological Survey, Reston, VA.

Luckey, R. R., E. D. Gutentag, and J. B. Weeks (1981), *Water-level and saturated thickness changes, predevelopment to 1980, in the High Plains Aquifer in parts of Colorado, Kansas, Nebraska, New Mexico, Oklahoma, South Dakota, Texas, and Wyoming*, HA-652, U.S. Geological Survey, Reston, VA.

Manning, D. T., R. Rockel, J. Schneekloth, A. Stoecker, and J. Warren (2018), Crop insurance, *Ogallala Aquifer Summit White Pap.*

Marsalis, M., T. Blaine, and R. Ghimire (2018), *New Mexico*, Ogallala Aquifer Summit White Papers.

McCarthy, B., A. D. Kendall, R. Anex, A. Anctil, and D. W. Hyndman (n.d.), The dynamic response of energy, water, and carbon emissions from pumping to irrigation technology change over the High Plains Aquifer, *In Prep.*

Mcguire, V. L. (2017), *Water-level and recoverable water in storage changes, High Plains aquifer, predevelopment to 2015 and 2013-15*, U.S. Geological Survey.

NAIP (2017), USDA Farm Service Agency National Agriculture Imagery Program, *USDA Farm Serv. Agency*, 2003–2017.

NASS (2017), Quick Stats API, *USDA Natl. Agric. Stat. Serv.* Available from: https://quickstats.nass.usda.gov/api (Accessed 1 January 2017)

NRCS (2016), SSURGO Web Soil Survey, *USDA Nat. Recources Conserv. Serv.*

NW KS GMD 4 (2016), Revised management program, Northwest Kansas Groundwater Management District.

NW KS GMD 4 (2017), The Water Table, Northwest Kansas Groundwater Management District, (Fall 2017), 1–32.

Ostrom, E. (2009a), A general framework for analyzing sustainability of social-ecological systems, *Science (80-.).*, *325*(5939), 419–422, doi:10.1126/science.1172133.

Ostrom, E. (2009b), A general framework for analyzing sustainability of social-ecological systems., *Science*, *325*(5939), 419–22, doi:10.1126/science.1172133.

Ostrom, E., R. Gardner, and J. Walker (1994), *Rules, games, and common-pool resources*, University of Michigan, Ann Arbor.

Ozdogan, M., and G. Gutman (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, doi:10.1016/j.rse.2008.04.010.

Ozdogan, M., C. E. Woodcock, G. D. Salvucci, and H. Demir (2006), Changes in summer irrigated crop area and water use in Southeastern Turkey from 1993 to 2002: Implications for current and future water resources, *Water Resour. Manag.*, *20*(3), 467–488, doi:10.1007/s11269-006-3087-0.

Ozdogan, M., Y. Yang, G. Allez, and C. Cervantes (2010a), Remote sensing of irrigated agriculture: Opportunities and challenges, *Remote Sens.*, *2*(9), 2274–2304, doi:10.3390/rs2092274.

Ozdogan, M., M. Rodell, H. K. Beaudoing, and D. L. Toll (2010b), Simulating the Effects of Irrigation over the United States in a Land Surface Model Based on Satellite-Derived Agricultural Data, *J. Hydrometeorol.*, *11*, 171–184, doi:10.1175/2009JHM1116.1.

Pastick, N. J., B. K. Wylie, and Z. Wu (2018), Spatiotemporal analysis of Landsat-8 and Sentinel-2 data to support monitoring of dryland ecosystems, *Remote Sens.*, *10*(5), 1–15, doi:10.3390/rs10050791.

Pebesma, E. J. (2004), Multivariable geostatistics in S: The gstat package, *Comput. Geosci.*, *30*(7), 683–691, doi:10.1016/j.cageo.2004.03.012.

Peck, J. C. (2003), Property rights in groundwater: Some lessons from the Kansas experience, *Kansas J. Law Public Policy*, *12*(3), 493–520.

