ORIGINAL PAPER

Mapping almond stem water potential using machine learning and multispectral imagery

Peter Savchik[1] · Mallika Nocco[2,3] · Isaya Kisekka[1,2]

© The Author(s) 2024

Abstract

Almonds are a major crop in California which produces 80% of all the world's almonds. Widespread drought and strict groundwater regulations pose significant challenges to growers. Irrigation regimes based on observed crop water status can help to optimize water use efficiency, but consistent and accurate measurement of water status can prove challenging. In almonds, crop water status is best represented by midday stem water potential measured using a pressure chamber, which despite its accuracy is impractical for growers to measure on a regular basis. This study aimed to use machine learning (ML) models to predict stem water potential in an almond orchard based on canopy spectral reflectance, soil moisture, and daily evapotranspiration. Both artificial neural network and random forest models were trained and used to produce high-resolution spatial maps of stem water potential covering the entire orchard. Also, for each ML model type, one model was trained to predict raw stem water potential values, while another was trained to predict baseline-adjusted values. Together, all models resulted in an average coefficient of correlation of $R^2 = 0.73$ and an average root mean squared error (RMSE) of 2.5 bars. Prediction accuracy decreased significantly when models were expanded to spatial maps ($R^2 = 0.33$, RMSE = 3.31 [avg]). These results indicate that both artificial neural networks and random forest frameworks can be used to predict stem water potential, but both approaches were unable to fully account for the spatial variability observed throughout the orchard. Overall, the most accurate maps were produced by the random forest model (raw stem water potential $R^2 = 0.47$, RMSE = 2.71). The ability to predict stem water potential spatially can aid in the implementation of variable rate irrigation. Future studies should attempt to train similar models with larger datasets and develop a simpler faster workflow for producing stress predictions from field measurements.

Introduction

Almonds are a major crop in the state of California, contributing an estimated $11 billion annually to the state's GDP (Sumner 2014). Along with their profitability, almonds are a highly water-intensive crop, requiring almost year-round irrigation to maintain optimal quality and yield at harvest. Recent droughts throughout the state coupled with groundwater regulations imposed through California's Sustainable Groundwater Management Act have created a major water scarcity problem for this lucrative industry, with nearly every almond producer in the state facing water supply challenges. Thus, there is a high demand for increased water use efficiency in almond production.

Regulated deficit irrigation, applying just the right amount of water or slightly less water at less sensitive growth stages, has been shown to greatly increase water use efficiency with only moderate reductions in quality and yield (Drechsler and Kisekka 2022; DeJonge et al. 2015). However, proper implementation of regulated deficit irrigation requires an accurate assessment of crop water status (DeJonge et al. 2015), which can be difficult to obtain. The most widely accepted method of determining an almond tree's water status is to measure its midday stem water potential (SWP) using a pressure chamber (Drechsler 2019). Obtaining SWP measurements in this way is tedious and time-consuming and thus is rarely used to assess crop water status in commercial settings. According to an Almond Board of California survey, less than 30% of almond growers use the pressure chamber to guide irrigation (Almond Board of California 2019).

✉ Isaya Kisekka
ikisekka@ucdavis.edu

1   Department of Biological and Agricultural Engineering, UC Davis, One Shields Ave, Davis, CA 95616, USA

2   Department of Land, Air, and Water Resources, UC Davis, One Shields Ave, Davis, CA 95616-8627, USA

3   Biological Systems Engineering, University of Wisconsin Madison, Madison, USA

The rationale for using SWP over leaf water potential by almond growers is that earlier provides a good indicator of the water status of the terminal stem and the canopy. SWP also tends to be less variable, therefore, almond growers prefer SWP over leaf water potential. More recent research also confirms the superiority of stem water potential over leaf water potential as a plant water status for guiding irrigation management decision making (e.g., Santisteban et al. 2019).

