Agriculture, Ecosystems and Environment 359 (2024) 108753

Contents lists available at ScienceDirect

Agriculture, Ecosystems and Environment

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

A geospatial assessment of soil properties to identify the potential for crop rotation in rice systems

Luke A. Salvato[a,*], Cameron M. Pittelkow[a], Anthony T. O'Geen[b], Bruce A. Linquist[a]

[a] Department of Plant Sciences, Univ. of California, One Shields Ave., Davis, CA 95616, USA
[b] Department of Land, Air, and Water Resources, University of California, One Shields Ave., Davis, CA 95616, USA

ARTICLE INFO

Keywords:
Crop rotation
Soil properties
Random forest model
Agroecological diversification
Sustainability
Rice systems

ABSTRACT

Crop rotation is one strategy for adapting agroecosystems to a framework that balances ecological diversity, sustainability, and food production. The Sacramento Valley, one of the most productive rice growing regions in the US, faces sustainability challenges including increasing herbicide resistant weed pressure and water use restrictions. Increasing crop diversity may help address these challenges, but this region has unique soil attributes including high clay content, salinity, alkalinity, and cemented subsurface layers, and the degree to which these soil properties influence crop rotation decisions remain unclear. The objectives of this study were to quantify the extent of crop rotation in this region, compare soil properties for rotated and continuous rice fields, and assess the potential for expanding rotations based on the geographic coverage of influential soil variables. Using satellite derived land cover data for 2007–2021, our analysis shows that only ~5000 ha are in rotation with rice, while 220,000 ha are in continuous rice production. This land cover information is fused with SSURGO soil maps in a spatial random forest model. The modeling approach indicates that fields with soil pH between 6.5 and 8, EC between 0.5 and 2 (ds m^{-1}), and saturated hydraulic conductivity less than 2 (μm s^{-1}) are more likely to be rotated. However, we estimate that only 11% of the continuous rice area has all three of these soil properties combined, suggesting soil limitations are an important constraint. This research highlights a method for evaluating land use decisions in relation to spatial variability of soil properties to better understand barriers to agroecological diversification.

1. Introduction

Increased crop rotation is being explored as one potential solution to global challenges in agricultural sustainability (Altieri et al., 2015; Cabell and Oelofse, 2012). Rice (*Oryza sativa* L.), the main staple food for nearly half the world's population (Awika, 2011; F.A.O, 2019; Yuan et al., 2021), is often grown in continuous cropping systems supporting one to three rice crops per year. Continuous flooded rice production can maintain high productivity due to biological and chemical soil processes unique to flooded agricultural soils (Pampolino et al., 2008; Cassman and Pingali, 1995; Waha et al., 2020; Bronson et al., 1998). However, there are agricultural sustainability challenges for modern continuous rice systems. To support these challenges, diversifying rice-based cropping systems with non-flooded crops is being explored in different contexts (Baste et al., 2021; Cassman and Grassini, 2020; Horton et al., 2021).

In California, rice production is concentrated in the Sacramento Valley, where it is grown on approximately 210,000 ha (*USDA - NASS, 2021*). California is the second largest rice growing state in the US, with some of the highest rice yields in the world (Hill et al., 2006). However, the long-term viability of the California rice industry is threatened by a number of challenges including increasing weed pressure and water scarcity (Hanson et al., 2014; Gebremichael et al., 2021). California rice has the highest number of herbicide resistant weed species of any other crop or region in the U.S. (Hanson et al., 2014). Also, populations of new weed species, such as weedy rice (also known as red rice), are evolving a suite of phenotypic traits that closely resemble cultivated rice, making them particularly difficult to manage (De Leon et al., 2019). Droughts across the Western US have led to severe water shortages, including water restrictions for growers (Hanak et al., 2019; Gebremichael et al., 2021). Gebremichael et al. (2021) found that fallow land across California's Central Valley tripled during drought years due to water use restrictions. In four of the last 10 years, during periods of severe drought, rice acreage declined leading to widespread fallowing throughout the

* Corresponding author.
E-mail address: lasalvato@ucdavis.edu (L.A. Salvato).

https://doi.org/10.1016/j.agee.2023.108753
Received 3 February 2023; Received in revised form 19 July 2023; Accepted 18 September 2023
Available online 4 October 2023
0167-8809/

rice growing region (USDA, 2020). Drought conditions are expected to increase in frequency and severity due to climate change (Cayan et al., 2010), and alternative cropping system strategies are needed to maintain agricultural sustainability.

Diversifying the number of crops grown in the region could be an important component of integrated weed management strategies and has potential benefits for water conservation. Rotations can be part of integrated weed management strategies by allowing for the use of different modes of herbicide action, and cultivation techniques and irrigation systems that are different from those used in typical continuous rice systems and can target different weed species including pervasive aquatic weeds (Kayeke et al., 2017; Beckie et al., 2004; Brim-DeForest et al., 2017). Rice is one of the most water intensive California crops, behind almonds, pistachios, and alfalfa (Cody and Johnson, 2015; Medellin-Azuara, 2022). Common alternative irrigated summer annual crops in the region such as processing tomatoes, dry beans, and safflower use 36%, 9.2%, and 3.4% the water that rice uses annually (Cody and Johnson, 2015). Rotating continuous rice with less water intensive annual crops could help maintain agricultural productivity while meeting water use restrictions.

Despite the potential benefits of crop rotation for weed management and water conservation in rice systems, soil constraints may be a major limitation. Large portions of the Sacramento Valley are reclaimed wetlands that have soil properties that are suitable to flooded rice production but make crop rotation difficult (Hill et al., 2006). Previous studies have reported that more than half of the rice growing region is considered 'rice only', where production of other summer or winter crops is, due to properties of the soil, expected to fail due to poor yield and high input costs (Carter et al., 1994; Hill et al., 2006). The remaining half of the region has been described as having limited rotation capability. Soil features such as floodplains, heavy clays, salinity and/or alkalinity, and cemented subsurface layers are widespread in the region and are perceived as major limitations to rotation (LaHue and Linquist, 2021; Rosenberg et al., 2022; Hill et al., 2006). Some rice growers are successfully rotating rice with summer irrigated crops (tomatoes, corn, safflower, and dry beans), vetch and wheat as a winter annual crop, alfalfa, and other forages (Rosenberg et al., 2022), with most rotation occurring in a limited portion the Southern Sacramento Valley (Rosenberg et al., 2022; Carter et al., 1994). However, these evaluations of soil properties were not based on digital USDA-NRCS Soil Survey Maps, which offer the most detailed soil map data for the US (Soil Survey Staff, 2014), and the relative amount of rice area under some form of rotation is poorly quantified. Thus, there is an opportunity to explore the feasibility of rotations based on soil limitations using soil survey maps to determine the relationship between crop rotation decisions and soil properties.

Machine learning methods are one potential approach for assessing the role of soil properties on future land use scenarios to improve natural resource management. Machine learnings models can be used to efficiently examine relationships between multiple, interacting soil parameters and their influence on different land use categories at large spatial scales. For example, these approaches have been used to predict crop rotation in the US Midwest (Socolar, 2021), to assess the potential for dryland agriculture in the High Plains region, USA (Deines et al., 2020), to identify floodplains at high resolution across the continental US (Woznicki et al., 2019), to simulate the conversion of grasslands to grain in the great plains (Olimb and Robinson, 2019), and to predict future cropland expansion (Rashford et al., 2011). These studies illustrate the novel insights that can be gained by integrating land use maps with underlying soil properties in a machine learning framework.

The overall goal of this study was to explore the potential for agroecological diversification of rice-based cropping systems based on a geospatial assessment of soil limitations. The method presented here uses land cover data, SSURGO soil data, and a spatial random forest model to identify key soil properties associated with continuous rice fields and rotated rice fields. The specific objectives were to: 1) quantify the total rice area under crop rotation and continuous rice; 2) evaluate differences in soil properties between rotated and continuous rice fields; and 3) estimate the potential continuous rice area that could support a high likelihood of rotations based on the most influential soil variables identified in random forest models. Results will help inform the feasibility of crop rotations as a tool for enhancing the long-term sustainability of California rice systems.

2. Materials and methods

2.1. Data sources and processing

2.1.1. Land use data

Land use maps for the Sacramento Valley were built by integrating the Crop Land Data Layer (CDL) with field boundary data provided by the California Department of Water Resources (DWR). The CDL is a high resolution (30 m) national land cover data set that provides crop-level information on a yearly basis. The CDL is generated from Landsat satellite missions and developed by the United States Department of Agriculture/ National Agricultural Statistics Service (USDA/NASS) CropScape project (NASS CDL, 2022). The CDL includes up to 141 land use classes, 117 of which are agricultural. In California, the CDL is currently available from 2007 to 2021, and all years were used in this analysis.

While the CDL is considered a powerful tool for understanding agricultural landscapes in the United States (Lark et al., 2017) it is still prone to uncertainties that result from land cover classification using satellite remote sensing data. To reduce these errors and improve overall accuracy, the CDL was integrated with a high accuracy land cover and field boundary data set provided by DWR (Verburg et al., 2011; Seo et al., 2014). The field boundary data set is prepared by LandIQ, a private mapping company based in Sacramento CA, and provided to the California DWR Regional Office Land Use office (DWR, 2022). The integration approach is as follows. Pixels, or portions of pixels, outside of field boundaries were excluded, managing errors where fields do not align with pixels, and where edge effects can influence acreage estimates (Lark et al., 2017). Within the field boundary, each field was reclassified as the dominant pixel type, mitigating errors where, for example, a few incorrect pixels are scattered across a rice field. Acreage estimates and change detection were performed on this reclassified, field level data set.

