Agricultural Water Management 271 (2022) 107764

Contents lists available at ScienceDirect

Agricultural Water Management

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

Remote sensing of field-scale irrigation withdrawals in the central Ogallala aquifer region

Steven K. Filippelli [b,*], Matthew R. Sloggy [c], Jody C. Vogeler [b], Dale T. Manning [d], Christopher Goemans [d], Gabriel B. Senay [a]

[a] US Geological Survey (USGS) Earth Resources Observation and Science Center, North Central Climate Adaptation Science Center, Fort Collins, CO 80523, USA
[b] Natural Resource Ecology Laboratory, Colorado State University, Fort Collins, CO 80523, USA
[c] USDA Forest Service, Pacific Southwest Research Station, Riverside, CA 92507, USA
[d] Department of Agricultural & Resource Economics, Colorado State University, Fort Collins, CO 80523, USA

ARTICLE INFO

Handling Editor - J.E. Fernández

Keywords:
Water abstraction
Groundwater pumping
Operational Simplified Surface Energy Balance (SSEBop)
Evapotranspiration
Local enhanced management area
Model transfer

ABSTRACT

For agricultural areas facing water scarcity, sustainable water use policy relies on irrigation information that is timely and at a high resolution, but existing publicly available water use data are often insufficient for monitoring compliance or understanding the influence of policy on individual farmer decisions. This study attempts to fill this data gap by using remote sensing to map annual irrigation quantity at the field-scale within the central Ogallala aquifer region of the United States. We compiled in situ annual irrigation volume data at the field scale in the Republican River Basin of Colorado for 2015–2018 and at the Public Land Survey System (PLSS) section scale in western Kansas for 2000–2016, which served as reference data in random forest models that relied on Landsat-based actual evapotranspiration from the Operational Simplified Surface Energy Balance model (SSEBop) along with maps of irrigated area, Landsat spectral indices, climate, soils, and derived hydrologic variables. The models explained 87% of the variability in irrigation volume in Colorado and 75% in Kansas, but accuracy declined when transferring the models in spatial cross-validation (Colorado R^2 =0.81; Kansas R^2 =0.51) and temporal cross-validation (Colorado R^2 =0.82; Kansas R^2 =0.68). Predicted annual totals of irrigation volume in western Kansas had a mean absolute error of 11.9%, which was slightly higher than the average annual change of 11%. Use of predicted irrigation maps also lead to an underestimated effect size for a water use restriction policy in Kansas. These results indicate that field- and section-scale irrigation can be mapped with reasonable accuracy within a region and time period that has adequate sample data, but that methods may need to be improved for applying the models more broadly in areas that lack extensive in situ irrigation data to support further research on water use and aid in structuring policy.

1. Introduction

Irrigation with groundwater and surface water represents a vital agricultural input responsible for substantial gains in agricultural productivity (Hornbeck and Keskin, 2014). Yet in some areas the sustainability of agricultural productivity is in jeopardy due to severe groundwater depletion (Deines et al., 2020). The Ogallala aquifer in the central Plains of the United States, in particular, accounts for a substantial percentage of the irrigation supplied to this agriculturally productive region, but excessive withdrawals from the aquifer have led to severe drawdown in some areas (Haacker et al., 2016; McGuire, 2017). This depletion of groundwater resources has been projected to result in a 24% loss in irrigated acreage by 2100, forcing these areas to transition to dryland crop agriculture or pasture (Deines et al., 2020). Failing to account for land use suitability can underestimate the county-level economic impacts of groundwater depletion by 12–45% (Deines et al., 2020).

To support sustainable groundwater use and agricultural productivity, accurate monitoring of groundwater withdrawals is needed for informing water policy and management decisions. Every five years the U.S. Geological Survey (USGS) provides a report on national water use with county-level summaries (Dieter et al., 2018), which is compiled from data collected across a multitude of entities using a variety of methods (Bradley, 2017). However, because decisions on groundwater

* Corresponding author.
E-mail address: Steven.Filippelli@colostate.edu (S.K. Filippelli).

https://doi.org/10.1016/j.agwat.2022.107764
Received 14 October 2021; Received in revised form 1 March 2022; Accepted 3 June 2022
Available online 15 June 2022

use are made at the individual field scale and policies are implemented within water districts that may not correspond to counties, irrigation monitoring is needed not just at national, state, or county levels but at the field level as well. Besides water management policies, several other factors drive individual irrigation decisions made at the field level such as crop type, soil characteristics, climate, and the type of irrigation system. Consistent and widely available data on field-level irrigation would be beneficial to understand how these factors interact to influence the decisions of individual farmers.

Several studies have examined the effects of policies, prices, and a handful of other determinants on groundwater extraction levels (Pfeiffer and Lin, 2014; Zhu et al., 2007), but these studies have focused on areas where high resolution irrigation information is available. However, there are other areas that use considerable groundwater for which high-resolution data are not available. For example, Nebraska has a collection of local regulations across the state that mandate reporting and limit extraction through allotment systems, drilling moratoria, and other means (Ifft et al., 2018). These regulations are often credited with improvements in aquifer levels (Andales et al., 2020), but such monitoring and regulation is by far the exception. Having access to a broader dataset of field-scale irrigation, both spatially and temporally, would greatly increase the potential scope of the aforementioned behavioral studies. Additionally, broadly available field-scale irrigation data could ensure water use compliance, which is often self-reported.

Recent advances in remote sensing have enabled spatially explicit estimation of water abstraction for irrigation at global to field-scales using a variety of approaches (Foster et al., 2020). Foster et al. (2020) categorized these approaches into those primarily based on satellite estimation of soil moisture, estimation of crop coefficients, and use of thermal infrared imaging for estimating evapotranspiration. Studies that primarily rely on remotely sensed soil moisture (Brocca et al., 2018; Jalilvand et al., 2019; Zaussinger et al., 2019) have been limited to mapping irrigation at gridded resolutions of 25 km or greater rather than field scales due to the coarse resolution of the synthetic aperture radar systems used, with the highest resolution soil moisture products now being 1–3 km from the fusion of Soil Moisture Active Passive and Sentinel-1 (Das et al., 2019). Crop coefficient-based methods infer vegetation condition over the growing season from remotely sensed vegetation indices from which crop water requirements are estimated. However, previous studies have shown that actual water use from in situ irrigation data is often not perfectly aligned with estimated crop water requirements (Foster et al., 2019), which could be due to imperfect information on conditions, regional water management policies, irrigation costs, or other factors.

A third class of approaches estimates irrigation by incorporating remote sensing-based actual evapotranspiration (ETa) into soil water balance models or predictive models. Recent advances in remote sensing of ETa have made this approach appealing because it can now be applied at field scales and deriving estimates of irrigation from ETa depends less upon assumptions about water application efficiency. Although ETa may be estimated from remote sensing through a variety of techniques (Zhang et al., 2016), many models use a surface-energy balance approach that utilizes land surface temperature obtained from remote sensing of thermal radiation (Allen et al., 2007; Bastiaanssen et al., 1998; Senay et al., 2013). These methods have recently been extended from coarse resolution sensors such as the Moderate Resolution Imaging Spectroradiometer (MODIS; 500 m) to Landsat (30 m) for relatively high-resolution mapping of ETa. Maps of Landsat-based ETa have been used to characterize trends in crop consumptive water use at regional scales (Senay et al., 2019) and aid in understanding the influence of water policy on consumptive water use (Senay et al., 2017). Some studies have further partitioned ETa into "blue water" originating from ground and surface water sources and "green water" originating from precipitation using a combination of remotely sensed ET and land surface models (Romaguera et al., 2012; Velpuri and Senay, 2017).

Foster et al. (2020) identified 20 studies that primarily relied upon thermal remote sensing of ET for estimating irrigation. However, only three of these studies (Folhes et al., 2009; López Valencia et al., 2020; Tazekrit et al., 2018) were applied at the field scale and incorporated the use of in situ irrigation data, and they were all limited to a single year and a single region. Little research has been done to test applicability of irrigation models to other time periods or regions that lack in situ irrigation data. Majumdar et al. (2020) leveraged 17 years of in situ irrigation data in Kansas to develop models of irrigation but did so at a 5-km resolution and primarily relied upon crop type maps that only extend back to 2006 in most areas of the United States. To achieve sustainable water use goals, this approach would need to be extended to a field scale across multiple years of data. Additionally, errors would need to be well characterized to avoid inappropriate policy decisions that could negatively affect farmers or water conservation (Foster et al., 2020).

The objective of this study was to develop and test high resolution mapping of irrigation quantity to support sustainable water management goals. We made use of advances in remote sensing of ETa and irrigated area along with other environmental data to develop empirical models of the volume and depth of irrigation applied to fields across multiple years. Models were tested at two scales to reflect two different scenarios of data availability: one in which individual field boundaries are available for known irrigated fields, and one in which field boundaries and irrigation status are unknown but where Public Land Survey System (PLSS) sections (~640 acres) can be used to aggregate several fields. Through development and testing of these models, we sought to answer the following questions:

1. How does irrigated map accuracy compare between using field-scale models and PLSS section-scale models?
2. What is the relative importance of predictor variables related to hydrology and irrigated area in modeling irrigation quantity?
3. What is the consistency of map accuracy when applying the models to new time periods and areas with similar conditions?
4. What is the potential effect of map error on downstream applications such as regional estimates of irrigation and economic analyses of water use policy?

Through answering these questions, this research makes progress towards a model framework that could be applied for high resolution mapping of irrigation in areas and time periods that lack extensive in situ water abstraction data to inform water management and policy.

2. Methods

2.1. Overview of methods

This study draws on extensive databases of in situ irrigation data and advances in remote sensing of ETa, irrigated area, and other environmental data to develop empirical models of irrigation across multiple scales and multiple years. Field-scale mapping of irrigation was tested in the Republican River Basin of Colorado where well-level data were available on water abstraction from 2015 to 2018. Mapping of irrigation was also tested in western Kansas where water abstraction data were available at the scale of Public Land Survey System (PLSS) sections (~640 acres) from 2000 to 2016. We developed and tested a series of random forest models of irrigation from this reference data utilizing predictor variables related to soil water balance components including Landsat-based ETa obtained from the Operational Simplified Surface Energy Balance (SSEBop) model (Senay et al., 2013), precipitation, and runoff along with soil water storage capacity, Landsat-based vegetation indices related to irrigated area and previously developed maps of irrigated area (Deines et al., 2019). Given the previous challenges with high uncertainty from remotely sensed irrigation, we tested the transferability of models across years and sub-regions within the study areas. We also compared total observed and predicted irrigation levels within the Kansas study area for 2000–2016, and we investigated how error in predicted irrigation could lead to different conclusions for an economic policy analysis.

2.2. Study Area

Modeling and mapping of field-level irrigation was conducted for the Republican River Basin of eastern Colorado (22,753 km^2) and western Kansas (85,444 km^2) (Fig. 1). This study area was chosen because it overlies the Ogallala aquifer (Qi, 2010) where groundwater depletion is an important issue of concern and because this area contains records of annual well-level groundwater pumping that could be joined to individual fields in Colorado or to PLSS sections in Kansas. Irrigation modeling in the Colorado study area reflects an application scenario in which field boundaries are known and irrigation is known to have occurred, and models of the Kansas study area reflect a scenario in which neither field boundaries nor irrigated area are known.

According to the 2016 Cropland Data Layer (CDL) (USDA-NASS, 2019), which maps crop type of both irrigated and non-irrigated cultivation, there were 9253 km^2 and 40,515 km^2 of various crops and hay within the Colorado and Kansas study areas, respectively. Based on the 2016 CDL, corn and winter wheat (including irrigated and non-irrigated) were the dominant crops in both the Colorado (48% winter wheat and 35% corn) and Kansas (46% winter wheat and 27% corn) study areas. These cropland areas and proportions of each crop were different for the data used in modeling after limiting to fields and sections with irrigation data and other filtering described below in Section 2.2. Corn is typically sown around late April and harvested around October. Winter wheat is the earliest sown crop that is planted in the fall of the previous year and harvested in the summer, while sorghum is the latest common crop with many fields being harvested in November (NASS, 2010).