Numerous past studies have attempted to predict water status through other means, such as crop water stress indices or predictive models. Jackson et al. (1981) derived the widely known *crop water stress index* (CWSI), which is defined as the ratio of a crop's canopy temperature minus the air temperature, relative to the temperature differential of a well-watered plant and a non-transpiring plant (DeJonge et al. 2015). While crop water stress index has been demonstrated to correlate well with SWP, it still requires the calculation of a baseline stress value using a well-watered crop leaf (often impractical to obtain), and its correlation with SWP has been found to vary widely throughout a growing season (Möller et al. 2007). While the CWSI has its drawbacks, its consistent correlation with SWP indicates that its key parameters, namely air temperature, leaf/canopy temperature, relative humidity, and vapor pressure deficit, may be useful inputs in a predictive model of SWP (Blaya-Ros et al. 2020). Numerous other soil, plant, and atmospheric parameters show promising potential for the prediction of water stress. As reference evapotranspiration (ET_o) is calculated based on relative humidity and vapor pressure deficit, it follows that ET_o may be useful in a water status prediction model. Elevation data may help guide spatial predictions as it relates closely to the hydraulic gradient of an orchard, however, this may vary from site to site based on specific topographic features. Soil electrical conductivity (EC_a) is known to correlate with numerous soil properties, including water content, texture, and porosity (Hawkins 2017). As such, it too may guide spatial predictions of SWP by helping to capture the variation of soil properties across an orchard's area.

In addition to crop water stress indices, remotely sensed vegetation indices (VI's), derived from reflectance values of specific electromagnetic wavelengths, may also be used to estimate crop water status. Water is known to absorb radiant energy from many wavelengths throughout the electromagnetic spectrum, particularly throughout the mid-infrared region 6 (1300–2500 nm), with significant absorption bands centered on 1450, 1940, and 2500 nm wavelengths (Carter 1991). By extension, absorption of 400–2500 nm radiation by water causes the reflectance of a plant's leaves to decrease. When water is lost from leaves, there is an increase in intracellular air space, consequently increasing the intensity of reflections within the leaves. Based on this phenomenon, certain reflectance values may assist in the detection

of water stress alongside other relevant data. While the Normalized Difference Vegetation Index (NDVI) is certainly the most widely used VI, the Normalized Difference Red Edge Index (NDRE) has been demonstrated to correlate well with plant water content (Zhang 2019). NDRE is obtained using the following equation (Virnodkar et al., 2019):

$$NDRE = (R_{790} - R_{720})/(R_{790} + R_{720}) \tag{1}$$

where R_{800} is the reflectance of the 800 nm near-infrared spectral band, and R_{720} is the reflectance of the 720 nm red edge spectral band. The exact center wavelengths may vary slightly depending on the specific equipment used.

In recent years, machine learning (ML) algorithms have emerged as a novel tool for modeling a wide variety of agricultural phenomena, including plant water status. Compared with more traditional modeling techniques, ML models can prove advantageous in capturing complex nonlinear relationships between variables, as is often seen when working with soil, plants, and atmospheric data. Two common ML approaches are Artificial Neural Networks (ANN) and Random Forest (RF). ANNs are comprised of input, hidden, and output layers, each of which contains a number of neurons representing predictive variables. They are particularly useful for modeling of highly non-linear phenomenon (Virdnokar et al. 2020). A random forest model is made up of many decision "trees" which may be used for regression and classification. The defining attributes of this model type are the number of trees "n" and the number of variables used to make decisions at each tree, or "mtry". Past studies have determined n = 500 as the optimal number of trees based on resulting error stabilization (Belgiu et al. 2016). "mtry" is typically set to the square root of the number of predictive variables (p), or p/3 if more precise integer value is desired (Gislason et al. 2006). RF models are advantageous as they are highly resistant to overfitting compared with other model types (Virdnokar et al. 2020).

ML models have already been applied in many case studies for water stress prediction. In a 2018 study, an ANN model utilizing thermal indices along with canopy and air temperature data predicted SWP in grapevines with relatively good accuracy [$R^2 = 0.61$] (Gutierrez et al. 2018). It was found that whether baseline temperatures for the indices were included as inputs or not, the model predicted SWP equally well. Several other studies have utilized remotely sensed multispectral imagery in ANN models to predict SWP with promising results, with coefficients of correlation ranging from 0.58 to 0.87 (Poblete et al. 2017; Romero et al. 2018). In another case, the predictive performance of an optimum regression equation was compared with that of an ANN model, utilizing various soil and environmental parameters (Marti et al. 2013). The ANN model and regression equation resulted in coefficients of correlation between