After reclassifying all fields within the Sacramento Valley region, we limited the analysis to the rice growing area by selecting all the fields where the dominant class was rice in at least one year of the 15-year data set. In total there were 13,120 fields covering 268,950 ha. This region has a high diversity of annual and perennial crop types. To simplify our analysis and to increase accuracy (Lark et al., 2017), we grouped CDL crop classes together into eleven dominant groups. These groups are: rice, fallow, summer annual, winter annual, other annual, alfalfa, grasses, walnut, almond, other perennial, and other (See table S1 for a complete list of the CDL classes in each group). We used this data set to examine land use changes in the region including yearly changes in rice area, the acreage under rotation with rice, and the area that has been converted to perennial tree crops such as almonds and walnuts. To assess the accuracy of our custom data set, we compared our county level rice acreage estimates to NASS data (USDA - NASS, 2020) for the eight major counties in the rice growing region.

2.1.2. Soil data

Spatial soil information in this region was obtained from the Soil Survey Geographic Database (SSURGO) developed by NRCS (NRCS Soils, 2022). SSURGO data was accessed using the FedData package (Bocinsky, 2019), which downloads federal geospatial data directly from the internet and loads it into RStudio. A range of soil variables important to agricultural production were used in the preliminary data analyses including chemical properties (acidity, salinity, sodium adsorption ratio), physical properties (soil texture, saturated hydraulic conductivity, linear extensibility), and general descriptors of soil type (soil order, soil series, irrigated capability class, or the presence of subsurface layers that are restrictive to root growth).

For all numeric soil variables, a depth weighted average was computed across all horizons in the rooting zone (top 30 cm of the soil profile). In this process, horizons that start below 30 cm depth are excluded and the thickness of all remaining horizons is computed. A weighted mean for each numeric soil property was calculated by multiplying the soil property value from each horizon by the thickness of the horizon, summing the value for all horizons present, and dividing by the total depth (30 cm). This depth weighted average was then applied across the SSURGO map unit area, and field level averages were computed for each field in the study area. This process was conducted using the sf package in R (Pebesma, 2018). Categorical soil variables, such as soil order or irrigated capability class, have only one value per component and do not require a depth weighted average. For this data, the dominant soil component (largest percent composition within the map unit) was selected, and the location of each field's centroid was used to determine the field level value.

2.2. Modelling process

2.2.1. Data preparation

To compare soil properties of continuous rice fields and rotated rice fields, a binomial classification approach was used. For this approach, two classes were determined from the land cover data set: '*continuous rice*' fields and '*rotated rice*' fields. These two groups had strict criteria: Fields that were in rice at least 12 of the 15-year data set and were fallowed in the alternate years (i.e. never planted in a summer annual, winter annual, alfalfa, or grass) were considered '*continuous rice*' fields; Fields that were in rice at least seven years and were rotated with summer annual crops, winter annual crops, alfalfa, or grasses on at least two separate occasions were considered '*rotated rice*' fields. We used these two categories to make a direct comparison between continuous rice fields and fields that are in rice almost half the time but also rotated with other crops. The core objective was to explore the potential for rotation in fields that are currently continuous rice, thus we set a medium requirement of years in rice for rotated fields rather than including fields that were predominantly other crops or not rotated with rice on two separate occasions. As a result, some rice fields were excluded from the model. Fields that were in rice 9 of the 15 years and other land uses the remaining years were placed into a distinct category called '*rice + other*' (15.8% of total area). This includes fields that transitioned out of rice to urban uses, or fields that were fallowed and/or planted with other crops such as alfalfa or annuals only on a few occasions. Fields that were in annual crops (summer annuals, winter annuals, alfalfa, and grasses) 9 of the 15 years were considered '*annual*' (6.3% of total area). Fields that were frequently fallowed (9 of the 15 years) were considered '*fallow*' (1.3% of total area).

2.2.2. Random forest

After randomly splitting the data set into training and validation subsets (75% to train and 25% to validate), a random forest classification model (Breiman, 2001) was trained to create a binary prediction for 'continuous rice' and 'rotated rice' based on soil variables. Random forest is an ensemble learning method that has recently become very popular because it combines the interpretability of decision trees with the performance of modern learning algorithms such as artificial neural networks and SVMs. Random forest models use multiple independently constructed decision trees, each with a unique bootstrap sample of the training data set, thus reducing the variance of single trees and improving prediction accuracy (Liaw and Wiener, 2002; Wiener; Prasad et al., 2006). Furthermore, random forest models are efficient, insensitive to overfitting, and are relatively straight forward to implement (Belgiu and Drăgu, 2016).

2.2.3. Model assessment

Following model training, model assessment was performed on our remaining validation data set. The primary objective was to correctly classify rotated fields within the rice area. Therefore, we examined classification accuracy using precision (P), recall (R), and F1 score of rotated fields. Precision is how often the classified product (rotated fields) is correct when compared to the source data set (Eq. (1)). Recall, also known as the hit rate, is how often the source dataset is correctly classified by the model (Eq. (2)). F1 measures classification accuracy of rotated fields by combining precision and recall using their harmonic mean (Eq. (3)).

$$P = \frac{T_P}{T_P + F_P} \tag{1}$$

$$R = \frac{T_P}{T_P + F_N} \tag{2}$$

$$F1 = 2\frac{P \quad x \quad R}{P + R} \tag{3}$$

Where T_p is the number of true positives (number of pixels correctly classified as rotated fields), F_p is the number of false positives (number of pixels incorrectly classified as rotated fields), and F_N is the number of false negatives (number of pixels incorrectly classified as continuous rice fields). In our modelling process, the F1 score was used as the primary measure for model evaluation because it balances precision and recall.

2.2.4. Model set up

In spatial data, observations that are relatively close tend to be more related to each other, which means that training and validation data sets are rarely independent, violating an important prerequisite of model building and leading to highly optimistic evaluations of predictive power (Arlot and Celisse, 2010; Ploton et al., 2020). Factors other than soil properties can influence the spatial distribution of crop rotation across a landscape, such as market distance, access to equipment, or economic factors. (Socolar et al., 2020; Rosenberg et al., 2022). One method to deal with spatial heterogeneity in is spatial cross validation. Spatially cross validated models of ecological data can have better performance at predicting error estimates and predicting to new data or predictor space, as well as for selecting causal predictors (Roberts, 2017). We used R's 'spatialsample' package (Silge, 2021), to implement spatial cross validation. In summary, our training data was split into ten cross validation groups using k-means clustering of the field's spatial coordinates.

Next, model hyperparameters were defined and tuned. Because the model was assessed primarily with F1, these combinations of hyperparameter values were optimized for F1. F1 was static after a minimum of 150 trees, so 150 trees were used to ensure adequate trees for all models. To define the number of variables randomly selected as candidates at each split (*mtries),* and the minimum number of data points in a node that is required for the node to be split further (*min _n)*, a hyperparameter grid search was performed with values from one to six and 20–40, respectively. *Mtries* = 3 and *min _n* = 31 were selected based on the highest F1 score.

Sampling strategy is another model parameter that can require careful calibration, especially when there is substantial class imbalance (Woznicki et al., 2019). In our binomial data set, continuous rice fields were 16-times more prevalent than rotated rice fields. Sampling strategies that adjust the prevalence of either class in the training data can affect precision and recall rates (Chen, 2004; Woznicki et al., 2019). Thus, we also optimized for F1 in our model tuning process. The minority class was up sampled randomly and at 10%, 20%, and 50% sampling ratios. The strategy using a 20% sampling ratio had the highest F1 score, so this strategy was used in the final model deployment. Woznicki et al. (2019) used this sampling strategy and similarly found that a 20% sampling regime was optimal because of its higher recall.

2.2.5. Variable importance

An important task in machine learning interpretation is to understand which predictor variables have the strongest influence on the predicted outcome. To accomplish this, a ranked order of variable importance in the classification model was determined using the permutation method based on AUC. In this method, AUC is computed for each tree after permuting each predictor variable (Greenwell et al., 2018). This method is considered more robust towards instances of class imbalance (Janitza et al., 2013). Variable importance was computed in R's 'vip' package (Greenwell et al., 2018).

To improve interpretability, we pruned our set of predictor variables to reduce model complexity without compromising accuracy. Redundant soil variables were removed if they were highly correlated ($r > 0.8$). Moreover, if model predictions for a variable did not show clear patterns in partial dependence plots (explained below) and omission of this variable in the model did not affect F1 scores when evaluated on the validation data set, variables were removed. This included linear extensibility (%), cation exchange capacity (CEC), drainage class, the presence of a cemented layer, and the soil series name. The remaining set of soil variables used in the modeling process were pH, electrical conductivity (EC), sodium adsorption ratio (SAR), saturated hydraulic conductivity (Ksat), taxonomic soil order (soil order), and irrigated capability class (ICC) (see Table 1 for a complete description of these soil variables).

Throughout our model tuning process, variable importance scores were often tied or closely ranked. To establish a clear order of variable importance, the variable importance computation was executed 200 times across the training data, and the variables were ranked based on their mean effect on the AUC score.

Table 1
SSURGO predictor variables used in the modeling process.