The Colorado study area and western Kansas have cold semi-arid climates (Koppen classification BSk) with warm dry summers and cold winters, and the climate becomes humid and subtropical (Koppen classification CFa) with hot humid summers and cold winters when moving farther east into Kansas (Kottek et al., 2006). Precipitation follows an increasing gradient from west to east with the Colorado portion receiving approximately 400 mm per year and the eastern edge of the study area near Wichita, Kansas receiving approximately 900 mm per year (Daly et al., 2008). Thirty-year normals (1981–2010) of mean, minimum, and maximum temperature and vapor pressure deficit similarly increase from northwest to southeast. Mean annual temperature is approximately 10 °C in the northwest Republican River Basin of Colorado and approximately 14 °C near Wichita, Kansas (Daly et al., 2008).

2.3. Data

2.3.1. Annual field- and section-level irrigation

For the Republican River Basin of Colorado, annual irrigation of each field was derived from attributing well pumping data in 2015–2018 to field boundaries (CDSS, 2019) supplied by the Colorado Division of Water Resources (CDWR). Annual pumping volume of each well was measured and then reported to the water district by the landowner. The CDWR field boundaries were delineated from aerial imagery and Landsat maximum Normalized Difference Vegetation Index (NDVI) in each year. This field boundary dataset also contains the list of wells serving each field. Fields ranged in size, including fields that were small corners of a parcel around a center pivot (~7 ac), portions of a center pivot field, standard center pivot fields (~125 ac), and occasionally larger fields. We joined the annual well data to the field boundaries and

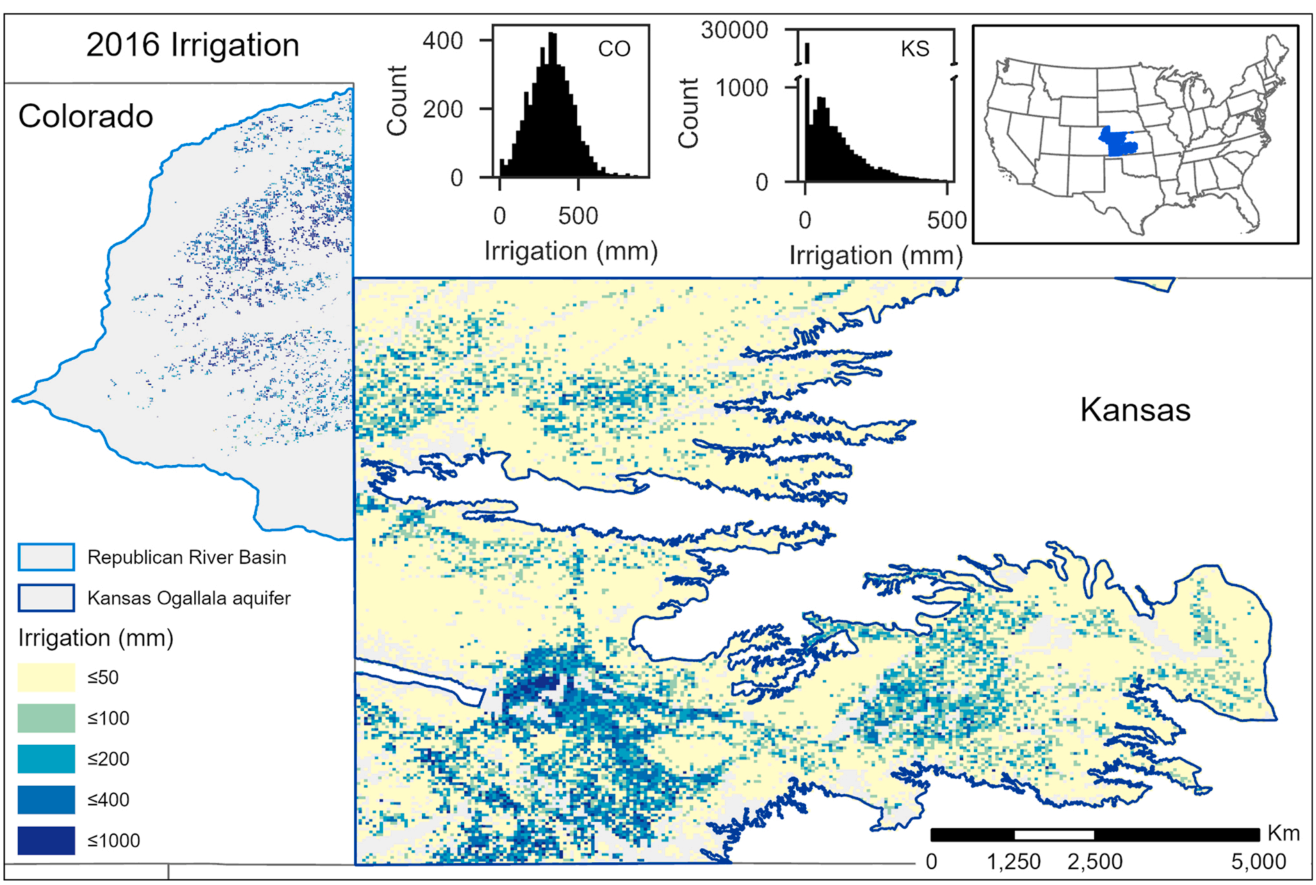

Fig. 1. Maps and histograms of 2016 depth of irrigation (mm) in individual fields in the Republican River Basin of Colorado and in PLSS sections over the Ogallala aquifer boundary (Qi, 2010) in Kansas.

summed irrigation from multiple wells serving a single field. Water from wells serving multiple fields were divided based on field area. For the 53.2% of wells that served multiple fields, a vast majority consisted of the water feeding a single circular central pivot field with the remainder going to adjacent small corner fields. Splitting total irrigation evenly across multiple fields based on field area, therefore, likely has little effect on the outcome. We removed the 0.77% of fields with unreasonably high irrigation levels (greater than 3 ac-ft per ac). Filtering the dataset left an average of 1863 km^2 of cropland in each year to be used in analysis, which was predominately corn (mean of 77% corn across years). Note that this is quite different from the proportion of crops in the region (48% winter wheat and 35% corn from 2016 CDL), and field boundaries with well data appeared to exclude some fields, likely because they were unirrigated. For example, National Agricultural Statistics Service (NASS) census data for Phillips and Yuma counties in Colorado (NASS, 2019), which are completely within the study area, show that only 8% of the harvested area of winter wheat was irrigated in 2017.

Although we were primarily interested in volume of water extracted, much of the variation in pumping volume was due to variability in field size (Pearson's r = 0.85), so pumping volume was divided by field area to obtain the irrigation quantity as depth in millimeters (mm). Modeling irrigation as depth and converting back to volume yielded slightly better performance than modeling volume directly and better illustrates how much variance is truly explained by the predictive modeling as opposed to simply knowing field size. Irrigation quantity had a normal distribution with a slight right tail (Fig. 1) and a mean of 328 mm in 2016; the distribution was similar in other years and the overall mean was 336 mm.

For the Kansas study area, 2000–2016 annual irrigation for each PLSS section was derived from well abstraction data from the Kansas Water Information Management and Analysis System (WIMAS), which is managed jointly by the Kansas Department of Agriculture, Kansas Division of Water Resources, and Kansas Geological Survey (2019). Landowners are required to self-report water use for each well on an annual basis. As of 2011, landowners using groundwater must have a flowmeter (State of Kansas, 2011), but prior to this regulation, water use may have been estimated by other means such as by multiplying pumping rate and operation time for constant rate pumps. We joined this well-level water withdrawal data to the Kansas PLSS section map and subset by lands overlapping the Ogallala aquifer. Land ownership and management is divided by these PLSS sections, and wells are typically located within the center of fields at center pivot locations. Because of this, we assumed all water from a well was applied to the same section that it was located in, although some water may have been distributed to other sections. The standard PLSS section is 640 acres, but section size varied between 13 and 909 acres because of sections on the edge of the state and principal meridians. Because sections were mostly the same size and entire sections were rarely irrigated, irrigation volume was less correlated with section area (Pearson's r = 0.07) but more closely related to estimates of irrigated area which we incorporated into other predictor variables as described in section 2.2.4. In 2016, depth of irrigation (mm) was zero-inflated and had a right skew (Fig. 1) with a mean of 39 mm; the distribution was similar in other years and the dataset had an overall mean of 43 mm. Because the irrigated area within each section was unknown and unlikely to be evenly distributed across the entire section area, we modeled volume of irrigation instead of depth for this Kansas dataset; the overall mean irrigation volume for all years was 90.7 ac-ft. We removed 0.14% of the observations with unreasonably high irrigation levels (≥3 ac-ft per ac) because the maximum allowable water use in counties of western Kansas is at most 2 ac-ft per ac (Barfield, 2017). Filtering the dataset left an average area of 83,211 km^2 in each year included in the analysis. In years with CDL data (2006–2016), 37,118 km^2 of this area was cropland on average, which was composed of 50% winter wheat, 24% corn, 16% sorghum, 4% soy, 3% alfalfa, and 3% other crops including double cropping.

2.3.2. Actual Evapotranspiration

Field-scale estimates of monthly actual and growing season evapotranspiration were obtained through application of the SSEBop model to Landsat imagery and gridMET and Daymet climate data in Google Earth Engine (Senay et al., 2017, 2016, 2013). SSEBop estimates the fraction of potential evapotranspiration (ETf) which is multiplied by a scaled (k) alfalfa reference potential evapotranspiration (ETo) obtained from the Gridded Surface Meteorological dataset (gridMET) (Abatzoglou, 2013) in order to calculate actual evapotranspiration (ETa; Eq. 1). ETf is obtained through a simplified surface energy balance approach by scaling the difference between the land surface temperature (T_s) and a cold/wet limit (T_c) by the surface psychrometric constant of a dry bare surface (γ^s).

$$ETa = (1 - \gamma^s (T_s - T_c)) * k * ETo \tag{1}$$

Land surface temperature is calculated from the corrected thermal radiance obtained from Landsat and emissivity based on the NDVI threshold method (Sobrino et al., 2004). The daily cold/wet limit of surface temperature and the surface psychrometric constant are obtained from daily maximum and minimum temperatures and average clear-sky net radiation calculated from Daymet (Thornton et al., 2016). Desert Research Institute has created a preliminary implementation of SSEBop for Landsat in Google Earth Engine (Gorelick et al., 2017) as part of the OpenET project (Morton et al., 2021), which is the implementation we have used in this study (OpenET core version 0.0.10 and SSEBop version 0.0.13). Ji et al. (2021) found 2015 monthly SSEBop estimates of ETa had high accuracy when compared against two Ameriflux eddy covariance towers in Kansas (US-KFS: R^2 = 0.93, bias = 5.5 mm, RMSE = 18.9 mm; US-KLS: R^2 = 0.87, bias = 5.7 mm, RMSE = 20.3 mm).

Landsat Tier 1 surface reflectance images from April to October for 2011–2018 were masked for clouds, cloud shadow, and snow using the provided Pixel QA band. We applied the SSEBop algorithm to these images to estimate daily ETa at 30 m and aggregated these estimates to monthly intervals by linearly interpolating daily values between available images. Field- or section-level ETa was then calculated as the mean of the monthly pixel values within each field or section boundary to reduce the influence of missing data and minor variations within a field that may not be due to actual differences in irrigation. The monthly ETa values were then summed for the growing season (April-October) as well.

2.3.3. Other predictors

Other predictor variables included those related to water balance components and irrigated area. Monthly precipitation from gridMET (2.5′, ~4 km) was incorporated in the modeling along with other hydrologic variables derived from combinations of the precipitation and runoff with ETa. To account for runoff, we obtained USGS estimates of Hydrologic Unit Code (HUC) level 8 watershed runoff, which are calculated through area-weighted averages of daily streamgage data within basins of the HUC8 watersheds (Brakebill et al., 2011) The HUC8 watershed intersecting both study areas ranged from 1059 to 9990 km^2 (mean=3706 km^2). Mean gridMET precipitation (*PRmean*) was calculated over each HUC8 watershed, and an effective precipitation coefficient (*PRcoef*) for each watershed was calculated as the fraction of watershed precipitation after subtracting *runoff* (Eq. 2):

$$PRcoef = \frac{PRmean - runoff}{PRmean} \tag{2}$$

This effective precipitation coefficient was then used to scale the pixel-level (~4 km) precipitation over each watershed, which produces pixel-level estimates of effective precipitation (*PREF*) available for crop water use. The mean effective precipitation was calculated for each field in Colorado and each section in Kansas. Runoff was generally low across watersheds and years and the precipitation coefficients were correspondingly quite high; the growing season (April-October) *PRcoef* averaged across years and watersheds for both study areas ranged from

0.69 to 1.00 and had a mean of 0.97. Because of this, the effective growing precipitation was lower than precipitation by only 2.6 mm on average for the fields in Colorado and by 10.3 mm on average for the sections in Kansas.