Peck, J. C. (2007), Groundwater management in the High Plains Aquifer in the USA: Legal problems and innovations, in *The Agricultural Groundwater Revolution: Opportunities and Threats to Development*, edited by M. Giordano and K. G. Villholth, pp. 296–319, CABI.

Peel, M., B. Finlayson, and T. McMahon (2007), Updated world map of the Köppen-Geiger climate classification, *Hydrol. Earth Syst. Sci.*, *11*, 1633–1644.

Pei, L. et al. (2016), Effects of Irrigation on Summer Precipitation over the United States, *J. Clim.*, *29*(10), 3541–3558, doi:10.1175/JCLI-D-15-0337.1.

Pekel, J.-F., A. Cottam, N. Gorelick, and A. S. Belward (2016), High-resolution mapping of global surface water and its long-term changes, *Nature*, *540*(7633), 1–19, doi:10.1038/nature20584.

Peña-Arancibia, J. L., T. R. McVicar, Z. Paydar, L. Li, J. P. Guerschman, R. J. Donohue, D. Dutta, G. M. Podger, A. I. J. M. van Dijk, and F. H. S. Chiew (2014), Dynamic identification of summer cropping irrigated areas in a large basin experiencing extreme climatic variability, *Remote Sens. Environ.*, *154*, 139–152, doi:10.1016/j.rse.2014.08.016.

Pervez, M. S., and J. F. Brown (2010), Mapping irrigated lands at 250-m scale by merging MODIS data and National Agricultural Statistics, *Remote Sens.*, *2*(10), 2388–2412, doi:10.3390/rs2102388.

Pervez, M. S., M. Budde, and J. Rowland (2014), Mapping irrigated areas in Afghanistan over the past decade using MODIS NDVI, *Remote Sens. Environ.*, *149*, 155–165, doi:10.1016/j.rse.2014.04.008.

Pfeiffer, L., and C. Y. C. Lin (2014a), Does efficient irrigation technology lead to reduced groundwater extraction? Empirical evidence, *J. Environ. Econ. Manage.*, *67*(2), 189–208, doi:10.1016/j.jeem.2013.12.002.

Pfeiffer, L., and C. Y. C. Lin (2014b), The effects of energy prices on agricultural groundwater extraction fromthe high plains aquifer, *Am. J. Agric. Econ.*, *96*(5), 1349–1362, doi:10.1093/ajae/aau020.

Postel, S. L. (2003), Securing water for people, crops, and ecosystems: New mindset and new priorities, *Nat. Resour. Forum*, *27*(2), 89–98, doi:10.1111/1477-8947.00044.

Qi, B. S. L., A. Konduris, D. W. Litke, and J. Dupree (2002), Classification of irrigated land using satellite imagery, the High Plains aquifer, nominal date 1992, U.S. Geological Survey Water Resources Investigation Report.

Quintana Ashwell, N. E., J. M. Peterson, and N. P. Hendricks (2018), Optimal groundwater management under climate change and technical progress ☪, *Resour. Energy Econ.*, *51*, 67–83, doi:10.1016/j.reseneeco.2017.10.005.

R Core Team (2014), *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria.

Ritchie, J. T., and S. Otter (1985), *Description and performance of CERES-Wheat: a user-oriented wheat yield model*.

Robinson, N. P., B. W. Allred, M. O. Jones, A. Moreno, J. S. Kimball, D. E. Naugle, T. A. Erickson, and A. D. Richardson (2017), A dynamic landsat derived normalized difference vegetation index (NDVI) product for the conterminous United States, *Remote Sens.*, *9*(8), 1–14, doi:10.3390/rs9080863.

Rockström, J., M. Falkenmark, M. Lannerstad, and L. Karlberg (2012), The planetary water

drama: Dual task of feeding humanity and curbing climate change, *Geophys. Res. Lett.*, *39*(15), 1–8, doi:10.1029/2012GL051688.

Rodell, M., I. Velicogna, and J. S. Famiglietti (2009), Satellite-based estimates of groundwater depletion in India, *Nature*, *460*, 999–1002, doi:10.1038/nature08238.