Variable name	Type	Description
EC	Continuous	The electrical conductivity of an extract from saturated soil paste (EC_e, ds m^{-1}). EC represents the ability of a soil to conduct or attenuate electrical current. It is a metric reflecting the content of soluble salts in the soil matrix, also known as salinity or ion concentration.
pH	Continuous	The negative logarithm to the base 10, of the hydrogen ion activity in the soil using the 1:1 soil-water ratio method. A numerical expression of the relative acidity or alkalinity of a soil sample.
Ksat	Continuous	*Saturated hydraulic conductivity* The ease with which pores of a saturated soil transmit water (μm s^{-1}); the amount of water that would move vertically through a unit area of saturated soil in unit time under unit hydraulic gradient.
SAR	Continuous	*Sodium adsorption Ratio* A measure of the amount of sodium (Na) relative to calcium (Ca) and magnesium (Mg) in the water extract from saturated soil paste. Soils with high SAR values, say greater than 13, may be characterized by an increased dispersion of organic matter and clay particles, reduced Ksat, and aeration, and a general degradation of soil structure.
Soil Order	Categorical	*Taxonomic soil order* The highest level in the soil classification system where soils are grouped into twelve orders based on distinct soil characteristics and ecological significance. Soil orders are typically defined by a dominant characteristic affecting soils in that location including the type of prevalent vegetation, the type of parent material, climate variables that influence soil formation, the amount of physical and chemical weathering present, and/or the relative amount of soil profile development that has taken place.
ICC	Categorical	*Irrigated capability class* Describes the general suitability of soils for most kinds of field crops where irrigation is used, including chemical, physical, and biological soil parameters. ICC values range from one to eight, and lower numbers indicate better growing conditions. For example, fields with ICC of two or less have few to moderate limitations that reduce the choice of plants that could be grown. While fields with ICC of three or greater require special conservation practices or have severe limitations that reduce the choice of plants.

2.2.6. Model application

Once a subset of the important features was identified, expected target responses were computed while accounting for the average effect of the other predictors in the model. This produces a partial dependence plot (PDP), which is a method of visualizing the effect of each soil variable on the model outcome (yhat) (Hastie et al., 2009). PDPs were built in R's 'DALEXtra' package (Maksymium et al., 2020). PDPs were built for the three most important soil variables from the variable importance plot.

Another primary objective was to estimate the area of continuous rice fields that have the soil features of a rotated rice field based on the three most important soil variables determined by the model. Rather than examining F_p, which can be heavily influenced by sampling schemes (i.e. up sampling or down sampling) (Woznicki et al., 2019), we developed a 'manual approach' using partial dependence data to determine thresholds, i.e. ranges for each of the three important soil variables supporting a higher likelihood of rotation. All continuous rice fields were then examined to determine how many of them met each of these three soil criteria, both individually and combined. For all the fields in each group, we computed the acreage (sum), and median predicted probability that the field is rotated (denoted as $yhat_m$ and shown as a percentage). This manual method allowed us to investigate how soil properties may act as a barrier to crop rotations and to quantify the acreage of continuous rice fields that have some of the properties of a rotated field.

3. Results and discussion

3.1. Continuous rice and rotated rice

According to our analysis, annual rice production area ranged from 227,000 to 161,000 ha, which is consistent with USDA reported acreage, and represents approximately 95% of California's rice growing area (USDA NASS, 2020). Across the eight major counties in the growing area, agreement between our data set and USDA NASS acreage was strong ($R^2 = 0.99$) and fell along the 1:1 line (Fig. 1).

58% of the study area was in continuous rice production (Fig. 2, Table 2). Rotation with rice occurred more in the southern portion of the Sacramento Valley (Colusa, Sutter, and Yolo counties), which is consistent with previous studies (Rosenberg et al., 2022; Carter et al., 1994). While the area under rotation area was considerably smaller than the area under continuous rice, there was high diversity of rotation schemes. Most fields under rotation transitioned from rice to another crop on two to three occasions. A small number of fields in the study area transitioned out of rice up to seven times during the 15-year study period, meaning these fields were planted back to rice every other year. According to our definition of continuous rice and rotated rice, which we used for binomial classification (Table 2), there were 155,640 ha in continuous rice (7550 fields) and 16,650 ha of rotated rice (470 fields) (Table 2). The other dominant field types in the region are 'rice + other' and 'annuals', which occupy 15.8% and 6.3% of the study area, respectively. Continuous rice production is common due to high prices for rice, consistent high yield, high efficiency of the production system (Hill et al., 2006), as well as farmer experiences or perceptions that these fields are not suitable for rotated crops (Rosenberg et al., 2022; Carter et al., 1994).

Approximately 3000 to 8000 ha were exchanged annually between rice and other crops including summer annuals, winter annuals, alfalfa, and grasses (Fig. 3). This exchange is dominated by summer annuals and

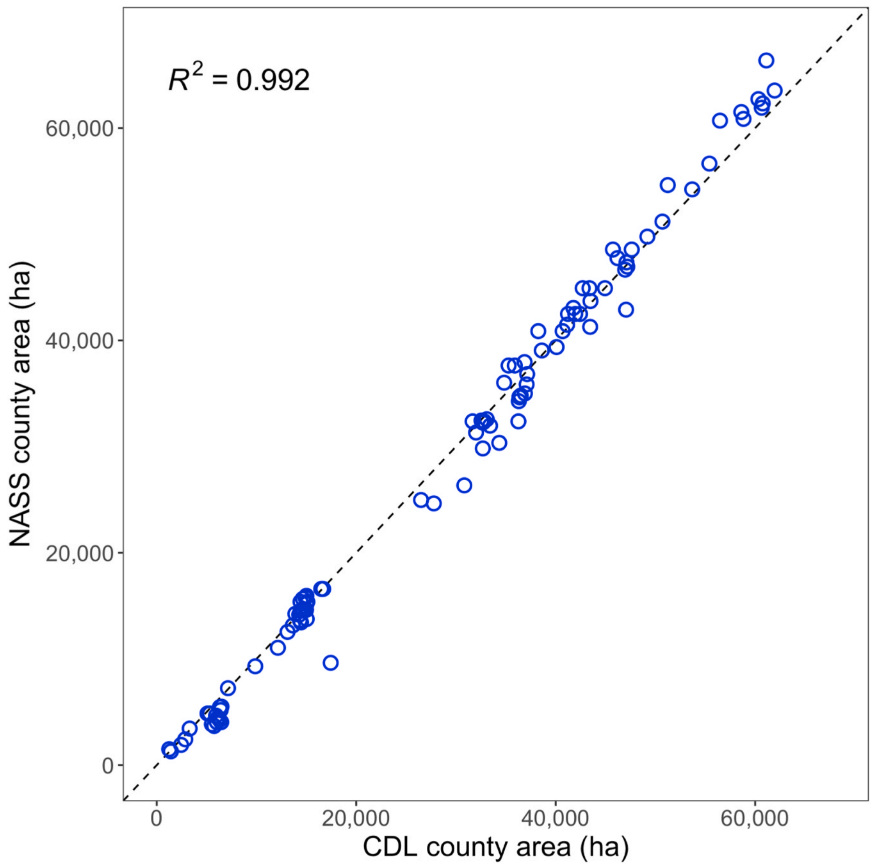

Fig. 1. County-level agreement between NASS planted area and reclassified CDL prediction of rice area. Agreement is measured by the coefficient of determinization, R^2, across the eight major rice growing counties in the study area from 2007 to 2021 (dashed line is 1:1).

winter annuals, on average 3030 ha and 2430 ha are exchanged between rice and summer annuals and rice and winter annuals each year, respectively. A smaller portion are rotated with grasses and alfalfa. The area in rotation also decreased throughout the study period ('08-'21) from roughly 8000 ha to 4000 ha. This decrease is likely because of specialization of agricultural operations. Interviews with rice farmers suggest that farmers that used to rotate have stopped due to market changes and labor and equipment requirements for alternative crops (Rosenberg et al., 2022). Furthermore, efficient irrigation methods such as subsurface drip are becoming increasingly widespread in non-flooded agricultural systems in the region due to political and economic motivations to maintain or improve production while using less water (Samuel Sandoval-Solis et al., 2022). Processing tomatoes, for example, have seen greater than 70% conversion to subsurface drip irrigation due to increased water savings (Ayars et al., 2015). These are semi-permanent crop specific systems that make it difficult to rotate with crops using different spacing or irrigation strategies, such as flooded rice.

Conversion to walnuts and almonds occupied 2.4% and 0.95% of the study area, respectively (Table 2). Conversion to perennial tree crops may be increasing due to increasing walnut and almond crop prices, despite increasing drought conditions (Gebremichael et al., 2021). This represents a shift from annual cropping to a system that is more permanent. Walnut and almond fields have comparable water use to rice (Cody and Johnson, 2015) but they must be watered annually to prevent tree mortality, meaning these fields cannot be fallowed during drought periods without significant economic loss.

3.2. Soil properties supporting crop rotations

The overall classification accuracy of continuous rice and rotated fields evaluated on the validation data set using random forest models was 93.9%. While our overall accuracy score was high, previous studies have suggested that, when there is class imbalance in the training data, other criteria for model evaluation should also be considered. After hyperparameter tuning and choosing an optimal sampling strategy, an F1 score of 0.62 was possible (precision was 0.52, recall was 0.76) (Table S2). Our F1 score of 0.62 suggests that the model performed well at binomial classification given the severity of class imbalance in the training data. A recall score of 0.76 suggests the model correctly classified 76% of the rotated rice area. This performance indicates that soil properties are a good predictor of crop rotation in the region, but that there may also be other important considerations not observed in this model, such as the economic, cultural, or logistical factors described by Rosenberg et al. (2022). Overall, this performance is consistent with other large-scale land cover modeling efforts using soil predictor variables (Woznicki et al., 2019; Olimb and Robinson, 2019; Sangwan and Merwade, 2015; Wing et al., 2017).

Another objective was to assess the relative importance of soil variables in predicting rotated fields in the rice growing area. Fig. 4 shows variable importance plots based on the mean decrease in AUC when each variable is permuted (Greenwell et al., 2018). The red dots show the variable importance score of the initial model execution. EC and Ksat had very similar variable importance scores, so we repeated the model execution 200 times and computed variable importance for each. The box and whisker plot shows the median and interquartile range of the 200 variable importance scores from this approach, while the violin plot shows the distribution of VI scores. Of the six soil properties included in our analysis, pH was the most influential, followed by EC and Ksat. The importance of each of these variables is discussed below. Soil order was the least important variable in the model.