Irrigation (*Irr*) could be calculated directly with a simple water balance model if all other factors were known (Eq. 3), including the input of precipitation (*pr*), and outputs of actual evapotranspiration (*ETa*), runoff (*ro*), and changes in soil water storage (Δs):

$$Irr = ETa + ro + \Delta s - pr \tag{3}$$

Because the total annual change in surface soil water storage is relatively small in comparison to the factors that we can estimate, such as *ETa, pr, ro*, simply subtracting the effective precipitation from ETa to obtain net irrigation (NETIRR; Table 1) should provide a strong predictor of water abstraction for irrigation in an empirical model. The mean net irrigation (NETIRR) was calculated on a monthly and seasonal basis for each Colorado field and Kansas section.

Ideally, the annual change in soil water storage would also be included in the model and the derivation of NETIRR, but current soil moisture datasets have at best a 1 km resolution (Das et al., 2019) which is likely insufficient to capture field-level spatial variability in soil moisture changes caused by irrigation practices. In lieu of having Δs, we included the mean available water storage in the upper 25, 50, 100, and 150 cm (Table 1) as additional predictors, which were calculated for each field and section from the SSURGO database (Soil Survey Staff, 2019) of information obtained from soil surveys.

Deines et al. (2019) found a suite of Landsat spectral indices to be important predictors in detecting irrigation across the High Plains aquifer, which contains the Ogallala aquifer, from 1984 to 2017. These spectral metrics may also capture variance in irrigation quantity not explained by the above water balance components. We incorporated the spectral indices which had the highest ranked importance by Gini score in their models, including the annual maximums of the Green Chlorophyll Vegetation Index (GCVI), Normalized Difference Moisture Index (NDMI), NDVI, and the Water-adjusted Green Index (WGI; annual maximum GCVI x annual maximum NDMI) shown in Table 1. These spectral indices were also normalized by dividing by the 15th percentile value of the index within a 50-km radius to create neighborhood greenness indices (NGCVI, NNDMI, NNDVI, NWGI) as in Deines et al. (2019). Prior to calculating the neighborhood values, the indices are first masked to exclude urban areas and water using the 2016 National Land Cover Database (Yang et al., 2018) because an abundance of these land cover types depresses the apparent regional greenness. Holes in the neighborhood image were then filled with the focal median within a 10-km radius. We also limited the neighborhood values to be at most the median of the region-wide spectral index value within each study area. These neighborhood normalized greenness indices provide more consistent metrics for applying the model across a wide range of conditions as they account for regional differences in overall greenness to highlight local variations.

2.3.4. Irrigated Area

In Colorado individual fields are delineated so the potential irrigated area is known, but within the PLSS sections of the Kansas study area, the area of active cultivation and irrigation varied greatly, with many sections containing no irrigated cropland at all (Fig. 1). To reduce the influence of non-irrigated areas on estimates of average ET and Landsat spectral indices, we incorporated the use of irrigated area maps created for the entire High Plains aquifer region at a 30-m resolution from 1985 to 2019 by Deines et al. (2019), which were generated from predictive modeling with Landsat neighborhood normalized spectral indices like those described in Section 2.2.3. Use of these maps has the advantage of focusing summaries of our predictor variables to estimated irrigated areas rather than entire crop areas, many of which may not be irrigated. Also, as shown for the Colorado study area, most of the variance in irrigation volume can be explained by area alone. The estimated irrigated area and percent irrigated area within each section in Kansas were included as stand-alone predictor variables, and the mean of the Landsat-based spectral indices and evapotranspiration were obtained only over the irrigated area. Other predictor variables had much lower resolutions (e.g., precipitation at ~4 km) such that summaries using only the irrigated area would be no different than using the entire section. For the hydrology-based predictor variables (ET, precipitation, and ET minus precipitation) we multiplied the previously calculated depth values by the estimated irrigated area within each section for the appropriate year to convert these variables to volume in acre-feet (Table 1). Sections with no irrigated area were given a predicted irrigation volume of 0 ac-ft.

Table 1
Variables used in random forest regression models of irrigation. Means of the predictor variables were taken over the entirety of each field in Colorado and within irrigated areas of each PLSS section in Kansas. For hydrologic variables, depth in mm was used for Colorado, and volume in ac-ft was used for Kansas by multiplying the depth variables by the predicted irrigated area in each year.

Hydrologic Variables		
Acronym	*Description*	*Reference*
ETa_grow	Growing season (Apr-Oct) sum of monthly SSEBop actual evapotranspiration	Senay et al. (2013)
ETa_04…10	Monthly SSEBop actual evapotranspiration from April to October	Senay et al. (2013)
PREF_grow/year	Annual (Jan-Dec) and growing season (Apr-Oct) sum of monthly effective precipitation	Abatzoglou (2013)
PREF_01…12	Monthly effective precipitation from gridMET and USGS HUC8 runoff from January to December	Abatzoglou (2013)
AWS_025…150	Soil available water storage in the upper 25 cm to upper 150 cm from SSURGO	Soil Survey Staff (2019)
AREA, AREA_pct	Colorado uses field area only. For Kansas, this is the annual predicted irrigated area and percent area within each section	Deines et al. (2019)
NETIRR_grow	Differences of seasonal ETa and effective precipitation (ETa_grow – PREF_grow)	
NETIRR_04…10	Difference of monthly ETa and effective precipitation (e.g. ETa_04 – PREF_04)	

Landsat Spectral indices			
Acronym	*Name*	*Formula*	*Reference*
NDVI	Annual maximum Normalized Difference Vegetation Index	$(NIR - Red)/(NIR + Red)$	Rouse et al. (1974)
NDMI	Annual maximum Normalized Difference Moisture Index	$(NIR - SWIR1)/(NIR + SWIR1)$	Gao (1996)
GCVI	Annual maximum Green Chlorophyll Vegetation Index	$NIR/Green - 1$	Gitelson et al. (2005)
WGI	Water-adjusted Green Index	$maxNDMI \times maxGCVI$	Deines et al. (2019)
NNDVI…NWGI	Above spectral indices divided by their neighborhood value. The neighborhood is calculated as the 15th percentile of the index within a 50-km radius after excluding urban areas and water using 2016 NLCD and filling holes with a 10-km radius focal median.		Deines et al. (2019)

2.4. Irrigation modeling

The irrigated area, actual evapotranspiration, effective precipitation, derived net irrigation, Landsat-based spectral indices, and soil available water storage variables served as predictors (Table 1) of field-level irrigation depth for the Colorado study area and section-level irrigation volume in the Kansas study area using random forest models (Breiman, 2001). The models used common random forest parameters of 500 trees and splitting at each node by considering the square root of the total number of predictors. Evaluation of each model was initially done using out-of-bag predictions, which provide unbiased estimates of generalization error (Breiman, 2001) and have been shown to be as accurate as using a test set of the same size as the training set (Breiman, 2001, 1996). Models were evaluated by the coefficient of determination (R^2), root-mean-squared-error (RMSE), and mean difference of predictions from the observations (bias). Differences in model accuracy between the two study areas were primarily compared with RMSE as a percentage of the mean observed value (RMSE%) because of differences in scale when comparing irrigation volume (i.e., fields vs. sections) and differences in the distributions of irrigation depth (Fig. 1).

The primary goal of our study was to develop and test models for high prediction accuracy rather than the development of parsimonious interpretable models. However, we compared the relative importance of predictor variables to aid future research and also verified that the model was capturing the relationships between key predictor variables and the response variables that would be expected from simple water balance models, hydrologic theory, and known irrigation practices.

Predictor variables of the full models for each study area were compared by their importance values, which were calculated as the mean across trees of the total decrease in mean squared error caused by the variable when it was used for splitting nodes of each tree; these importance values were normalized so that they sum to one. The models included some highly correlated predictor variables such as adjacent months for hydrologic variables, the different Landsat spectral indices, and the available water storage variables, which were not filtered because multicollinearity does not reduce random forest performance. However, multicollinearity can confound the ability to interpret variable importance values because correlated predictors will often have similar but diluted importance scores while other predictors that may have contributed more novel information for improving model performance could be pushed down in rank (Kuhn and Johnson, 2013). A correlation heatmap is shown for both study areas to aid in interpreting the variable importance scores.

One-dimensional Accumulated Local Effects (ALE) plots (Apley and Zhu, 2020) were created to capture the main effect of the top ranked predictor variable from each of the six types of predictors: ETa, PREF, NETIRR, AWS, AREA, and spectral indices (Table 1). These plots illustrate the difference of the predicted response from the overall mean prediction across a range of values for a given predictor variable. ALE plots have the advantages of isolating the main effect for the feature of interest and preventing unreasonable combinations for correlated predictor variables (Molnar, 2019), which are problems common to the more widely used M-plots and partial dependence plots, respectively.

Because we are primarily interested in the application of these models for other regions and years with no available pumping data, we also tested transferability of the models by training on one region and testing on another region (leave-one-side-out validation) as well as training on a subset of years and testing on the withheld year (leave-one-year-out validation). These forms of validation serve to test for applicability of the models across space and time but within the same domain (i.e. environment and conditions) as the model training space because extrapolation often leads to poor results (Kuhn and Johnson, 2013). For the leave-one-side-out validation, the data were split spatially into halves for training and testing. For example, fields or sections with their centroid lying in the eastern half of the study area were used for training and the other half were used for testing, and then the training and testing sides were flipped. The same process was repeated with a north-south split. For spatial and temporal cross-validation the sample size weighted mean of the accuracy statistics from each test set are given. For the Colorado models, we also provide accuracy statistics once predicted irrigation depth is converted back into volume using field area because estimating volume was of primary interest.

For the Kansas study area, where a census of observed irrigation was available, we performed an additional analysis comparing the total annual irrigation volume from WIMAS as well as the observed total irrigation of sections that were included in modeling after filtering for outliers and missing data against the corresponding total of predicted irrigation when that year was the test set of leave-one-year-out validation. This comparison illustrates how section-level error translates to expected levels of error when aggregating over larger areas and applying the model in years lacking training data. Often models even with relatively high error can produce reasonable aggregated estimates so long as the model has relatively low bias.

2.5. Irrigation in Local Enhanced Management Areas

One of our objectives was to develop an irrigation model that could be used for examining the influence of past water management policies on local and regional water use decisions for areas that lack comprehensive irrigation data. Because most policy analyses evaluate the average effect of a policy on relevant outcomes, we are primarily interested in obtaining predictions of irrigation that are unbiased and errors that are not correlated with the primary factors of interest (e.g., areas where the policy is applied or not applied). To test this application, we examined the influence of using predicted irrigation versus observed irrigation in an analysis of policy effects on water use decisions. This analysis replicated a simplified version of the model presented in Drysdale and Hendricks (2018) using both our observed irrigation data and the predicted irrigation data at the scale of PLSS sections in Kansas.

Kansas previously implemented a system by which local authorities can self-regulate groundwater extraction by establishing a Local Enhanced Management Area (LEMA). In 2013, a LEMA was established in the area of Sheridan County in Kansas, which focused on reducing groundwater extractions. The study by Drysdale and Hendricks (2018) compared water usage inside this LEMA to the immediate surrounding area. They used a spatial discontinuity approach using water withdrawal data from a sample of wells from 2007 to 2016 within the LEMA boundary, and then all wells no greater than 5 miles and no less than 2 miles from the LEMA boundary. They ran a series of mixed effect regression models with logged dependent variables (total irrigation, irrigation intensity, and irrigated acres) and a set of weather controls as independent variables, including temperature and precipitation, as well as water-rights level fixed effects, and farmer-year fixed effects. A dummy independent variable was also included to indicate whether the well was within the LEMA or in the buffer zone outside of the LEMA. Their results showed that groundwater restrictions within the LEMA led to a reduction in groundwater use, mostly through reductions in irrigation intensity.