Rodell, M., J. S. Famiglietti, D. N. Wiese, J. T. Reager, H. K. Beaudoing, F. W. Landerer, and M.-H. Lo (2018), Emerging trends in global freshwater availability, *Nature*, doi:10.1038/s41586-018-0123-1.

Rosegrant, M. W., C. Ringler, and T. Zhu (2009), Water for agriculture: Maintaining food security under growing scarcity, *Annu. Rev. Environ. Resour.*, *34*(1), 205–222, doi:10.1146/annurev.environ.030308.090351.

RRCA (2003), *Republican River Compact Administration Ground Water Model*. Republican River Compact Administration. 76 pp.

Rufin, P., C. Levers, M. Baumann, J. Jägermeyr, T. Krueger, T. Kuemmerle, and P. Hostert (2018), Global-scale patterns and determinants of cropping frequency in irrigation dam command areas, *Glob. Environ. Chang.*, *50*(January), 110–122, doi:10.1016/j.gloenvcha.2018.02.011.

Scanlon, B. R., R. C. Reedy, D. a. Stonestrom, D. E. Prudic, and K. F. Dennehy (2005), Impact of land use and land cover change on groundwater recharge and quality in the southwestern US, *Glob. Chang. Biol.*, *11*(10), 1577–1593, doi:10.1111/j.1365-2486.2005.01026.x.

Scanlon, B. R., C. C. Faunt, L. Longuevergne, R. C. Reedy, W. M. Alley, V. L. Mcguire, and P. B. Mcmahon (2012), Groundwater depletion and sustainability of irrigation in the US High Plains and Central Valley, *Proc. Natl. Acad. Sci.*, *109*, 9320–9325, doi:10.1073/pnas.1200311109/-/DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1200311109.

Shiklomanov, I. A. (2000), Appraisal and Assessment of world water resources, *Water Int.*, *25*(1), 11–32, doi:10.1080/02508060008686794.

Shroyer, J. P., C. Thompson, R. Brown, P. D. Ohlenbusch, D. L. Fjell, S. Staggenborg, S. Duncan, and G. L. Kilgore (1996), *Kansas Crop Planting Guide*, Manhattan, KS.

Siebert, S., J. Burke, J. M. Faures, K. Frenken, J. Hoogeveen, P. Döll, and F. T. Portmann (2010), Groundwater use for irrigation - A global inventory, *Hydrol. Earth Syst. Sci.*, *14*(10), 1863–1880, doi:10.5194/hess-14-1863-2010.

Smidt, S. J., E. M. K. Haacker, A. D. Kendall, J. M. Deines, L. Pei, K. a. Cotterman, H. Li, X. Liu, B. Basso, and D. W. Hyndman (2016), Complex water management in modern agriculture: Trends in the water energy-food nexus over the High Plains Aquifer, *Agric. Water Manag.*, *566–567*, 988–1001, doi:10.1017/CBO9781107415324.004.

Stanton, J. S., S. L. Qi, D. W. Ryter, S. E. Falk, N. a. Houston, S. M. Peterson, S. M.

Westenbroek, and S. C. Christenson (2011), Selected approaches to estimate water-budget components of the High Plains, 1940 through 1949 and 2000 through 2009, , 92.

Streeter, T., S. Metzger, C. Beightel, S. Stover, J. Aguilar, and B. B. Golden (2018), Kansas, *Ogallala Aquifer Summit White Pap.*

Teluguntla, P., P. S. Thenkabail, J. Xiong, M. K. Gumma, R. G. Congalton, A. Oliphant, J. Poehnelt, K. Yadav, M. Rao, and R. Massey (2017), Spectral matching techniques (SMTs) and automated cropland classification algorithms (ACCAs) for mapping croplands of Australia using MODIS 250-m time-series (2000–2015) data, *Int. J. Digit. Earth*, *8947*(January), 1–34, doi:10.1080/17538947.2016.1267269.