While variable importance plots can help rank the influence of

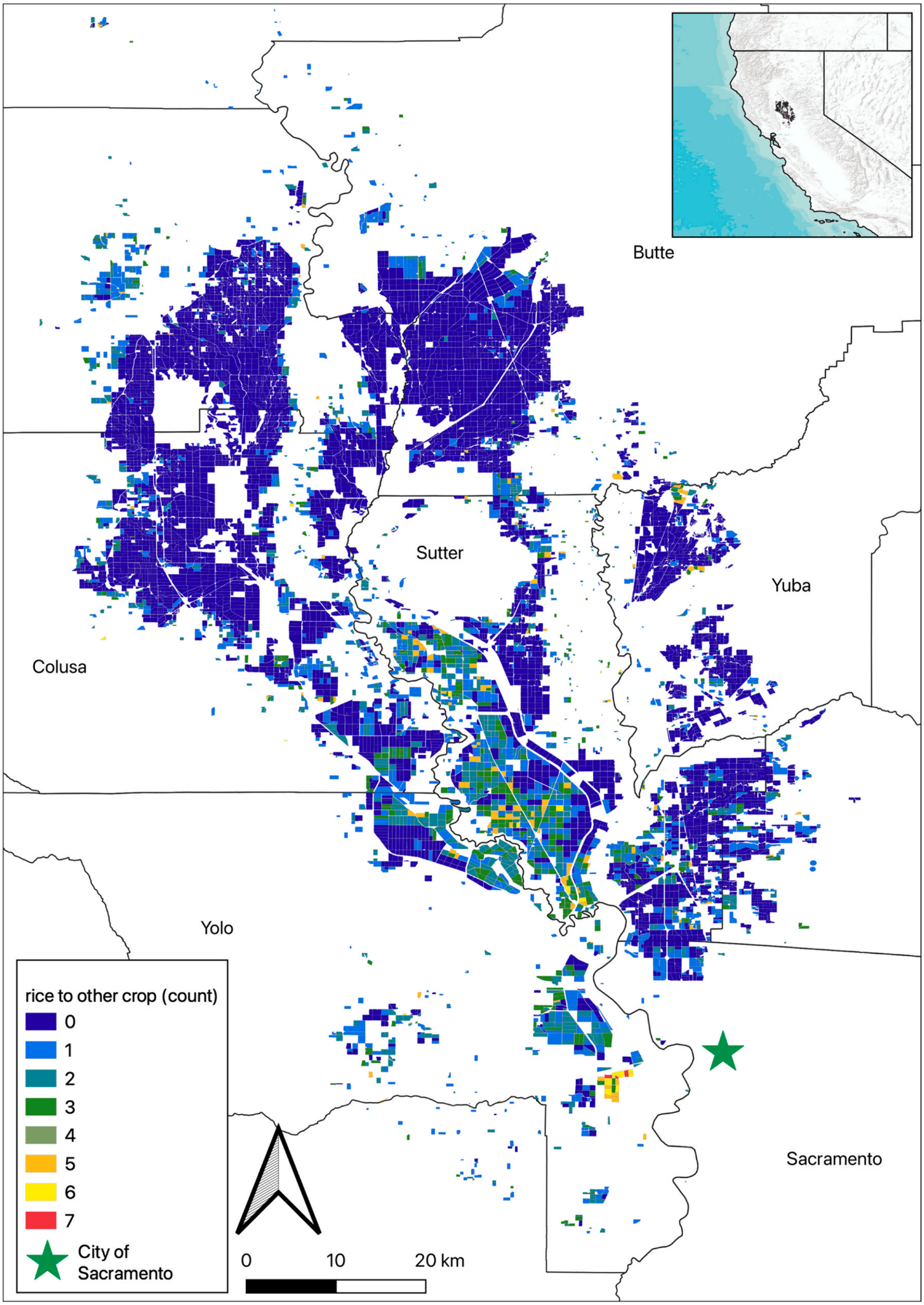

Fig. 2. Rice rotation frequency map showing the number of times a field changed from rice to alternate crop (i.e. annual crop, alfalfa, forage, pasture, etc.) in the 15-year data set. Fields converted to perennial trees were excluded. A count of 0 implies continuous rice, while a count of 6 or 7 implies the field is rotated annually.

Table 2
Area of major field types in the study area.

Field Type	Hectares	% Of Area	Criteria
Continuous Rice	155,640	57.9	Rice at least 12 out of 15 years, else fallow
Rotated Rice	16,650	6.2	Rice at least 7 out of 15 years, annual crop (summer annuals, winter annuals, grasses, and alfalfa) at least 2 out of 15 years, and rotated from rice to annual on at least two separate occasions
Annual	16,990	6.3	Annual crop (summer annual crop, winter annual crop, alfalfa, and grasses) at least 9 out of 15 years
Rice + other	42,370	15.8	Rice at least 9 out of 15 years
Fallow	3440	1.3	Fallow at least 9 out of 15 years
Walnut	6350	2.4	Walnut at least 4 out of 15 years
Almond	2550	0.95	Almond at least 4 out of 15 years
Alfalfa	650	0.25	Alfalfa at least 9 out of 15 years
Other	24,400	9.1	Doesn't meet any of the above criteria
Total:	268,950	100%	

different variables, they do not indicate the behavior (i.e. linear, monotonic, or more complex) or direction (i.e. positive or negative) of the interaction between an input feature and the target response (Hastie et al., 2009). To understand how the three most important variables (pH, EC and Ksat) influenced the likelihood of rotations, we used PDP to predict outcomes across each variable while marginalizing the model output over the distribution of the other features (Fig. 5a-c). Including data density curves with each PDP allows us to examine the distribution of values for each soil variable, which aids our interpretation of the PDP.

Soil pH in the study region ranged from five to greater than nine (Fig. 5a). The partial dependence data indicates that rotated fields are more likely between 6.5 and 8.0, while fields with pH less than 6.5 or greater than 8.0 are likely to be continuous rice. Annual crops, including rice, have highest productivity in neutral pH ranges (Havlin, 2020). Acidic soils can be managed with limestone, and alkaline soils can be managed with elemental sulfur, but both soil types can be costly and difficult to remediate, particularly alkaline soils (Fernandez and Hoeft, 2021). Soil flooding for rice production, however, results in the convergence of alkaline or acidic soil pH to neutral, allowing rice growers to maintain high yields without additional inputs (Sahrawat, 2012; Ponnamperuma and Kozlowski, 1984). Furthermore, flooding for rice production improves the availability of nutrients such as ammonium, phosphorous, potassium, and other exchangeable cations, which are mobilized in soil solution (Ponnamperuma, 1972).

Soil electrical conductivity (EC) is a metric of the salt content (salinity) in the soil, which is an indicator of mineral nutrients in the soil that can be quickly utilized by plants, and an indicator of salt ions in soil that could limit crop growth (Friedman, 2005). Most of the fields in the study area had relatively low EC (Fig. 5b). Fields with EC ranges between 0.5 and 1.5 had a higher likelihood of being rotated, while fields with higher EC were more likely to be in continuous rice. Low EC values could indicate that nutrients needed for plant growth are insufficient (Friedman, 2005), while high salinity has been shown to reduce agricultural productivity by causing reduced water uptake by plants (Machado and Serralheiro, 2017). High salinity can cause reduced osmotic pressure and ion imbalance as plants accumulate salt ions over time (Munns and Tester, 2008). Previous studies have indicated that the salinity threshold for field crops ranges from 1 to 2.5 dS m^{-1} (Ayers and Westcot, 1985; Maas and Grattan, 2015; Grattan et al., 2002; Machado