Our analysis is modeled after the Drysdale and Hendricks (2018) study. We used a similar set of weather controls as independent variables, including potential evapotranspiration (ET_o) in mm, average temperature (*Temp*) in °C, precipitation (*Precip*) in mm, and the square of precipitation ($Precip^2$) in mm^2 during the growing season calculated from gridMET daily weather estimates (Abatzoglou, 2013). We included precipitation as a quadratic term in the regression to control for non-linear effects precipitation may have on irrigation. Because our observation unit was a section, rather than a water right, our fixed effects were at the section level, along with year fixed effects. There were many zeros in our dependent variable, and so we excluded wells that reported no irrigation, focusing instead on sections that were irrigated. We used the level of irrigation volume rather than a log transformation as in Drysdale and Hendricks (2018) to aid in interpretation. Our empirical specification was the following:

$$v_{it} = \alpha_i + \gamma_t + \beta I_{it} + \delta X + \varepsilon_{it}$$

Where v_{it} is the volume in acre-feet of water pumped within some section i in year t. We indicated whether a section was within the LEMA with the variable I_{it}. We had section fixed effects a_i and year fixed effects γ_t. X is the set of weather controls with a set of corresponding coefficients, δ. The term ε_{it} is an independent and identically distributed (i.i.d.) random error.

We ran this regression on four sets of data covering the same extent as in Drysdale and Hendricks (2018) and years 2000–2016. The first dataset included the full set of observed section-level irrigation values in and around the LEMA. The second was the set of out-of-bag predicted irrigation values from the model of irrigation volume using all data in the Kansas study area. The third was the set of predictions from temporal cross-validation where one year of data was withheld for validation while samples from the remaining years were used for model training. The fourth was the set of predicted irrigation values from the spatial cross-validation model where the model was trained on the sections in the southern half of Kansas and then validated on the northern half of Kansas, which includes the LEMA study area.

We also ran four different versions of our model. The first included only fixed effects. The second included fixed effects and the total potential evapotranspiration (ET_o) during the growing season. The third included the fixed effects plus the weather controls of mean temperature (*Temp*), total precipitation (*Precip*), and the square of total precipitation ($Precip^2$) during the growing season. The final model included the fixed effects, evapotranspiration, and the weather controls. All results feature standard errors that were clustered at the section level. For each set of models, we compared differences in the coefficients between using the observed irrigation data and using the three predicted irrigation datasets to see if different conclusions would be drawn depending on the data source. We also evaluated whether presence of a section within the LEMA had a different average level of error in predicted irrigation volume than outside the LEMA, which could be a major factor in influencing our interpretations from the above regression models.

3. Results

3.1. Field-level irrigation in Colorado

In Colorado, the random forest model explained about half of the variability in field-level irrigation quantity (R^2 =0.49; RMSE%=30%) when modeled as depth (Fig. 2a). However, the model explains 87% of the variability in irrigation volume, which was calculated by multiplying by field size (Fig. 2d). As previously mentioned, field area alone explains 72% of the variance (Pearson's r = 0.85) in irrigation volume because of the array of field sizes in this study area. Overall bias was low (1 mm), but the two-dimensional density plots (i.e., hexbin plots) of observations and predictions show the model to underestimate high levels of irrigation with predictions rarely exceeding 500 mm while low levels of irrigation have a slight positive bias (Fig. 2a). This compression of predictions towards the mean observed value is common in models using an ensemble of decision trees like random forest. For 2016, the spatial distribution of errors from the full model in Colorado appeared to be random with no strong clustering in any one region, and the histogram of residuals was approximately normal and centered on zero (Fig. 3).

Model accuracy declined slightly when transferring the model to other time periods through leave-one-year-out cross-validation (Fig. 2c; RMSE%=35%) and also for transfer to other areas with leave-one-side-out cross-validation (Fig. 2b; RMSE%=35%) in comparison to the full model (Fig. 2a; RMSE%=30%). When converting to volume the full, spatial cross-validation, and temporal cross-validation models had slightly greater differences in RMSE% but smaller difference in R^2 and bias (Fig. 2d-f). More notable is the increase in bias with spatial and

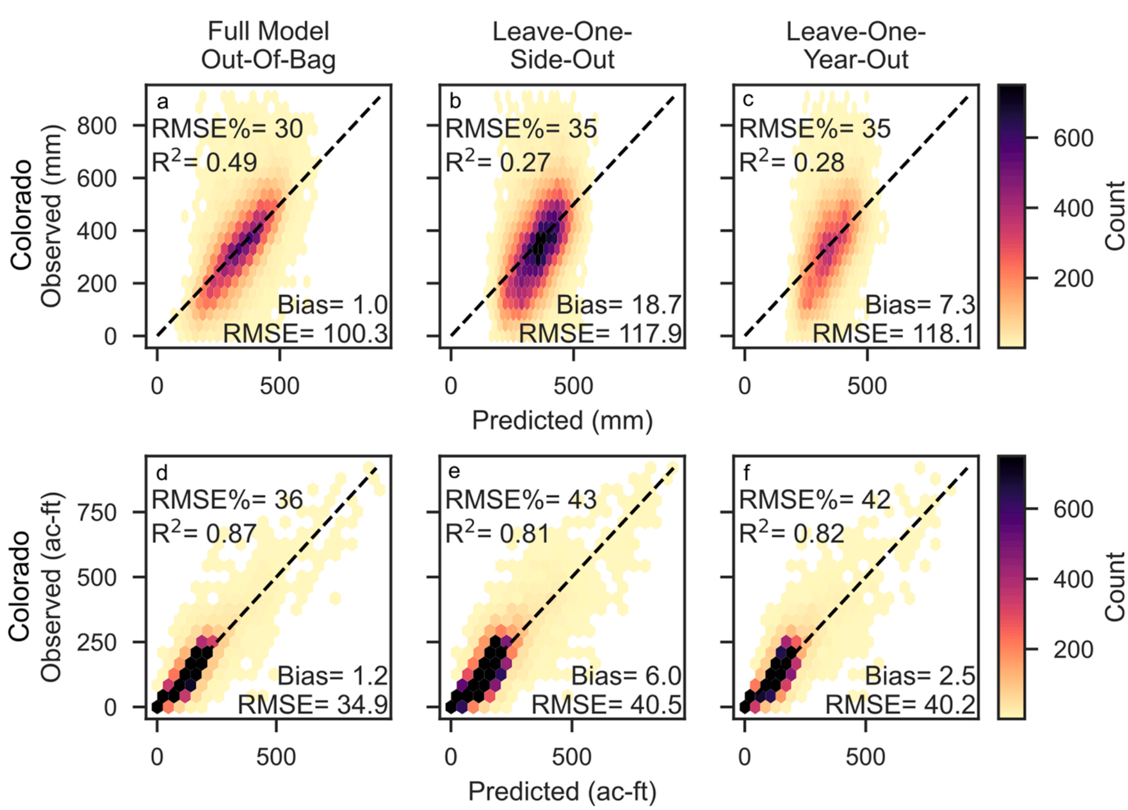

Fig. 2. Models of irrigation in the Colorado study area for a full model evaluated with out-of-bag predictions and cross-validation using one-side out at a time (i.e., east, west, north, and south halves) and one-year-out at a time. Irrigation depth in mm was modeled (a-c) before converting to volume with field area (d-f).

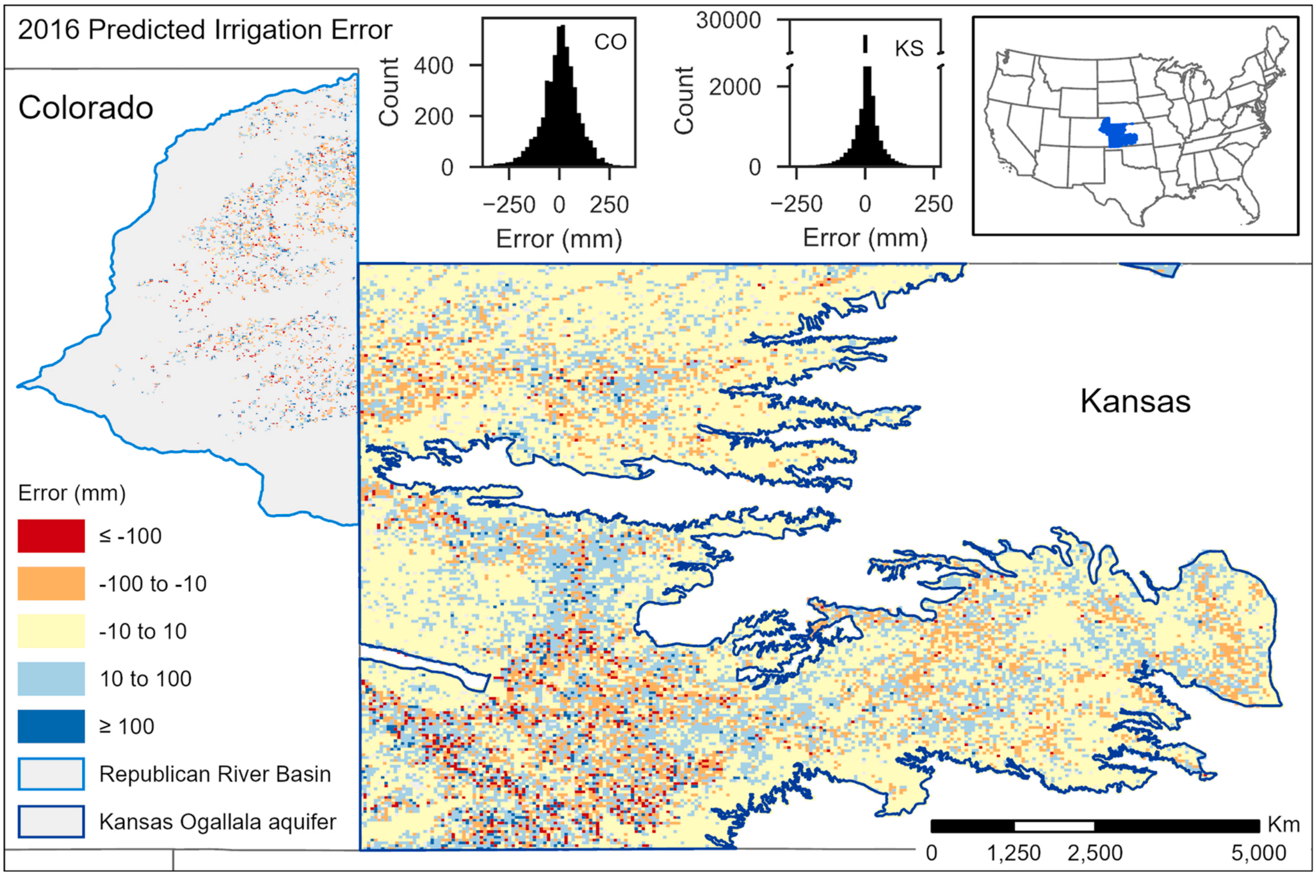

Fig. 3. Map of error for 2016 with out-of-bag predictions made from the "full" models trained with all data in each state. In Kansas, irrigation volume (ac-ft) was modeled and then converted to depth in mm using PLSS section area, but for most sections only a portion of their area was irrigated.

temporal cross-validation for the depth-based results (Fig. 2a-c), particularly for fields with low irrigation. The decreases in accuracy between the full model and when performing spatial and temporal cross-validation are likely due to spatial and temporal dependence in the dataset. One source of spatial dependence mentioned previously is the predominantly east-west precipitation gradient that exists in this region. By training and testing on different halves of the dataset we can see that the model would likely have slightly decreased accuracy when transferring to other regions without available training data, possibly even those with only slightly different conditions such as modestly higher or lower precipitation. The same can be said for transferability to different years, but this may be because with only four years in the dataset (2015–2018), the model has yet to be exposed to a wide range of different temporal conditions for training such as multiple wet or dry years.