Thenkabail, P. S., and Z. Wu (2012), An automated cropland classification algorithm (ACCA) for Tajikistan by combining landsat, MODIS, and secondary data, *Remote Sens.*, *4*(10), 2890–2918, doi:10.3390/rs4102890.

Tilman, D., C. Balzer, J. Hill, and B. L. Befort (2011), Global food demand and the sustainable intensification of agriculture, *Proc. Natl. Acad. Sci.*, *108*(50), 20260–20264, doi:10.1073/pnas.1116437108.

Tiwari, V. M., J. Wahr, and S. Swenson (2009), Dwindling groundwater resources in northern India, from satellite gravity observations, *Geophys. Res. Lett.*, *36*(18), 1–5, doi:10.1029/2009GL039401.

Tringali, C., V. Re, G. Siciliano, N. Chkir, C. Tuci, and K. Zouari (2017), Insights and participatory actions driven by a socio-hydrogeological approach for groundwater management: the Grombalia Basin case study (Tunisia), *Hydrogeol. J.*, *25*(5), 1241–1255, doi:10.1007/s10040-017-1542-z.

Troy, T. J., C. Kipgen, and I. Pal (2015), The impact of climate extremes and irrigation on US crop yields, *Environ. Res. Lett.*, *10*(5), doi:10.1088/1748-9326/10/5/054013.

USDA-NASS (2017), USDA National Agricultural Statistics Service Cropland Data Layers, , 2002–2016, U.S Departmernt of Agriculture.

USDA-NASS-RRD Spatial Analysis Research Section (2017), Cropland data layer metadata, Available from: https://www.nass.usda.gov/Research_and_Science/Cropland/metadata/meta.php

USGS (2012), National Elevation Dataset, *U.S. Geol. Surv.*

USGS (2015), Water use data for the nation, *U.S. Geol. Surv.*, 2000–2010.

USGS (2017a), *Landsat Collection 1 Level 1 Product Definition Version 1.0*.

USGS (2017b), Product guide: Landsat 4-7 climate data record (CDR) surface reflectance, *U.S. Geol. Surv.*, (Version 7.5), 36.

USGS (2017c), Product guide: Landsat 8 surface reflectance code (LaSRC) Product, *U.S. Geol. Surv.*, (Version 3.7), 33.

Velpuri, N. M., P. S. Thenkabail, M. K. Gumma, C. Biradar, V. Dheeravath, P. Noojipady, and L. Yuanjie (2009), Influence of resolution in irrigated area mapping and area estimation, *Photogramm. Eng. Remote Sens.*, *75*(12), 1383–1395, doi:10.14358/PERS.75.12.1383.

Vogelmann, J. E., A. L. Gallant, H. Shi, and Z. Zhu (2015), Perspectives on monitoring gradual change across the continuity of Landsat sensors using time-series data, *Remote Sens. Environ.*, (March), doi:10.1016/j.rse.2016.02.060.

Wada, Y., and L. Heinrich (2013), Assessment of transboundary aquifers of the world—vulnerability arising from human water use, *Environ. Res. Lett.*, *8*(2), 24003, doi:10.1088/1748-9326/8/2/024003.

Wada, Y., L. P. H. Van Beek, and M. F. P. Bierkens (2011), Modelling global water stress of the recent past: On the relative importance of trends in water demand and climate variability, *Hydrol. Earth Syst. Sci.*, *15*(12), 3785–3808, doi:10.5194/hess-15-3785-2011.

Wada, Y., L. P. H. van Beek, and M. F. P. Bierkens (2012), Nonsustainable groundwater sustaining irrigation: A global assessment, *Water Resour. Res.*, *48*(November 2011), W00L06, doi:10.1029/2011WR010562.

Wada, Y. et al. (2013), Multimodel projections and uncertainties of irrigation water demand under climate change, *Geophys. Res. Lett.*, *40*(17), 4626–4632, doi:10.1002/grl.50686.

Wada, Y. et al. (2016), Modeling global water use for the 21st century: The Water Futures and Solutions (WFaS) initiative and its approaches, *Geosci. Model Dev.*, *9*(1), 175–222, doi:10.5194/gmd-9-175-2016.