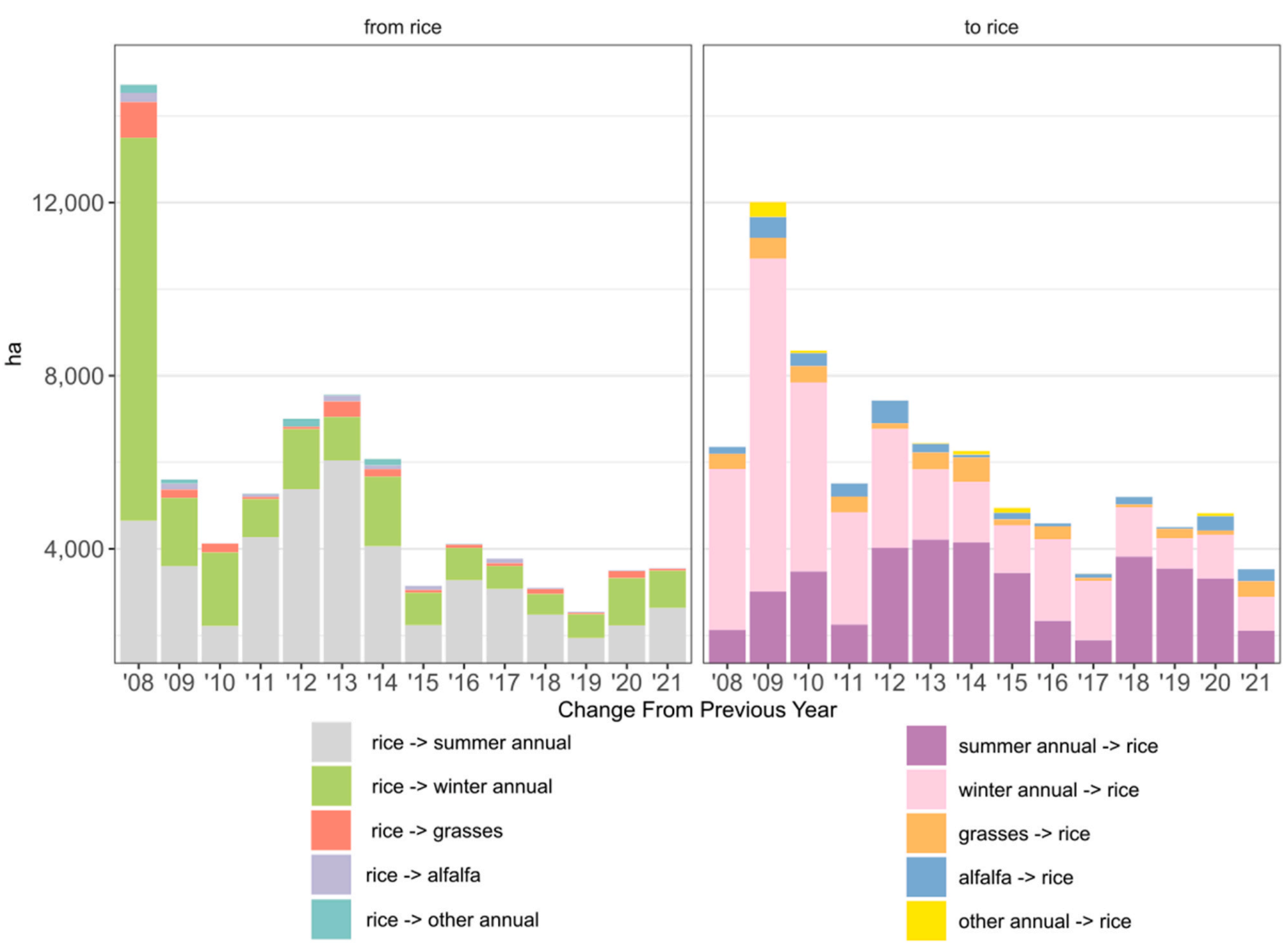

Fig. 3. Yearly area exchanged between rice and annuals, other annuals, grasses, and alfalfa from the 15-year data set.

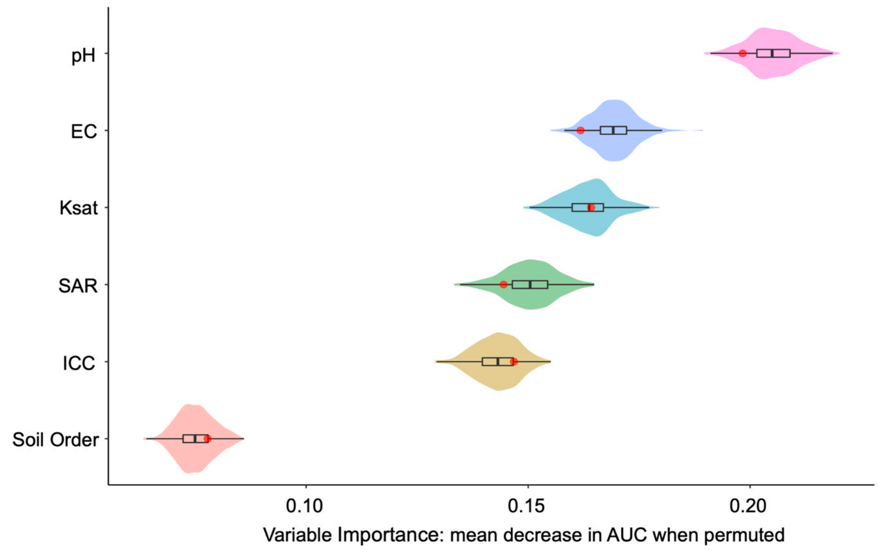

Fig. 4. Variable importance (VI) scores for the soil variables used in the random forest model based on the permutation method using AUC as the metric. To stabilize the VI scores, the model was repeated 200 times. The violin plot shows the distribution of the 200 VI scores, the left and right side of the box are the upper and lower quartiles, the vertical line inside the box is the median, and the whiskers extend to 1.5 times the interquartile range. The red dot shows the VI score from the initial model execution.

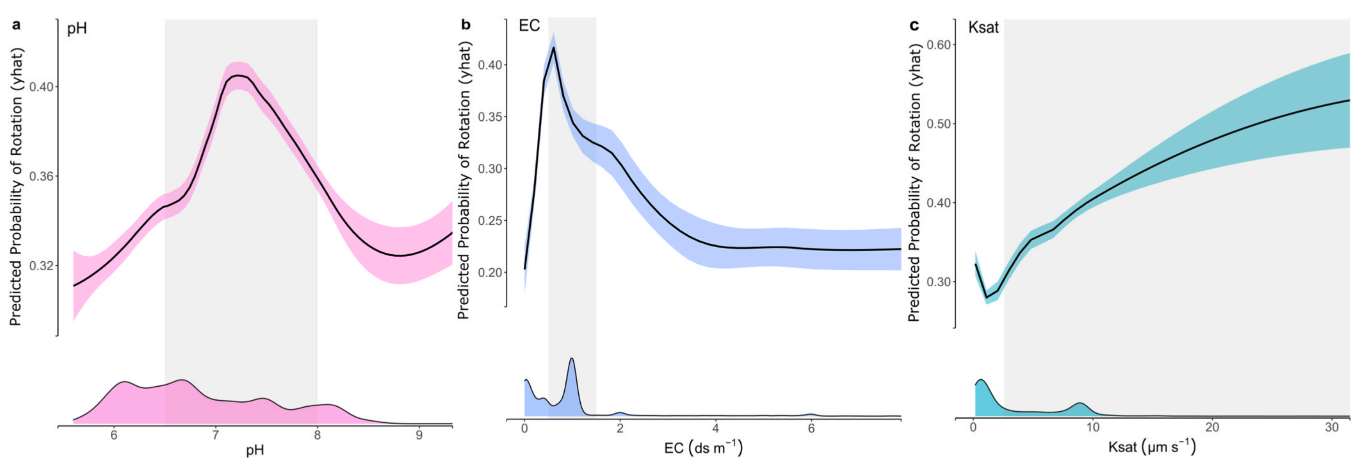

Fig. 5. Partial dependency plots of the three most important soil variables in the random forest model. The black lines show regression point estimates from spline fits on the partial dependence data, and the colored ribbons show bootstrapped 95% CIs. The histograms at the bottom show the distribution of fields in the training data set. The grey background indicates the manually selected thresholds for each variable where the probability of rotation (yhat) is higher, used to estimate proportion of the continuous rice region that could accommodate rotations.

and Serralheiro, 2017). In non-flooded agriculture, salinity can be managed by leaching salts, but in regions where there is also poor drainage this practice often requires installing costly drainage systems (Hanson et al., 2006). In rice production, however, high soil salinity can be managed with flood irrigation. For example, in the growing season, maintaining high water depth and allowing for tailwater drainage early in the season can help manage salinity (Marcos et al., 2018). In the winter season, flooding of rice fields, which is commonly done to decompose rice straw and to promote habitat for waterbirds in this region (Linquist et al., 2006), can lead to diffusion of salts into the water column, where it can potentially be percolated out of the root zone or exported in surface water drainage (Bachand et al., 2014).

Ksat represents how easily water can pass through saturated soil. Fields with low Ksat values will have little water loss to percolation and relatively high-water use efficiency for flooded crops (LaHue and Linquist, 2021). Ksat values in the study ranged from 0 to greater than 50 µm s^{-1} and most fields in the study area have Ksat values below 15 µm s^{-1} (Fig. 5c). Where Ksat is above 2 µm s^{-1}, fields were increasingly likely to be rotated. Ksat values can vary based on a range of soil and hydrologic factors including soil texture, soil structure, bulk density, field water height, and ground water elevation (Bouman et al., 2007; LaHue and Linquist, 2021). Many fields in the region have low Ksat either because they have very high clay content or because they have a cemented subsurface soil layer. In some parts of Glenn and Colusa counties, clay content was greater than 60%. These clayey soils are used for continuous rice because high clay content can lead to poor tilth, making it difficult to prepare a seed bed, and low water availability and poor aeration in non-flooded soils (Lund, 1959). A cemented subsurface soil layer can lead to poor root-ability and poor workability, which is also determinantal to plant growth for non-flooded crops (Dexter, 2004).

While sodium adsorption ratio (SAR) was not one of the three most important variables in the model (Fig. 4), some fields in the study area had sodic and saline-sodic soil properties, which likely limits their suitability for rotation. Sodic soils have high pH (> 8.5) and are also high in exchangeable sodium (Na^+) (>15%) (Sumner, 1993).

Saline-sodic soils have both high salinity and high Na^+. Sodium toxicity causes dispersion of soil particles leading to soil degradation and poor tilth, making them detrimental to growth of most plants (Qadir and Oster, 2004). Ameliorating sodic soils requires increasing calcium (Ca^{2+}) to replace Na^+ on the exchange, then leaching with excessive irrigation. This process is difficult, costly, and time consuming, especially in soils with low Ksat limiting drainage capacity (Qadir and Oster, 2004). Water management in rice, however, can help the crop tolerate sodic and saline-sodic soil properties with irrigation techniques such as maintaining flooded conditions and excess drainage (Munns and Tester, 2008).

3.3. Feasibility of expanding crop rotations

We used the partial dependence data to determine ranges of the three soil properties (pH, EC, Ksat) that have a higher likelihood of rotation given historical land use decisions in this region (grey shading in Fig. 5a-c). Fields with pH between 6.5 and 8, EC values between 0.5 and 1.5 ds m^{-1}, and Ksat values > 2 μm s^{-1} had a higher likelihood of rotation. Most data lay within these ranges for pH and EC, however most of the fields in our study region have low Ksat (Fig. 5a-c).

We examined all the continuous rice fields in our study that meet each of these soil criteria to determine the extent and location of similar soil properties associated with rotations (Table 3). Around 69,000 ha (47%) of the continuous rice region has pH values between 6.5 and 8.0. These fields are mostly in the northern and central portions of the study region (Butte, Glen, and Sutter counties) (Fig. 6a). Meanwhile 73,000 ha (50% of the continuous rice fields) had EC values between 0.5 and 1.5 dS m^{-1}. These fields are mostly in the center and west of the region (Sutter and Colusa counties) (Fig. 6b). Lastly 55,000 ha (37% of the continuous rice area) had Ksat > 2 μm s^{-1}. These fields are in the east (Yuba and Sutter counties) (Fig. 6c). For all the continuous rice fields in each group, we computed the median predicted probability of rotation ($yhat_m$). Continuous rice fields that met the pH criteria had a $yhat_m$ value of 32%, while fields meeting EC and Ksat criteria had a median $yhat_m$ of 28% and 21%, respectively.