The available water storage variables (AWS_025…100) were all ranked as the most important in the full model (Fig. 4a), each having an importance value of ~0.04–0.05. These variables were also very highly correlated (Pearson's r = 0.91–0.99; Fig. S1), and each had a monotonically decreasing but non-linear effect on predicted irrigation (Fig. S2). Growing season and summer monthly net irrigation variables (e.g. NETIRR_grow, NETIRR_07…09) were the next group of variables with the highest importance ranking (importance=0.03–0.04). NETIRR_grow had ALE values that increased in an approximately linear fashion when above ~100 mm (Fig. S2). Landsat spectral indices ranked fairly low in feature importance for this model even though they were important in models of irrigated area by Deines et al. (2019). The common variable importance method we employ here is known to be biased towards variables with higher cardinality (Strobl et al., 2007). However, all the predictor variables were continuous, and the most important variables, available water storage, had the lowest number of unique values in the set of predictors. Variable rankings were also similar when using a permutation-based variable importance method on the training data (results not shown).

3.2. Section-level irrigation in Kansas

The full section-level model of irrigation volume in Kansas explained a relatively large proportion of the variance (R^2 =0.75; Fig. 5a) but tended to overestimate in sections with low observed irrigation. Even in this region dominated by irrigated cultivation, sections with low pumping volume were common (72% had 0 ac-ft), which led to a high RMSE as a percentage of the mean observed value (RMSE%=114%). When disregarding sections with zero irrigation for calculating the accuracy statistics, RMSE% dropped to 54% for this model even though RMSE increased to 176 ac-ft. Overall model accuracy improved only slightly to an RMSE of 101.4 when only considering data from 2011 and later, after the flowmeter installation requirements took effect. The predicted irrigation volume was converted to depth in mm (Fig. 5d-f) using the PLSS section areas for comparison with the results from the Colorado study area, but this conversion likely underestimates the amount of irrigation concentrated on the irrigated area because it effectively assumes that the entire section was irrigated. The RMSE% and R^2 for the volume estimates (Fig. 5a-c) were nearly the same as depth-based converted results (Fig. 5d-f) because the section area explained little variation in irrigation volume; results were nearly identical if modeling irrigation depth directly and when excluding sections with zero irrigation. For 2016, the spatial distribution of errors in mm from the full model appeared to show some clustering with small regions of over- and under-estimation, but as in Colorado the residuals

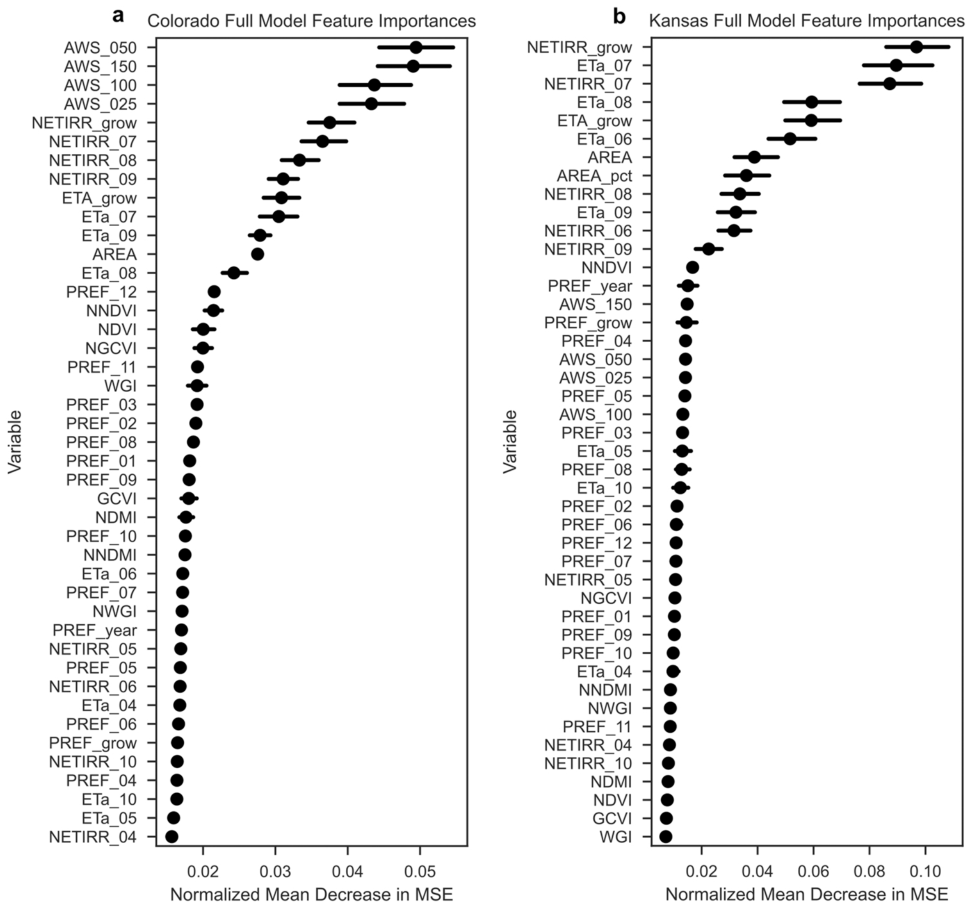

Fig. 4. Random forest variable importance values for the full models in Colorado (a) and Kansas (b). Variable importance was calculated as the normalized mean decrease in mean square error across the trees (black dot) with the 95% confidence interval (black line) calculated from bootstrapped draws (n = 1000) of the feature importance from individual trees. See Table 1 for variable definitions.

were centered on zero and approximately normal (Fig. 3).

The net irrigation and ETa variables summed over the growing season (NETIRR_grow and ETa_grow) and in the summer (NETIRR_07, ETa_07, ETa_08) were the top five predictor variables (Fig. 4b). All five were highly multicollinear (Pearson's r = 0.88–0.97; Fig. S3), in part because they were all transformed to volume using irrigated area. The ALE plots of the ETa and NETIRR predictors show steep increases of predicted irrigation as their values increased in the middle ranges but with flat responses at the ends (Fig. S4). The Deines et al. (2019) predicted irrigated area (AREA) alone explained 50% of the variance in irrigation volume (Pearson's r = 0.71) and had a linear effect on predictions until approaching an asymptote after about 450 acres (Fig. S4).

Although engineering the hydrologic predictors to be in volume units by incorporating irrigated area greatly improved their explanatory power, the resulting multicollinearity does confound our ability to determine their relative importance. To parse out the effect of irrigated area, we evaluated variable importance for another model of Kansas irrigation volume with hydrologic predictors used in their original depth form rather than converted to volume using irrigated area (Fig. S5). In this model irrigated area and percent had the highest importance values (0.20 and 0.18, respectively) followed by net irrigation and ETa variables for the growing season (0.05 and 0.03, respectively) and July (0.04 and 0.04, respectively). Because using the hydrologic variables as depth reduces some of their multicollinearity (Pearson's r = 0.47–0.83 for the same five previously listed), other variables that helped to improve model accuracy rose in rank, such as NGCVI and the available water storage variables.

As in the Colorado study area, model accuracy declined when examining spatial and temporal transferability with leave-one-side-out validation (Fig. 5b) and leave-one-year-out validation (Fig. 5c). Attempting to train and test the model on split halves caused many samples with high observed irrigation to be underestimated (Fig. 5b), which led to an overall negative bias. This underestimation with spatial transfer is likely due to differences in the distribution of the response between the training and testing halves (Table 2). For example, training on the north half which has a mean of 36.5 ac-ft and 83% zeros leads to a bias of − 44.2 ac-ft when testing on the south half which had a much higher mean of 123.2 ac-ft and only 66% zeros. Irrigation volume was much more consistent temporally with annual mean irrigation having a coefficient of variation (CV) of 13% as opposed to the 47% CV of spatial halves. This greater temporal consistency led to lower error with temporal cross-validation (RMSE%=126%) than spatial cross-validation (RMSE%=159%), although the apparent error was still fairly high because of zero-inflation depressing the mean of the observations. The annual sample-size weighted mean of annual irrigation including zeros was 91 ac-ft while it was 327 ac-ft when excluding zeros. Even though the overall bias for sections with zero irrigation was reasonably low (25 ac-ft) in temporal cross-validation, the accuracy appears to be better when not considering sections with zero irrigation (RMSE%=58%) because of zero-inflation.

We examined the effect of model error when totaling irrigation over the entire western Kansas study area (Fig. 6). The WIMAS census of irrigation allowed us to make a direct comparison of the total predicted irrigation (i.e., using predictions of each withheld year from temporal cross-validation) to the unfiltered total irrigation. However, because sections with outlier irrigation values (i.e., > 3 ac-ft per ac) and missing data were filtered out in model development, the predicted total irrigation is best compared to the same set of observations used in

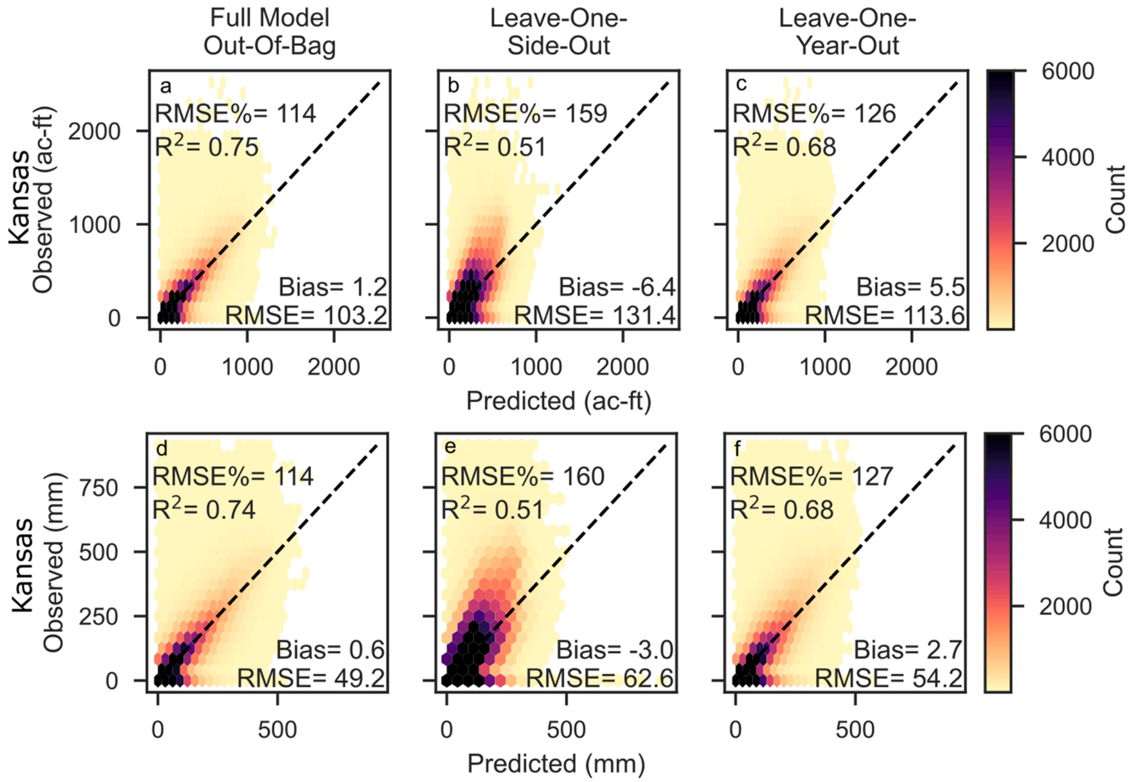

Fig. 5. Models of irrigation in the Kansas study area for a full model evaluated with out-of-bag predictions and cross-validation using one-side out at a time (i.e., east, west, north, and south halves) and one-year-out at a time. Irrigation volume in ac-ft was modeled (a-c), and the predicted volume was converted to irrigation depth in mm using PLSS section area (d-f). Note that this conversion likely underestimates the depth of irrigation over the irrigated area because it assumes that an entire section was irrigated.

Table 2
Summary statistics of each half of the Kansas irrigation volume modeling dataset from leave-one-side-out cross-validation as well as its accuracy statistics from when it was the test set.