Wang, T., S. C. Park, and H. Jin (2015), Will farmers save water ? A theoretical analysis of groundwater conservation policies, *Water Resour. Econ.*, *12*, 27–39, doi:10.1016/j.wre.2015.10.002.

Wardlow, B. D., and S. L. Egbert (2008), Large-area crop mapping using time-series MODIS 250 m NDVI data: An assessment for the U.S. Central Great Plains, *Remote Sens. Environ.*, *112*(3), 1096–1116, doi:10.1016/j.rse.2007.07.019.

Waskom, R. (2017), The Ogallala Aquifer: Director's Letter, *Color. Water*, (November/December), 1.

Waskom, R., J. Pritchett, and J. Schneeklot (2006), Outlook on the High Plains aquifer: What's in store for irrigated agriculture?, in *Great Plains Soil Fertility Conference*, pp. 122–128, Denver, CO.

Weeks, J. B., E. D. Gutentag, F. J. Heimes, and R. R. Luckey (1988), *Summary of the High Plains Aquifer-System analysis in parts of Colorado, Kansas, Nebraska, New Mexico, Oklahoma, South Dakota, Texas, and Wyoming. U.S. Geological Survey Professional Paper*

1400-A.

West, C., D. Porter, B. Guerrero, V. Uddameri, J. Bordovsky, J. Bell, and J. Tracy (2018), *Texas*.

White, J. C. et al. (2014), Pixel-based image compositing for large-area dense time series applications and science, *Can. J. Remote Sens.*, *40*(3), 192–212, doi:10.1080/07038992.2014.945827.

Whittemore, D. O., J. J. Butler, and B. B. Wilson (2016), Assessing the major drivers of water-level declines: new insights into the future of heavily stressed aquifers, *Hydrol. Sci. J.*, *61*, 134–145, doi:10.1080/02626667.2014.959958.

Wisser, D., S. Frolking, E. M. Douglas, B. M. Fekete, C. J. Vörösmarty, and A. H. Schumann (2008), Global irrigation water demand: Variability and uncertainties arising from agricultural and climate data sets, *Geophys. Res. Lett.*, *35*(24), 1–5, doi:10.1029/2008GL035296.

Woodcock, C. E. et al. (2008), Free access to Landsat imagery, *Science (80-.).*, *320*, 1011, doi:10.1126/science.320.5879.1011a.

WSEO (2018), *Wyoming*.

Wulder, M. A., J. C. White, S. N. Goward, J. G. Masek, J. R. Irons, M. Herold, W. B. Cohen, T. R. Loveland, and C. E. Woodcock (2008), Landsat continuity: Issues and opportunities for land cover monitoring, *Remote Sens. Environ.*, *112*(3), 955–969, doi:10.1016/j.rse.2007.07.004.

Wulder, M. A., J. G. Masek, W. B. Cohen, T. R. Loveland, and C. E. Woodcock (2012), Opening the archive: How free data has enabled the science and monitoring promise of Landsat, *Remote Sens. Environ.*, *122*, 2–10, doi:10.1016/j.rse.2012.01.010.

Wulder, M. A., N. C. Coops, D. P. Roy, J. C. White, and T. Hermosilla (2018), Land cover 2.0, *Int. J. Remote Sens.*, *39*(12), 4254–4284, doi:10.1080/01431161.2018.1452075.

Zheng, B., S. W. Myint, P. S. Thenkabail, and R. M. Aggarwal (2015), A support vector machine to identify irrigated crop types using time-series Landsat NDVI data, *Int. J. Appl. Earth Obs. Geoinf.*, *34*(1), 103–112, doi:10.1016/j.jag.2014.07.002.

Zipper, S. C. et al. (2016), Drought effects on US maize and soybean production: spatiotemporal patterns and historical changes, *Environ. Res. Lett.*, *11*(9), 94021, doi:10.1088/1748-9326/11/9/094021.