Combining these thresholds allowed us to examine how multiple soil factors affect the scope for agroecological diversification in this rice-based system (Table 3). A total of 38,720 ha met the combined pH and EC criteria. This accounts for about 26% of the continuous rice fields, and these fields have a 50.7% median predicted probability of rotation. Only 11% (16,710 ha) of the continuous rice area met all three of the combined criteria. These fields have $yhat_m$ of 54.1%. Most of these fields are nearby and to the east of current rotated rice fields, which are in the southern and central portion of the rice growing region (Sutter, Yolo, and southeastern Colusa Counties; Fig. 6d), a region known for having a high diversity of agricultural systems including continuous rice (Carter et al., 1994; Rosenberg et al., 2022). The area that meets all three criteria is approximately 12,000 ha larger than the size of the decrease in rotation area over the past ten years (~4000 ha) (Fig. 3). The remaining area under continuous rice production that does not meet the combined pH and EC criteria (74%), or all three combined criteria (89%) is 115,170 ha and 138,910 ha, respectively. This finding is comparable to the Carter et al. (1994) report on the Sacramento Valley rice area which stated that "on at least [120,000 ha] (…) it would be very difficult under any circumstances to produce another crop".

Table 3
Soil criteria where fields are more likely to be rotated, the area that meet the criteria, and the median predicted probability of rotation for all fields that meet the criteria. Criteria were created manually based on partial dependence plots in Fig. 3.

Criteria	Area that meets the criteria (ha)	Proportion of rice area (%)	Mean predicted probability (%) that the field is rotated
Single Criteria			
pH 6.5–8	68,970	46.8	32.0
EC 0.5–1.5 (dS m^{-1})	73,590	50.0	28.0
Ksat > 2 (μm s^{-1})	54,730	37.1	21.1
Two criteria met:			
pH + EC	38,720	26.3	50.7
pH + Ksat	26,910	18.3	34.0
EC + Ksat	24,870	18.9	40.2
All three criteria met:			
pH + EC + Ksat	16,710	11.3	54.1

While only 11% of the continuous rice area meets all three criteria, incentivizing rotation in this area could help manage weeds concurrently with reduced water use while maintaining agricultural revenue for farmers. In California, pesticide regulations have limited the number of herbicides available to farmers and have limited how the existing herbicides can be applied, leaving limited options for chemical weed management aside from increasing the number of herbicide applications, which has increased herbicide resistance challenges (Hill et al., 2006; Rosenberg et al., 2022). Crop rotation allows for integrated weed management approaches including aerobic irrigation and cultivation techniques and the use of herbicides with different modes of action, which can support weed control and limit herbicide resistance (Beckie et al., 2004; Kayeke et al., 2017; Vencill et al., 2012).

Currently, due to severe drought conditions in the region, many rice farmers are forced to fallow their fields in water districts with limited access to water rights (Pancorbo, 2023; Medellin-Azuara, 2022). Fallowing, however, is not an ideal solution to water scarcity as fallow fields inherently do not provide a harvestable cash crop or other ecosystem services such as wildlife habitat, and leaving bare soil causes soil erosion and degradation due to wind and water exposure (Pimentel and Burgess, 2013; Kaspar and Singer, 2015; Wendt et al., 1986). Alternative summer annual crops in the region such as processing tomatoes or dry beans use 30% and 5%, respectively, of the annual water requirements for flooded rice (Cody and Johnson, 2015). Winter annual crops such as wheat, oats, and rye, are predominantly rainfed and require little to no irrigation water unless it is a drought year. As droughts increase in severity (Cayan et al., 2010), rotating rice with less water intensive crops, where possible, could help limit agricultural demand for water across the region while maintaining agricultural productivity.

3.4. Limitations of the study

Agricultural systems are coupled human-natural systems that depend on complex food supply chains and international trade (Liu et al., 2007). There are numerous drivers fueling agricultural decision making including soil type, economic variables, socio-cultural factors, government policy, climate change, and distance to networks and terminal markets for agricultural products (Flora et al., 2019; Rosenberg et al., 2022). This study focuses only on the role of soil properties. To do so, this study utilizes soil survey (SSURGO) information, which is only one of multiple options for investigating soil barriers to rotation. SSURGO information in California offers a comprehensive, detailed spatial assessment of soil variables and is an excellent resource for local and regional land use planning. However, SSURGO has a few key limitations. SSURGO does not always integrate land use information and changes to soil management over time, it has variable spatial detail between soil surveys of different vintage, and there are sometimes artificial discontinuities at political boundaries (Li, 2012; Du, 2015; Gatzke, 2011; Subburayalu, 2014; Nauman and Thompson, 2014). Results of our study align with Rosenberg et al. (2022) and other's (Hill et al., 2006; Carter et al., 1994); however, SSURGO data does not replace on-the-ground soil sampling or field-based experiments that test the efficacy of planting

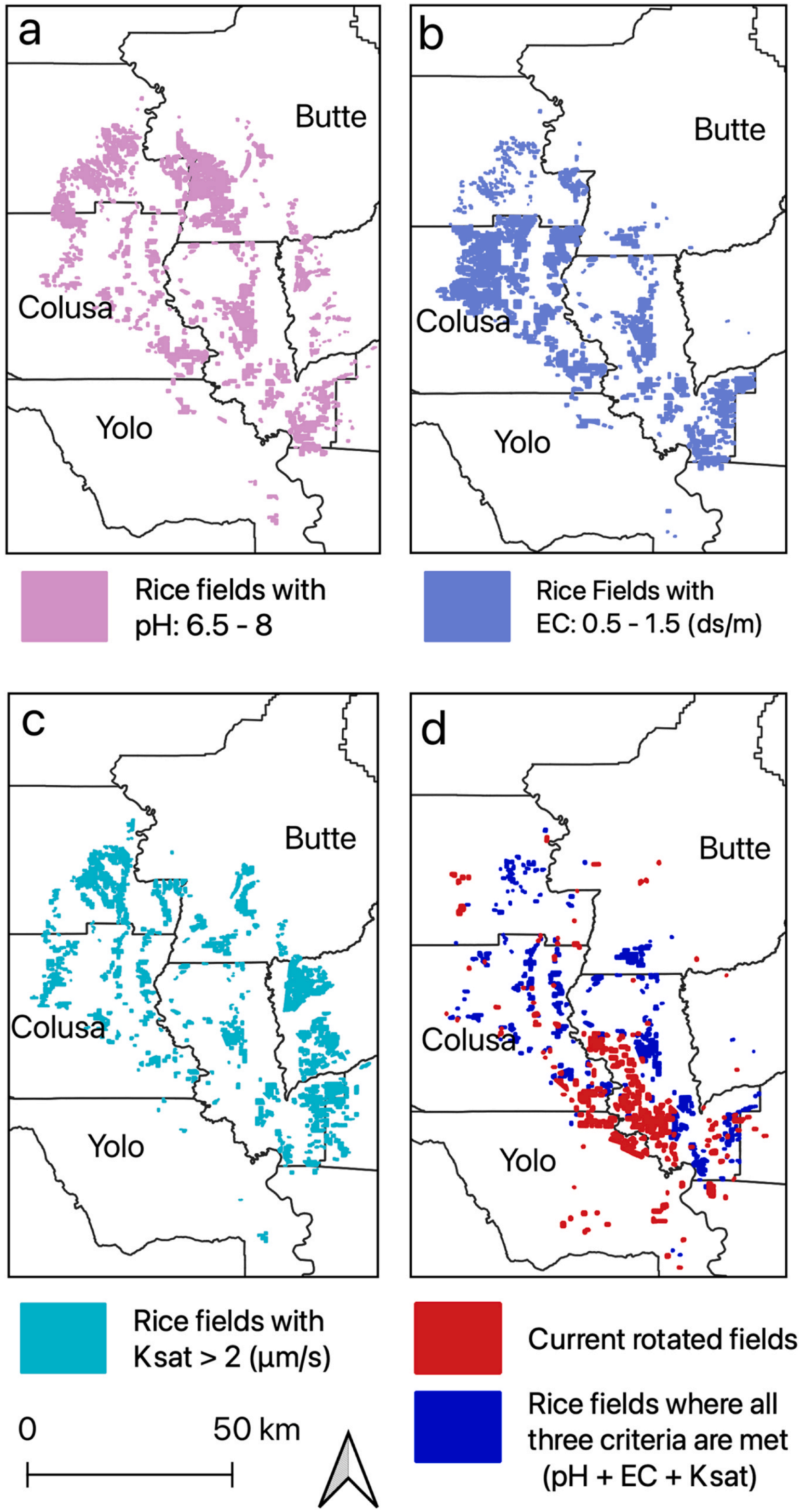

Fig. 6. Map a-c show continuous rice fields that meet each of the three soil criteria for rotation (pH, EC, and Ksat). Map d shows current rotated rice fields (red) and continuous rice fields that meet all three of the soil criteria.

row crops in unfavorable soils.