Side	Count	% Zeros	Mean	Std. Dev	R^2	RMSE	Bias
East	172095	72	64.9	150.7	0.39	117.4	44.5
North	205366	83	36.5	106.2	0.31	88.0	21.4
South	342572	66	123.2	239.6	0.53	165.1	-44.2
West	375843	72	102.5	224.0	0.66	130.9	-10.5

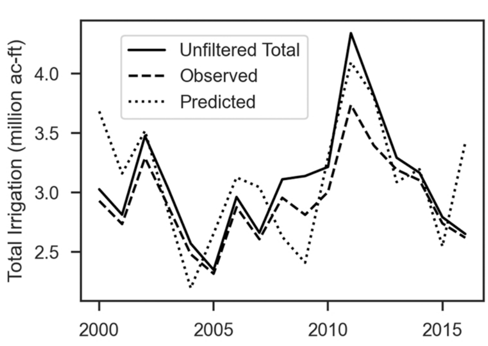

Fig. 6. Comparison of the total annual irrigation in western Kansas provided by WIMAS before any filtering for modeling, the total observed irrigation used in modeling, and the total predicted irrigation for each year when it was left out of the model in leave-one-year-out cross-validation.

modeling. The mean annual "observed" irrigation was 2.92 million ac-ft and the mean absolute error across years was 0.34 million ac-ft. The absolute error was 11.9% of the observations on average, which slightly exceeded the mean absolute percent change between subsequent years of 11.0%. Although we expected error to be highest in the drought year of 2012, irrigation volume was overestimated the most in 2000 (26%) and 2016 (31%), possibly because of overestimation in irrigated area. These two years also had the highest mean percent irrigated area (AREA_pct) for sections that actually had an observed irrigation of 0 ac-ft. Overestimation in sections with zero observed irrigation accounted for on average 63% of the total overestimation volume across years. No one factor explained the differences in trends between the observed and predicted total irrigation, but predicted irrigation most closely tracked the trend of the annual mean of the top predictor (NETIRR_grow).

3.3. Effect of data source on LEMA policy considerations

We tested whether the predicted irrigation maps would be effective in an application that investigates the average effects of water policy on groundwater extraction using an analysis akin to that in Drysdale and Hendricks (2018). We found that for the observed data, being within the LEMA was associated with an approximate 21.746 ac-ft reduction in water volume extraction (Table 3). Using irrigation from out-of-bag estimates and temporal and spatial cross-validation also showed the LEMA being associated with reductions in water extraction, albeit with lower magnitudes of − 9.774, − 8.697, and − 9.714 ac-ft, respectively.

In order to evaluate whether the differences in the LEMA coefficients were significant between models using observed and predicted irrigation, we conducted a bootstrapping procedure to obtain their distributions. We iteratively sampled the sections in our data with replacement 1000 times, re-estimating our models for each sample set and storing the coefficient values to generate a distribution for the LEMA indicator

Table 3
Coefficients from linear mixed effect models of observed irrigation and predicted irrigation for PLSS sections in a LEMA of Kansas and the immediate surrounding area when using all weather control independent variables. Predicted irrigation volumes were obtained from out-of-bag estimates of the model trained on the entire Kansas study area, predictions from leave-one-year-out cross-validation ("Temporal CV") and leave-one-side-out cross-validation ("Spatial CV"). Standard errors of the coefficients are in parentheses. df is degrees of freedom.

	Observed	Out-of-Bag	Temporal CV	Spatial CV
LEMA (I_{it})	-21.746*	-9.774*	-8.691*	-9.714*
	(2.730)	(1.565)	(1.408)	(1.638)
Temp	-24.136	-110.350*	-132.430*	-134.769*
	(20.143)	(18.951)	(20.183)	(22.226)
ET_o	0.059	0.190	0.167	0.283
	(0.324)	(0.209)	(0.253)	(0.238)
Precip	3.666	50.789*	60.721*	69.858*
	(2.688)	(3.045)	(3.194)	(3.553)
$Precip^2$	-0.736*	-3.108*	-3.517*	-3.654*
	(0.246)	(0.300)	(0.301)	(0.280)
Observations	3960	3960	3960	3960
R^2	0.887	0.918	0.925	0.936
Adjusted R^2	0.878	0.912	0.920	0.931
Residual Std Error (df=3675)	61.162	48.964	51.044	54.879

* $p < 0.01$

coefficient. The bootstrapped difference in the LEMA coefficients between using the observed irrigation and different sets of the predicted irrigations (i.e., out-of-bag, temporal cross-validation, and spatial cross-validation) were evaluated with a one-sample t-test with the null hypothesis that the difference was 0. Table 4 shows that the LEMA coefficients were very similar when comparing different sets of independent variables (i.e., weather controls, potential evapotranspiration, or only fixed effects) for the same source for the response variable. However, the LEMA coefficient values had small but significant differences of − 7.494 to − 8.468 ac-ft depending on whether observed irrigation values or one of the sets of predicted irrigation estimates were used for the dependent variable (Table 4). When comparing across just the fixed effects model, the relative effect of the LEMA was 46%, 41%, and 46% for the out-of-bag, temporal cross-validation, and spatial cross-validation irrigation sources, respectively, in comparison to the effect of the LEMA when using the observed irrigation values.

As shown in Figs. 2 and 5, low observed irrigation values tend to be overestimated and high observed irrigation values tend to be underestimated for each model whether evaluated by out-of-bag estimates, spatial cross-validation, or temporal cross-validation. Because of these biases caused by random forest modeling, we decided to also test the effect that the LEMA could have had on predicted irrigation. A bias or error in the predicted irrigation that was correlated with our policy factor of interest (the LEMA) could undermine our ability to use the predicted irrigation maps for policy analysis. In this case, if the LEMA was effective in decreasing irrigation then it may have lower irrigation values than outside the LEMA on average, and if those lower observed irrigation values in the LEMA were then overestimated by the random forest model it could lead to a lower apparent magnitude of difference in irrigation between inside and outside the LEMA (i.e., a smaller LEMA coefficient compared to using the observed irrigation). To test this potential effect of differences in irrigation prediction error or residual (i.e., *observed – predicted*) between inside and outside of the LEMA, we performed another regression with the prediction error as the response for each of the predicted irrigation sources (i.e., out-of-bag, temporal CV, and spatial CV). In each case the irrigation model tended to have a significantly higher overestimate of irrigation inside the LEMA than outside the LEMA (Table 5). This discrepancy in error was mostly composed of lower values (~200 ac-ft) in the LEMA that were overestimated but also a greater abundance of high irrigation sections (>400 ac-ft) outside the LEMA that were being highly underestimated.

4. Discussion

4.1. Comparison of high resolution irrigation models

The increasing availability of Landsat-based ETa has enabled examination of field-scale patterns in crop water consumptive use in agriculture regions such as the Central Valley of California (Schauer and Senay, 2019), the Upper Rio Grande Basin (Senay et al., 2019), and Colorado River Basin (Senay et al., 2016). However, in regions such as the Ogallala aquifer that primarily rely on groundwater pumping and also have some precipitation input, high resolution maps of groundwater withdrawal are useful for water planning, monitoring policy compliance, and understanding the influence of policy on water use. In this study, we leveraged high-resolution ETa from SSEBop and Landsat along with data related to other water balance components for field-scale mapping of irrigation quantity from groundwater withdrawal. We compared modeling at different spatial scales as well as spatial and temporal model transferability for potential application of these models in places and time periods where field-level irrigation data are unavailable.

Comparing modeling at a section level with the Kansas study area to field-level modeling in Colorado was complicated by several confounding factors besides scale, such as differences in their distributions (e.g. zero-inflated), differences in which crops were dominant (corn in

Table 5
Coefficients for the location of a section in the LEMA in a regression of the residuals (observed – predicted) for each model of irrigation volume as well as model fit statistics.

	Out-of-Bag	Temporal CV	Spatial CV
LEMA (I_{it})	-11.763*	-12.762*	-11.577*
	(2.167)	(2.335)	(2.468)
Observations	3960	3960	3960
R^2	0.523	0.544	0.574
Adjusted R^2	0.487	0.510	0.542
Residual Std Error (df=3679)	65.683	71.240	76.319

* $p < 0.01$

Table 4
Mean LEMA coefficients (β) from bootstrapping and their differences between models using different sources for the dependent variable (i.e. observed versus predicted irrigation). Results are shown for different sets of independent variables (i.e. fixed effects, weather, evapotranspiration, all variables [total]) and different data sources for the predicted response (i.e. out-of-bag, temporal cross-validation, and spatial cross-validation).

Model	Observed	Out-of-Bag OOB-Difference		Temporal CV Difference		Spatial CV Difference	
	β	β	Difference	β	Difference	β	Difference
Fixed Effects	-13.962	-6.357	-7.605*	-5.722	-8.240*	-6.457	-7.505*
Weather	-14.059	-6.291	-7.768*	-5.591	-8.468*	-6.215	-7.844*
Evapotranspiration	-13.949	-6.356	-7.594*	-5.724	-8.255*	-6.456	-7.494*
All Controls	14.047	-6.281	-7.767*	-5.582	-8.466*	-6.199	-7.848*

* $p < 0.01$

Colorado and wheat in Kansas), and other regional differences such as climate. Under ideal circumstances, differences in scale would be compared in the same region, but in Colorado the irrigation status of every field was not known (see Section 2.2.1), which could introduce substantial error when aggregating to the section level or a raster grid. Comparing the two scales in different study areas illustrated that knowledge of irrigated area was of primary importance in determining the accuracy of models of irrigation quantity whether as depth or volume at either scale because even in a region with some of most extensive irrigated cropland in the United States, much of the land area is still non-irrigated. Because of this, field-scale models of irrigation depth with exact irrigated field boundaries in Colorado had higher accuracy (RMSE %=30%) than similar models in Kansas for which irrigated area was predicted (RMSE%=54%), even when sections with zero irrigation were disregarded.

4.2. Consideration of predictors

Majumdar et al. (2020) similarly found density of agriculture derived from 2015 CDL to be the most important predictor in modeling irrigation depth in Kansas at a scale of 5 km even though the model was applied across multiple years and also included other environmental variables such as MODIS-based ETa and PRISM climate data. This is expected given the high degree of zero-inflation. For example, 87% of sections in Kansas between 2011 and 2016, the years in one test set used by Majumdar et al. (2020), have 0 ac-ft of irrigation and most are not cultivated at all. Even at the scale of individual sections, we found simply using the annual percent area of major CDL classes as predictors in random forest of irrigation depth for all of Kansas between 2011 and 2016 yielded an RMSE of 45.79 mm. Aggregating well-level reference data to a coarser resolution raster grid also reduces the variance that needs to be explained by a model, often improving accuracy statistics. However, aggregation into a grid assumes irrigation is evenly distributed within the area of each grid cell even though grid cells likely do not align with fields or land ownerships, which can introduce error in the gridded reference data used for validation. These factors of scale and incorporating eastern Kansas, which has less agriculture, may in part explain the comparatively high accuracy in tests by Majumdar et al. (2020), which contrasts with the higher resolution and focus on the densely cultivated region of western Kansas used in this study. On the other hand, Ji et al. (2021) found a stronger relationship between irrigation and ETa in the more arid western Kansas region, where generalized efficiency of water use (ETa / (precipitation + irrigation – change in soil moisture) is very high.

The central dilemma with field-level modeling of irrigation is that field boundaries are rarely known in the past or over broad regions, even in the case where irrigation volumes are known. The Common Land Unit dataset produced by the U.S. Department of Agriculture delineated field boundaries from aerial imagery, but it was removed from public distribution in 2008. In comparison, maps of PLSS sections are readily available and the Kansas WIMAS database of annual well abstractions can be aggregated at this level to provide an extensive training and testing dataset for future model development, which could potentially be applied in similar regions back to the start of the Landsat archive. At this scale, high resolution maps of irrigated area, such as those from Deines et al. (2019), became essential and were by far the most important variables in predicting irrigation quantity. Given their importance, efforts to map field-scale irrigation may be best spent refining and extending high resolution irrigated area maps. Recent development of Landsat-based ETa maps such as those from SSEBop (Senay et al., 2013) and the OpenET project (https://openetdata.org/) could aid greatly in this effort.