4. Conclusions

This research uses satellite-derived land cover information and soil survey data to examine the feasibility of crop rotation in California's Sacramento Valley, a region with a history of continuous rice production and growing sustainability challenges. Our analysis shows that rotation occurs in a limited portion of the region, and that there is a high diversity of rotation schemes including rotation with summer annuals, winter annuals, alfalfa, and grasses. By comparing the soil properties of continuous rice fields to rotated rice fields using a random forest model, our analysis suggests that chemical and physical soil properties such as alkalinity, salinity, and low saturated hydraulic conductivity are key variables that may limit the potential for crop rotations to be easily implemented in the region. This research highlights the importance of including biophysical considerations such as soil properties into broader efforts to diversify modern agricultural systems. Research and extension efforts to implement crop rotation practice in the region should focus on identifying pathways to overcome soil barriers alongside access to markets and equipment for rotated crops. Field scale experiments may be necessary to better understand potential rotated crops that can tolerate the soil conditions in this region while providing water savings, weed management benefits, and economic value.

Funding

This work was supported by the California Rice Research Board [grant number RR19-7 (RM-11).

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

We would like to thank Sara Rosenberg for helping inspire this research and offering feedback based on her survey of California rice farmers, Scott Devine for his guidance on working with the SSURGO data set, Matthieu Stigler for being an instrumental mentor of geospatial analysis in RStudio and Google Earth Engine, and the Linquist Lab and Pittelkow Lab for their ongoing support. We would like to thank the California Rice Research Board and the University of California, Davis Department of Plant Scf project. We would like to acknowledge the Department of Plant Sciences, UC Davis for the award of a GSR scholarship funded by endowments, particularly the James Monroe McDonald Endowment, administered by UCANR.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.agee.2023.108753.

References

Altieri, Miguel A., et al., 2015. Agroecology and the design of climate change-resilient farming systems. Agronom.Sustain. Dev. 869–890. https://doi.org/10.1007/S13593-015-0285-2/TABLES/2.

Arlot, Sylvain, and Alain Celisse. "A Survey of Cross-Validation Procedures for Model Selection." Https://Doi.Org/10.1214/09-SS054, vol. 4, no. none, Amer. Statist. Assoc., the Bernoulli Soc., the Inst. Math. Statist., and the Statist. Soc. Canada, Jan. 2010, pp. 40–79, doi:10.1214/09-SS054.

Awika, Joseph M., 2011. Major cereal grains production and use around the world. ACS Symposium Ser. 1–13. https://doi.org/10.1021/BK-2011-1089.CH001/ASSET/IMAGES/LARGE/BK-2011-00480Z_G003.JPEG.

Ayars, J.E., et al., 2015. Subsurface drip irrigation in california-here to stay? Agricult. Water Manag. 39–47. https://doi.org/10.1016/j.agwat.2015.01.001.

Ayers, R.S., and D.W. Westcot. "Water Quality for Agriculture. FAO Irrigation and Drainage." FAO Irrigation and Drainage Paper, vol. No.29, 1985.

Bachand, P.A.M., et al., 2014. Differentiating transpiration from evaporation in seasonal agricultural wetlands and the link to advective fluxes in the root zone. Sci. Total Environ. 232–248. https://doi.org/10.1016/j.scitotenv.2013.11.026.

Baste, Ivar A., et al., 2021. Making Peace with Nature: A Scientific Blueprint to Tackle the Climate, Biodiversity and Pollution Emergencies. United Nations Environment Programme.

Beckie, Hugh J., et al., 2004. Management practices influencing herbicide resistance in wild oat. Weed Technol. 853–859. https://doi.org/10.1614/WT-03-124R.

Belgiu, Mariana, Drăgu, Lucian, 2016. Random forest in remote sensing: a review of applications and future directions. ISPRS J. Photogramm. Remote Sens. 24–31. https://doi.org/10.1016/J.ISPRSJPRS.2016.01.011.

Bocinsky, R.Kyle, 2019. FedData: functions to automate downloading geospatial data available from several federated data sources. R. Package Version 2 (*5*), *7*.

Bouman, B.A.M., et al., 2007. Rice Water 187–237. https://doi.org/10.1016/S0065-2113(04)92004-4.

Brandon M. Greenwell, Bradley C. Boehmke, Andrew J. McCarthy.Greenwell, Brandon M., et al. A Simple and Effective Model-Based Variable Importance Measure. 2018.

Breiman, Leo, 2001. Random Forests. (Oct). In: Machine Learning 2001 45:1, vol. 45 Springer, pp. 5–32. https://doi.org/10.1023/A:1010933404324 (Oct).

Brim-DeForest, Whitney B., et al., 2017. Weed community dynamics and system productivity in alternative irrigation systems in California rice. Weed Sci. 177–188. https://doi.org/10.1614/WS-D-16-00064.1.

Bronson, K.F., et al., 1998. Soil carbon dynamics in different cropping systems in principal ecoregions of Asia. Manag. Carbon Sequestration Soil 35–57.

Cabell, Joshua F., Oelofse, Myles, 2012. An indicator framework for assessing agroecosystem resilience. Society Vol. 17 (1). https://doi.org/10.5751/ES-04666-170118.

Carter, H., et al. Maintaining the Competetive Edge in California's Rice Industry. University of California, Agricultural Issues Center, 1994, https://aic.ucdavis.edu/publications/oldanrpubs/rice.pdf.

Cassman, K.G., Pingali, P.L., 1995. Intensification of irrigated rice systems: learning from the past to meet future Challenges. GeoJournal 35 (3). https://doi.org/10.1007/BF00989137.

Cassman, Kenneth G., Grassini, Patricio, 2020. A global perspective on sustainable intensification research. Nat. Sustain. 3 (4) https://doi.org/10.1038/s41893-020-0507-8.

Cayan, Daniel R., et al., 2010. Future dryness in the Southwest US and the hydrology of the early 21st century drought. Proc. Natl. Acad. Sci. USA vol. 107 (50), 21271–21276. https://doi.org/10.1073/pnas.0912391107.

Chen, C., et al., 2004. Using random forest to learn imbalanced data. Univ. Calif. Berkeley vol. 110 (24), 1–12. ⟨https://scholar.google.com/scholar?cluster=18437624849810011430&hl=en&as_sdt=2005&sciodt=0,5⟩.

Cody, Betsy A., and Renee Johnson. California Agricultural Production and Irrigated Water Use Renée Johnson Specialist in Agricultural Policy Specialist in Natural Resources Policy. 2015, pp. 1–28, www.crs.gov.

CropScape - NASS CDL Program. ⟨https://nassgeodata.gmu.edu/CropScape/⟩. Accessed 23 Nov. 2022.

De Leon, Teresa B., et al., 2019. Genetic variation and possible origins of weedy rice found in California. Ecol. Evol. 5835–5848. https://doi.org/10.1002/ECE3.5167.

Deines, Jillian M., et al., 2020. Transitions from irrigated to dryland agriculture in the ogallala aquifer: land use suitability and regional economic impacts. Agric. Water Manag. vol. 233 (October 2019) https://doi.org/10.1016/j.agwat.2020.106061.

Department of Water Resources. Statewide Crop Mapping - Datasets - California Natural Resources Agency Open Data. ⟨https://data.cnra.ca.gov/dataset/statewide-crop-mapping⟩. Accessed 31 Oct. 2022.

Dexter, A.R., 2004. Soil physical quality: Part I. theory, effects of soil texture, density, and organic matter, and effects on root growth (June). Geoderma vol. 120 (3–4), 201–214. https://doi.org/10.1016/J.GEODERMA.2003.09.004.

Du, Fei, et al., 2015. Soil property variation mapping through data mining of soil category maps. Hydrol. Process. 2491–2503. https://doi.org/10.1002/HYP.10383.

F.A.O. "Food and Agriculture Organizations of the United Nations-Statistic Division." Faostat, 2019.

Fernandez, Fabian G., and Robert G. Hoeft. Managing Soil PH and Crop Nutrients. Illinois Agronomy Textbook, 2021, https://extension.illinois.edu/global/agronomy-handbook.

Friedman, Shmulik P., 2005. Soil properties influencing apparent electrical conductivity: a review. Comput. Electron. Agricult. 45–70. https://doi.org/10.1016/J.COMPAG.2004.11.001.

Gatzke, Sarah E., et al., 2011. Aggregation strategies for SSURGO data: effects on SWAT soil inputs and hydrologic Output. Soil Science Soc. Am. J. 1908–1921. https://doi.org/10.2136/SSSAJ2010.0418.

Gebremichael, Mekonnen, et al., 2021. What drives crop land use change during multi-year droughts in California ' s Central Valley. Prices. Or. Concern Water.

Grattan, Stephen R., et al. 1837. no. December, 2002, pp. 189–195.

Hanak, Ellen, et al. Water and the Future of the San Joaquin Valley OVERVIEW. 2019.

Hanson, Bradley D., et al., 2014. Herbicide-resistant weeds challenge some signature cropping systems. California Agricult. 142–152. https://doi.org/10.3733/CA.V068N04P142.

Hanson, Blaine R., et al. Agricultural Salinity and Drainage. 2006.

T. Hastie, R. Tibshirani, J. Friedman. The Elements of Statistical Learning. 2nd ed., Springer, 2009.

Havlin, John L. "Soil: Fertility and Nutrient Management." Landscape and Land Capacity, CRC Press, May 2020, pp. 251–265, doi:10.1201/9780429445552–34.

Hill, J.E., et al., 2006. The California rice cropping system: agronomic and natural resource issues for long-term sustainability. Paddy Water Environ. vol. 4 (1), 13–19. https://doi.org/10.1007/s10333-005-0026-2.

Horton, Peter, et al., 2021. Technologies to deliver food and climate security through agriculture. Nat. Plants 2021 7 (*3*), 250–255. https://doi.org/10.1038/s41477-021-00877-2.

Jan L. Flora et al.L.Flora Jan et al. Rural Communities Study Guide 2011.4324/9780429305009.

Janitza, Silke, et al., 2013. An AUC-based permutation variable importance measure for random Forests. BMC Bioinform. 1–11. https://doi.org/10.1186/1471-2105-14-119/FIGURES/5.

Kaspar, T.C., J. W, 2015. The use of cover crops to manage soil. Soil Manag. Build. Stable Base Agricult. 321–337. https://doi.org/10.2136/2011.SOILMANAGEMENT.C21.