We sought to include datasets closely related to the components of a simple water balance model when constructing our predictor variables. Even though they are still under development, Landsat-based maps of growing season ETa (ETa_grow) and derived net irrigation accounting for precipitation and runoff (NETIRR_grow) were the most important predictors at the section level, although unexpectedly, available water storage was more important at the field scale. Landsat-based ETa with SSEBop shows strong agreement with estimates from eddy covariance flux-towers (Ji et al., 2021; Senay et al., 2020) and well represents relative landscape patterns of ETa, but future efforts may benefit from local calibration to remove bias when integrating with other datasets such as precipitation and runoff. Estimation of local runoff rather than watershed-level runoff could also improve models of irrigation, though we believe runoff to be a comparatively minor factor as discussed in the methods (Section 2.3.3). The neighborhood normalized Landsat-based variables (e.g. NGCVI), which have proven effective in mapping irrigated area (Deines et al., 2019), may have had lower importance in our models simply because the spatial patterns in vegetation greenness they represent were already well captured by the irrigated area maps based upon these variables as well as the Landsat-based ETa variables. In the course of compiling datasets, we initially attempted to include a much larger suite of climate data from gridMET, such as mean vapor pressure deficit, and soil data from SSURGO, such as drainage class, but even with 400 + variables the models had negligible improvements of at most 2.7 mm RMSE and 0.8 ac-ft RMSE in Colorado and Kansas, respectively. We also initially included additional crop area variables derived from CDL, such as the area of major crops like corn, but these provided little additional explanatory power (i.e., <1% reduction in RMSE) and would limit application of models to years with CDL availability, which is 2006 at the earliest in Kansas. These comparisons show that predictor variables less directly related to field-scale hydrology increase model complexity for little benefit and may even reduce model transferability. Advances in mapping water balance components such as further downscaling precipitation and soil moisture, disaggregating runoff to fields, or improving the accuracy of ETa may be the best avenues for improvements in mapping field-scale irrigation quantity.

4.3. Spatial and temporal transferability of models

Through our analysis, we discovered challenges to developing models of irrigation quantity that could be applied reliably in times and regions outside the spatiotemporal extent of the training data. For example, when evaluating temporal transfer for estimating total irrigation in western Kansas, we found the average error slightly exceeded the average annual change, meaning evaluation of annual changes over a region could be attributed to model error. Majumdar et al. (2020) also found decreases in accuracy when applying irrigation models to other time periods, although they chose to test on select groups of years rather than performing a full temporal cross-validation. Majumdar et al. (2020) also found only modest decreases in accuracy when testing a model on a withheld sliver through the central portion of their study area. To reflect the intended application of our model to similar regions that are completely outside the training area, we tested on withheld halves of the study area. We found more substantial increases in model error through this approach (e.g., increase in RMSE% of 45 in Kansas with spatial cross-validation), which could largely be attributed to differences in the distribution of irrigation across regions. Although we did not evaluate irrigation totals in the field-scale models, the model error was lower in both spatial and temporal transfer (RMSE%=35%) than at the section-scale, but application of these models would require field boundaries or assuming that field boundaries remain consistent in years without data. Problems with irrigation reference data itself may also exist given that 0.14% of the western Kansas sections had reported water use greater than 3 ac-ft per ac (much greater than the regulated maximums) but still comprised 4.6% of the total irrigation volume in the study area across all years. Considering the degree of these problems, it would be best to formally account for uncertainty either through Monte-Carlo methods (Coulston et al., 2016) or using a model-based inference framework with parametric models (McRoberts, 2011) when applying remote sensing-based models of irrigation to evaluate irrigation trends, particularly for times and areas outside the training data.

4.4. Implications of irrigation maps and error

A recent review by Foster et al. (2020) cautions against the use of remote sensing-based irrigation mapping without accounting for error and demonstrates how inaccurate maps could have negative consequences for water management and policy. In our policy analysis example, the true effect of the LEMA was underestimated by 7.5–8.5 ac-ft, which could lead policy analysts to erroneously believe that the LEMA was not as effective and that water use limits need to be further decreased to account for overuse. However, depending upon the application, that level of error in estimating the effect of the LEMA may not make a meaningful difference for decision-making. Foster et al. (2020) notes that even unbiased errors in irrigation estimates can result in overly strict water use regulations that lead to substantial economic losses for farmers because of the asymptotic relationship between irrigation and crop yield. We also found that differences in the error of predicted irrigation between inside and outside the LEMA (Table 5) was at least one major factor driving down the apparent effect size of the LEMA. Thus, it is important that analysts also are aware of the potential for spatial or temporal variations in error for predicted irrigation, not just overall error, because of its potential influence on subsequent analysis. Machine learning models like random forest commonly overestimate low values and underestimate high values, driving estimates more towards the mean observed value. Developing models that have lower bias in predictions at the ends of the distribution is one potential avenue to making maps of irrigation that would be more useful in policy applications. Given the relatively high levels of overall error in predicted irrigation from this study and others (Foster et al., 2020; Majumdar et al., 2020), clear communication between policy and remote sensing analysts on the acceptable and expected levels of uncertainty and bias would be beneficial when developing and using irrigation maps for policy applications.

Despite the aforementioned challenges, improved models and the resulting maps of irrigation would be beneficial for a host of applications. Water use monitoring data is lacking in most regions, and although additional in situ data will still be a requirement for model development, remote sensing holds the potential to fill data gaps by estimating irrigation in very similar regions, such as adjacent water districts, or in other time periods. Models could be extended across the Landsat archive (1984-present) to analyze spatiotemporal trends of irrigation in regions with similar crops and climate regimes as long as care is taken to avoid extrapolating beyond the range of the training data. More accurate annual maps could be used to monitor compliance and serve as a check in areas that rely on self-reporting (Foster et al., 2020). They could also serve to identify illegally drilled wells, which may be prevalent in low-income countries with many small-scale farmers (Al Naber and Molle, 2017). Perhaps the most important application of irrigation maps would be for evaluating the effectiveness of water management policies to ensure the long-term sustainability of groundwater resources in the many regions that lack in situ monitoring.

5. Conclusion

The recent development of high-resolution remote sensing of ETa and irrigated area have enabled the potential to map field-scale irrigation quantity to aid water use monitoring and evaluate the effects of water use policy in regions and times for which field-scale data are currently unavailable. In this research we combined remotely sensed data and other environmental information with well-level data in two regions for testing empirical modeling of irrigation at different scales, regions, and time periods. We found that models at the field-scale had slightly higher accuracy than those at the section scale, and that accuracy declined with spatial and temporal transfer for models at both scales. Knowledge of irrigated area, whether through exact field boundaries or remote sensing-based maps, was of primary importance in predicting irrigation volume and improved the importance of hydrologic variables in our random forest models. Given the level of model error, as well as potential errors in the reference field data, caution is warranted when applying such models for evaluating irrigation trends. Future efforts may look to improve model accuracy through improvements in mapping irrigated area and possible local calibration of remotely sensed evapotranspiration. Remote sensing of irrigation has the potential to serve several beneficial applications such as checking compliance with water regulations, understanding the average effects of water policy, and estimating aquifer draw-down. However, given the currently high error of remotely sensed irrigation maps and the propensity of people to regard maps as truth, careful communication of uncertainty by the scientific community would be necessary to avoid leading map users and policymakers astray.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by USDA National Institute of Food and Agriculture grant number 2018–67023-27693, “Crop Insurance and Groundwater in the Ogallala Region.” Thank you to Michael Falkowski for co-leading original project development. Charles Morton provided advice for using the Google Earth Engine implementation of the Operational Simplified Surface Energy Balance model. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.agwat.2022.107764.

References

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

Al Naber, M., Molle, F., 2017. Controlling groundwater over abstraction: state policies vs local practices in the Jordan highlands. Water Policy 19, 692–708. https://doi.org/10.2166/wp.2017.127.

Allen, R.G., Tasumi, M., Trezza, R., 2007. Satellite-Based energy balance for mapping evapotranspiration with internalized calibration (METRIC)—Model. J. Irrig. Drain. Eng. 133, 380–394. https://doi.org/10.1061/(ASCE)0733-9437(2007)133:4(380).

Andales, A.A., Rudnick, D., Chávez, J.L., 2020. Special issue: improving irrigation management across the Ogallala aquifer, USA. Irrig. Sci. 38, 481–483. https://doi.org/10.1007/s00271-020-00704-0.

Apley, D.W., Zhu, J., 2020. Visualizing the effects of predictor variables in black box supervised learning models. J. R. Stat. Soc. Ser. B Stat. Methodol. 82, 1059–1086. https://doi.org/10.1111/rssb.12377.

Barfield, D.W., 2017. Kansas Water Appropriation Act: Rules and Regulations. https://agriculture.ks.gov/docs/default-source/statues-water/kwaa_rules_regs57C3ADA8D515.pdf.

Bastiaanssen, W.G.M., Menenti, M., Feddes, R.A., Holtslag, A.A.M., 1998. A remote sensing surface energy balance algorithm for land (SEBAL). 1. Formulation. J. Hydrol. 212–213, 198–212. https://doi.org/10.1016/S0022-1694(98)00253-4.

Bradley, M.W., 2017. Guidelines for preparation of State water-use estimates for 2015 (USGS Numbered Series No. 2017–1029), Guidelines for preparation of State water-use estimates for 2015. Open-File Report. U. S. Geol. Surv., Rest., Va. https://doi.org/10.3133/ofr20171029.

Brakebill, J.W., Wolock, D.M., Terziotti, S.E., 2011. Digital hydrologic networks supporting applications related to spatially referenced regression modeling1. JAWRA. J. Am. Water Resour. Assoc. 47, 916–932. https://doi.org/10.1111/j.1752-1688.2011.00578.x.

Breiman, L., 2001. Random forests. Mach. Learn 45, 5–32. https://doi.org/10.1023/A:1010933404324.

Breiman, L., 1996. Out-of-bag estimation. https://www.stat.berkeley.edu/~breiman/OOBestimation.pdf.

Brocca, L., Tarpanelli, A., Filippucci, P., Dorigo, W., Zaussinger, F., Gruber, A., Fernández-Prieto, D., 2018. How much water is used for irrigation? A new approach exploiting coarse resolution satellite soil moisture products. Int. J. Appl. Earth Obs. Geoinf. 73, 752–766. https://doi.org/10.1016/j.jag.2018.08.023.

CDSS, 2019. Coloado's Decision Support Systems (CDSS): Republican Basin Irrigated Lands Geodatabase (includes all years 2010, 2015–2020).

Coulston, J.W., Blinn, C.E., Thomas, V.A., Wynne, R.H., 2016. Approximating prediction uncertainty for random forest regression models. Photogramm. Eng. Remote Sens. 82, 189–197. https://doi.org/10.14358/PERS.82.3.189.

Daly, C., Halbleib, M., Smith, J.I., Gibson, W.P., Doggett, M.K., Taylor, G.H., Curtis, J., Pasteris, P.P., 2008. Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. Int. J. Climatol. 28, 2031–2064. https://doi.org/10.1002/joc.1688.

Das, N.N., Entekhabi, D., Dunbar, R.S., Chaubell, M.J., Colliander, A., Yueh, S., Jagdhuber, T., Chen, F., Crow, W., O'Neill, P.E., Walker, J.P., Berg, A., Bosch, D.D., Caldwell, T., Cosh, M.H., Collins, C.H., Lopez-Baeza, E., Thibeault, M., 2019. The SMAP and Copernicus Sentinel 1A/B microwave active-passive high resolution surface soil moisture product. Remote Sens. Environ. 233, 111380 https://doi.org/10.1016/j.rse.2019.111380.

Deines, J.M., Kendall, A.D., Crowley, M.A., Rapp, J., Cardille, J.A., Hyndman, D.W., 2019. Mapping three decades of annual irrigation across the US high plains aquifer using landsat and google earth engine. Remote Sens. Environ. 233, 111400 https://doi.org/10.1016/j.rse.2019.111400.

Deines, J.M., Schipanski, M.E., Golden, B., Zipper, S.C., Nozari, S., Rottler, C., Guerrero, B., Sharda, V., 2020. Transitions from irrigated to dryland agriculture in the Ogallala Aquifer: land use suitability and regional economic impacts. Agric. Water Manag. 233, 106061 https://doi.org/10.1016/j.agwat.2020.106061.