Kayeke, Mohamed J., et al., 2017. Reducing risk of weed infestation and labor burden of weed management in cropping systems. Smart Technol. Sustain. Smallholder Agricult. Upscal. Dev. Countries 123–143. https://doi.org/10.1016/B978-0-12-810521-4.00006-2.

LaHue, Gabriel T., Linquist, Bruce A., 2021. The contribution of percolation to water balances in water-seeded rice systems. Agricult. Water Manag. https://doi.org/10.1016/J.AGWAT.2020.106445.

Lark, Tyler J., et al., 2017. Measuring land-use aof agriculture's cropland data layer: cautions and recommendations. Int. J. Appl. Earth Obs. Geoinf. https://doi.org/10.1016/j.jag.2017.06.007.

Li, Runkui, et al., 2012. Effects of spatial aggregation of soil spatial information on watershed hydrological modelling. Hydrol. Process. 1390–1404. https://doi.org/10.1002/HYP.8277.

Liaw, Anddy; Wiener, Matthew. "Classification and Regression by RandomForest." R News, vol. Vol. 2/3, no. ISSN 1609–3631, 2002.

Linquist, Bruce A., Brouder, Sylvie M., Hill, James E., 2006. Winter straw and water management effects on soil nitrogen dynamics in California rice systems. Agronomy Journal 98 (4), 1050–1059.

Liu, Jianguo, et al., 2007. Complexity of coupled human and natural systems. Science 1513–1516. https://doi.org/10.1126/SCIENCE.1144004/SUPPL_FILE/LIU.SOM.PDF.

Lund, Zane F., 1959. Available water-holding capacity of alluvial soils in Louisiana1. Soil Sci. Soc. Am. J. 1. https://doi.org/10.2136/SSSAJ1959.03615995002300010009X.

Maas, E.V., and S.R. Grattan. Crop Yields as Affected by Salinity. Oct. 2015, pp. 55–108, doi:10.2134/AGRONMONOGR38.C3.

Machado, Almeida, Rui Manuel, Serralheiro, Ricardo Paulo, 2017. Soil salinity: effect on vegetable crop growth. management practices to prevent and mitigate soil salinization. Horticulturae vol. 3 (2). https://doi.org/10.3390/horticulturae3020030.

Maksymium, S., et al., 2020. Landscape of R packages for explainable artificial intelligence. ArXiv 6–15. ⟨https://arxiv.org/abs/2009.13248⟩.

Marcos, Mathias, et al., 2018. Spatio-temporal salinity dynamics and yield response of rice in water-seeded rice fields. Agricult. Water Manag. 37–46. https://doi.org/10.1016/j.agwat.2017.09.016.

Medellin-Azuara, Josue, et al Economic Impacts of the 2022 Drought on California Agriculture Preliminary Report Prepared for: The California Department of Food and Agriculture. 2.

Munns, Rana, Tester, Mark, 2008. Mechanisms of salinity tolerance. Annu. Rev. Plant Biol. vol. 59, 651–681. https://doi.org/10.1146/annurev.arplant.59.032607.092911.

Nauman, Travis W., Thompson, James A., 2014. Semi-automated disaggregation of conventional soil maps using knowledge driven data mining and classification trees. Geoderma 385–399. https://doi.org/10.1016/J.GEODERMA.2013.08.024.

NRCS Soils. ⟨https://www.nrcs.usda.gov/wps/portal/nrcs/detail/soils/home/?cid=nrcs142p2_053368⟩. Accessed 27 July 2022.

Olimb, Sarah K., Robinson, Barry, 2019. Grass to grain: probabilistic modeling of agricultural conversion in the North American great plains. Ecol. Indicat. 237–245. https://doi.org/10.1016/J.ECOLIND.2019.02.042.

Pampolino, Mirasol F., et al., 2008. Soil carbon and nitrogen changes in long-term continuous lowland rice cropping. Soil Sci. Soc. Am. J. 798–807. https://doi.org/10.2136/SSSAJ2006.0334.

Pancorbo, J.L., et al., 2023. Drought impact on cropland use monitored with AVIRIS imagery in Central Valley, California. Sci. Total Environ. https://doi.org/10.1016/J.SCITOTENV.2022.160198.

Pebesma, E., 2018. Simple features for R: standardized support for spatial vector data. R. J. 10 (*1*), *439–446*,. https://doi.org/10.32614/RJ-2018-009.

Pimentel, David, Burgess, Michael, 2013. Soil erosion threatens food production. Agriculture *443–463*. https://doi.org/10.3390/AGRICULTURE3030443.

Ploton, Pierre, et al., 2020. Spatial validation reveals poor predictive performance of large-scale ecological mapping models. Nat. Commun. 11 (1), 1–11. https://doi.org/10.1038/s41467-020-18321-y.

Ponnamperuma, F.N., 1972. The chemistry of submerged soils. Adv. Agron. 29–96. https://doi.org/10.1016/S0065-2113(08)60633-1.

Ponnamperuma, F.N., Kozlowski, T.T., 1984. Flooding and plant growth. Acad. Press vol. 9.

Prasad, Anantha, M., et al., 2006. Newer classification and regression tree techniques: bagging and random forests for ecological prediction. Ecosystems 181–199. https://doi.org/10.1007/S10021-005-0054-1/TABLES/5.

Qadir, M., Oster, J.D., 2004. Crop and irrigation management strategies for saline-sodic soils and waters aimed at environmentally sustainable agriculture. Sci. Total Environ. 1–19. https://doi.org/10.1016/J.SCITOTENV.2003.10.012.

Rashford, Benjamin S., et al., 2011. Agricultural land-use change in prairie canada: implications for wetland and waterfowl habitat conservation (June). Can. J. Agric. Econ. vol. 59 (2), 185–205. https://doi.org/10.1111/J.1744-7976.2010.01212.X.

Roberts, David R., et al., 2017. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography 913–929. https://doi.org/10.1111/ECOG.02881.

Rosenberg, Sara, et al., 2022. Crop rotations in California rice systems: assessment of barriers and opportunities. Front. Agron. vol. 4 (June) https://doi.org/10.3389/fagro.2022.806572.

Sahrawat, Kanwar Lal., 2012. Soil Fertility in Flooded and Non-Flooded Irrigated Rice Systems. In: Http://Dx.Doi.Org/10.1080/03650340.2010.522993, 58. Taylor & Francis, pp. 423–436. https://doi.org/10.1080/03650340.2010.522993.

Samuel Sandoval-Solis, et al. Application Efficiency. ⟨http://watermanagement.ucdavis.edu/research/application-efficiency/⟩. Accessed 28 Nov. 2022.

Sangwan, Nikhil, Merwade, Venkatesh, 2015. A faster and economical approach to floodplain mapping Using Soil Information. JAWRA J. Am. Water Resourc. Assoc. 1286–1304. https://doi.org/10.1111/1752-1688.12306.

Seo, B., et al., 2014. Deriving a per-field land use and land cover map in an agricultural mosaic catchment. Earth Syst. Sci. Data 339–352. https://doi.org/10.5194/ESSD-6-339-2014.

Silge, Julia, 2021. Spatialsample: spatial resampling infrastructure. R. Package Version 0 (*1*), *0*.

Socolar, Yvonne, et al., 2021. Biophysical and policy factors predict simplified crop rotations in the US Midwest. Environ. Res. Lett., 054045 https://doi.org/10.1088/1748-9326/ABF9CA.

Staff, Soil Survey. Soil Survey Field and Laboratory Methods Manual. Soil Survey Investigations Report No. 51 Version 2.0. Lincoln, NE: USDA-NRCS, 2014.

Subburayalu, S.K., et al., 2014. Disaggregation of component soil series on an Ohio County soil survey map using possibilistic decision trees. Geoderma 334–345. https://doi.org/10.1016/J.GEODERMA.2013.08.018.

Sumner, Malcolm E., 1993. Sodic soils - new perspectives. Soil Res. 683–750. https://doi.org/10.1071/SR9930683.

USDA - National Agricultural Statistics Service - California. ⟨https://www.nass.usda.gov/Statistics_by_State/California/⟩. Accessed 12 Feb. 2020.

USDA/NASS. 2020, ⟨https://quickstats.nass.usda.gov/⟩.

Vencill, William K., et al., 2012. Herbicide resistance: toward an understanding of resistance development and the impact of herbicide-resistant crops. Weed Sci. 2–30. https://doi.org/10.1614/WS-D-11-00206.1.

Verburg, Peter H., et al., 2011. Challenges in using land use and land cover data for global change studies (Feb.). Glob. Change Biol. vol. 17 (2), 974–989. https://doi.org/10.1111/J.1365-2486.2010.02307.X.

Waha, Katharina, et al., 2020. Multiple cropping systems of the world and the potential for increasing cropping intensity. Glob. Environ. Change. https://doi.org/10.1016/J.GLOENVCHA.2020.102131.

Wendt, R.C., et al., 1986. Variability of runoff and soil loss from fallow experimental plots. Soil Sci. Soc. Am. J. 730–736. https://doi.org/10.2136/SSSAJ1986.03615995005000030035X.

Wing, Oliver E.J., et al., . 2017. Validation of a 30 m resolution flood hazard model of the conterminous United States. Water Resourc. Res. 7968–7986. https://doi.org/10.1002/2017WR020917.

Woznicki, S.A., Baynes, J., Panlasigui, S., Mehaffey, M., Neale, A., 2019. Development of a spatially complete floodplain map of the conterminous united states using random forest. Sci. Total Environ. 942–953. https://doi.org/10.1016/j.scitotenv.2018.07.353.

Yuan, Shen, et al., 2021. Sustainable intensification for a larger global rice bowl. Nat. Commun. https://doi.org/10.1038/s41467-021-27424-z.