Dieter, C.A., Maupin, M.A., Caldwell, R.R., Harris, M.A., Ivahnenko, T.I., Lovelace, J.K., Barber, N.L., Linsey, K.S., 2018. Estimated use of water in the United States in 2015. Circular 1441. U.S. geological survey. Rest., Va. https://doi.org/10.3133/cir1441.

Drysdale, K.M., Hendricks, N.P., 2018. Adaptation to an irrigation water restriction imposed through local governance. J. Environ. Econ. Manag. 91, 150–165. https://doi.org/10.1016/j.jeem.2018.08.002.

Folhes, M.T., Rennó, C.D., Soares, J.V., 2009. Remote sensing for irrigation water management in the semi-arid Northeast of Brazil. Agric. Water Manag. 96, 1398–1408. https://doi.org/10.1016/j.agwat.2009.04.021.

Foster, T., Gonçalves, I.Z., Campos, I., Neale, C.M.U., Brozović, N., 2019. Assessing landscape scale heterogeneity in irrigation water use with remote sensing and in situ monitoring. Environ. Res. Lett. 14, 024004 https://doi.org/10.1088/1748-9326/aaf2be.

Foster, T., Mieno, T., Brozović, N., 2020. Satellite-based monitoring of irrigation water use: assessing measurement errors and their implications for agricultural water management policy. Water Resour. Res. 56 e2020WR028378. https://doi.org/10.1029/2020WR028378.

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

Gitelson, A.A., Viña, A., Ciganda, V., Rundquist, D.C., Arkebauer, T.J., 2005. Remote estimation of canopy chlorophyll content in crops. Geophys. Res. Lett. 32. https://doi.org/10.1029/2005GL022688.

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

Haacker, E.M.K., Kendall, A.D., Hyndman, D.W., 2016. Water Level Declines in the High Plains Aquifer: Predevelopment to Resource Senescence. Groundwater 54, 231–242. https://doi.org/10.1111/gwat.12350.

Hornbeck, R., Keskin, P., 2014. The historically evolving impact of the ogallala aquifer: agricultural adaptation to groundwater and drought. Am. Econ. J. Appl. Econ. 6, 190–219. https://doi.org/10.1257/app.6.1.190.

Ifft, J., Bigelow, D.P., Savage, J., 2018. The Impact of Irrigation Restrictions on Cropland Values in Nebraska. J. Agric. Resour. Econ. 43, 195–214.

Jalilvand, E., Tajrishy, M., Ghazi Zadeh Hashemi, S.A., Brocca, L., 2019. Quantification of irrigation water using remote sensing of soil moisture in a semi-arid region. Remote Sens. Environ. 231, 111226 https://doi.org/10.1016/j.rse.2019.111226.

Ji, L., Senay, G.B., Friedrichs, M., Schauer, M., Boiko, O., 2021. Characterization of water use and water balance for the croplands of Kansas using satellite, climate, and irrigation data. Agric. Water Manag. 256, 107106 https://doi.org/10.1016/j.agwat.2021.107106.

Kansas Department of Agriculture, Division of Water Resources (DWR) and Kansas Geological Survey (KGS), 2019. Kansas Water Information Management and Analysis System (WIMAS), version 5 [WWW Document]. URL https://www.sciencebase.gov/catalog/item/56a7edb4e4b0b28f1184da78 (accessed 1.1.19).

Kottek, M., Grieser, J., Beck, C., Rudolf, B., Rubel, F., 2006. World map of the Köppen-Geiger climate classification updated. Meteorol. Z. 15, 259–263.

Kuhn, M., Johnson, K., 2013. Applied Predictive Modeling. Springer, New York, NY.

López Valencia, O.M., Johansen, K., Aragón Solorio, B.J.L., Li, T., Houborg, R., Malbeteau, Y., AlMashharawi, S., Altaf, M.U., Fallatah, E.M., Dasari, H.P., Hoteit, I., McCabe, M.F., 2020. Mapping groundwater abstractions from irrigated agriculture: big data, inverse modeling, and a satellite–model fusion approach. Hydrol. Earth Syst. Sci. 24, 5251–5277. https://doi.org/10.5194/hess-24-5251-2020.

Majumdar, S., Smith, R., Butler, J.J., Lakshmi, V., 2020. Groundwater withdrawal prediction using integrated multitemporal remote sensing data sets and machine learning. Water Resour. Res. 56 e2020WR028059. https://doi.org/10.1029/2020WR028059.

McGuire, V.L., 2017. Water-level and recoverable water in storage changes, High Plains aquifer, predevelopment to 2015 and 2013–15 (USGS Numbered Series No. 2017–5040), Water-level and recoverable water in storage changes, High Plains aquifer, predevelopment to 2015 and 2013–15, Scientific Investigations Report. U.S. Geological Survey, Reston, VA. https://doi.org/10.3133/sir20175040.

McRoberts, R.E., 2011. Satellite image-based maps: scientific inference or pretty pictures? Remote Sens. Environ. 115, 715–724. https://doi.org/10.1016/j.rse.2010.10.013.

Molnar, C., 2019. Interpretable. Mach. Learn.: A Guide Mak. Black Box. Models Explain.

Morton, C., Harding, J., Erickson, T., Carrara, W., 2021. OpenET: Filling a Critical Data Gap in Water Management for the Western United States. Wiley. ⟨https://openetdata.org⟩.

NASS, 2019. 2017 Census of Agriculture (No. AC-17-A-51), geographic area series. Natl. Agric. Stat. Serv., Usda.

NASS, 2010. Field Crops Usual Planting and Harvesting Dates, Agricultural Handbook Number 628. National Agricultural Statistics Service, USDA.

Pfeiffer, L., Lin, C.-Y.C., 2014. Does efficient irrigation technology lead to reduced groundwater extraction? Empirical evidence. J. Environ. Econ. Manag. 67, 189–208. https://doi.org/10.1016/j.jeem.2013.12.002.

Qi, S., 2010. Digital map of the aquifer boundary for the High Plains aquifer in parts of Colorado, Kansas, Nebraska, New Mexico, Oklahoma, South Dakota, Texas, and Wyoming (USGS Numbered Series No. 543), Digital map of the aquifer boundary for the High Plains aquifer in parts of Colorado, Kansas, Nebraska, New Mexico, Oklahoma, South Dakota, Texas, and Wyoming, Data Series. U.S. Geological Survey, Reston, VA. https://doi.org/10.3133/ds543.

Romaguera, M., Krol, M.S., Salama, Mhd, S., Hoekstra, A.Y., Su, Z., 2012. Determining irrigated areas and quantifying blue water use in europe using remote sensing Meteosat Second Generation (MSG) products and Global Land Data Assimilation System (GLDAS) data. Photogramm. Eng. Remote Sens. 78, 861–873. https://doi.org/10.14358/PERS.78.8.861.

Rouse, J.W., Haas, R.H., Schell, J.A., Deering, D.W., 1974. Monitoring vegetation systems in the Great Plains with ERTS. Proc. Presente 3rd Earth Resour. Technol. Satell. (ERTS) Symp. 48–62.

Schauer, M., Senay, G.B., 2019. Characterizing crop water use dynamics in the central valley of california using landsat-derived evapotranspiration. Remote Sens 11, 1782. https://doi.org/10.3390/rs11151782.

Senay, G.B., Bohms, S., Singh, R.K., Gowda, P.H., Velpuri, N.M., Alemu, H., Verdin, J.P., 2013. Operational evapotranspiration mapping using remote sensing and weather datasets: a new parameterization for the SSEB approach. JAWRA J. Am. Water Resour. Assoc. 49, 577–591. https://doi.org/10.1111/jawr.12057.

Senay, G.B., Friedrichs, M., Singh, R.K., Velpuri, N.M., 2016. Evaluating Landsat 8 evapotranspiration for water use mapping in the Colorado River Basin. Remote Sens. Environ. Landsat 8 Sci. Results 185, 171–185. https://doi.org/10.1016/j.rse.2015.12.043.

Senay, G.B., Kagone, S., Velpuri, N.M., 2020. Operational global actual evapotranspiration: development, evaluation, and dissemination. Sensors 20, 1915. https://doi.org/10.3390/s20071915.

Senay, G.B., Schauer, M., Friedrichs, M., Velpuri, N.M., Singh, R.K., 2017. Satellite-based water use dynamics using historical Landsat data (1984–2014) in the southwestern United States. Remote Sens. Environ., Big Remotely Sensed. Data: tools, Appl. Exp. 202, 98–112. https://doi.org/10.1016/j.rse.2017.05.005.

Senay, G.B., Schauer, M., Velpuri, N.M., Singh, R.K., Kagone, S., Friedrichs, M., Litvak, M.E., Douglas-Mankin, K.R., 2019. Long-Term (1986–2015) crop water use characterization over the upper rio grande basin of United States and Mexico Using Landsat-Based Evapotranspiration. Remote Sens 11, 1587. https://doi.org/10.3390/rs11131587.

Sobrino, J.A., Jiménez-Muñoz, J.C., Paolini, L., 2004. Land surface temperature retrieval from LANDSAT TM 5. Remote Sens. Environ. 90, 434–440. https://doi.org/10.1016/j.rse.2004.02.003.

Soil Survey Staff, 2019. Soil Survey Geographic (SSURGO) Database. https://sdmdataaccess.sc.egov.usda.gov.

State of Kansas, 2011. Kansas Water Flowmeter Regulations. K. A. R. 5–1–1. https://agriculture.ks.gov/docs/default-source/dwr-water-appropriation-forms/meterregs.pdf.

Strobl, C., Boulesteix, A.-L., Zeileis, A., Hothorn, T., 2007. Bias in random forest variable importance measures: Illustrations, sources and a solution. BMC Bioinforma. 8, 25. https://doi.org/10.1186/1471-2105-8-25.

Tazekrit, I., Benslimane, M., Simonneaux, V., Hartani, T., Hamimed, A., Tazekrit, I., Benslimane, M., Simonneaux, V., Hartani, T., Hamimed, A., 2018. Estimation of irrigation water pumping by remote sensing: application of the SAMIR model to citrus under mediterranean climate conditions. Rev. Bras. Meteorol. 33, 391–400. https://doi.org/10.1590/0102-7786333016.

Thornton, P.E., Thornton, M.M., Mayer, B.W., Wei, Y., Devarakonda, R., Vose, R.S., Cook, R.B., 2016. Daymet: daily surface weather data on a 1-km grid for North America. Version 3. ORNL DAAC. https://doi.org/10.3334/ORNLDAAC/1328.

USDA-NASS, 2019. USDA National Agricultural Statistics Service Cropland Data Layer. Published crop-specific data layer. https://nassgeodata.gmu.edu/CropScape.

Velpuri, N.M., Senay, G.B., 2017. Partitioning evapotranspiration into green and blue water sources in the conterminous United States. Sci. Rep. 7. https://doi.org/10.1038/s41598-017-06359-w.

Yang, L., Jin, S., Danielson, P., Homer, C., Gass, L., Bender, S.M., Case, A., Costello, C., Dewitz, J., Fry, J., Funk, M., Granneman, B., Liknes, G.C., Rigge, M., Xian, G., 2018. A new generation of the United States national land cover database: requirements, research priorities, design, and implementation strategies. ISPRS J. Photogramm. Remote Sens. 146, 108–123. https://doi.org/10.1016/j.isprsjprs.2018.09.006.

Zaussinger, F., Dorigo, W., Gruber, A., Tarpanelli, A., Filippucci, P., Brocca, L., 2019. Estimating irrigation water use over the contiguous United States by combining satellite and reanalysis soil moisture data. Hydrol. Earth Syst. Sci. 23, 897–923. https://doi.org/10.5194/hess-23-897-2019.

Zhang, K., Kimball, J.S., Running, S.W., 2016. A review of remote sensing based actual evapotranspiration estimation. WIREs Water 3, 834–853. https://doi.org/10.1002/wat2.1168.

Zhu, T., Ringler, C., Cai, X., 2007. Energy price and groundwater extraction for agriculture: exploring the energy-water-food nexus at the global and basin levels. Int. Conf. Link. Energy Water Manag. Agric. Dev. Ctries., Hyderabad, India.