observed and predicted values of 0.93 and 0.85, respectively, indicating the superiority of the ANN framework for water stress prediction. Applications of random forest for crop water status are more limited however Yang et al. (2021) used a random forest model to predict canopy temperature of Chinese Brassica based on various meteorological data ($R^2 = 0.77-0.9$). Predicted temperatures were then used for the calculation of CWSI. Currently, applications of machine learning for water status prediction in almonds are relatively limited. In a 2018 study, an ANN model was used to predict water status in almonds with moderate success ($R^2 = 0.78$). In this case, the model was trained to predict $T_{\text{diff dry}}$, an alternative plant water status indicator, and utilized SWP as a predictive variable (Meyers et al. 2019).

Previous work has demonstrated that machine learning may be used effectively to predict water status in plants, based on soil and environmental variables, as well as information obtained through remote sensing. This study aimed to investigate whether machine learning-based models could be trained to accurately predict SWP in almonds, utilizing a combination of soil, environmental, and UAV-based variables. Also, we investigated the ability of ANN and RF models to map SWP spatially across an entire orchard.

Materials and methods

Study site and experimental layout

The site used in this study was a 1.6 hectare orchard located at Nickels Soil Lab near Arbuckle, CA. Data from the California Irrigation Management System (CIMIS) database shows that the average air temperature and monthly precipitation from January 2021 through December 2022 was 17°C and 19.6 mm/month, respectively (CIMIS 2024). The plot contained a total of 15 rows of 50 trees, with 5 rows each of nonpareil, butte, and Aldrich almond varieties. Throughout the two seasons during which the study took place, each row received a standard irrigation treatment determined by the grower. Each week from bloom until harvest, approximately 70 mm of irrigation was applied over two 18-h periods, one Sunday and the other Wednesday. For each date, irrigation began at 3 pm and lasted until 9 am the next day. The irrigation regime aimed to replace water losses from averaged daily evapotranspiration, adjusted by a crop coefficient (ET_c). The system itself consisted of double drip lines spanning the length of each row, with emitters spaced approximately 6 feet apart.

For data collection, twenty focus trees were chosen throughout the orchard to serve as sampling locations, with four locations occurring in each nonpareil row (Fig. 1). At each location, a 1.5 m aluminum access tube was installed near the base of the tree for collecting soil moisture
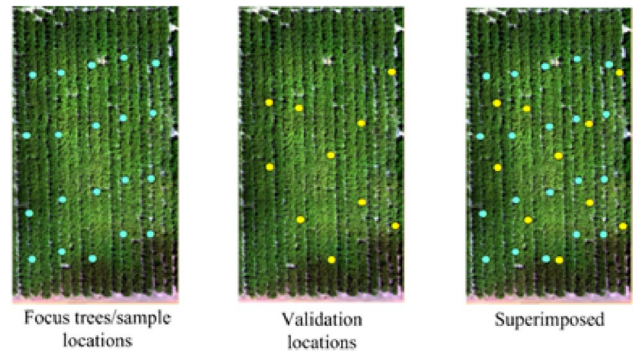

Fig. 1 Locations of focus (cyan dots) and validation (yellow dots) trees for stem water potential measurements at the Nickels Soil Lab near Arbuckle, California

measurements using neutron attenuation with a neutron probe (model: Instrotek 503 Hydroprobe, San Francisco, CA). An additional ten locations were designated for validation of SWP. Each validation measurement location occurred in a row of butte variety, while all main sampling locations occurred in nonpareil rows. Almond variety was not considered as a factor in observed SWP data, as a model that can assess water stress across an orchard's entirety regardless of tree variety may have more robust applications. Furthermore, a 2019 study conducted in the same plot found no increase in marketable kernel yield as a result of variety specific implementation of regulated deficit irrigation (Dreschler and Kisekka 2022). Thus, differences in almond variety were not considered to significantly affect observed SWP values. Data collection was carried out once every one to two weeks, typically between the hours of 11 am—2:30 pm, in order to allow for the majority of measurements to be taken as close to solar noon as possible. The days of the week on which data were collected varied over time in an effort to capture a variety of water stress conditions (some before and some after irrigation events).

Midday stem water potential and soil moisture measurements

On each measurement date, soil moisture was measured at each focus tree using a neutron probe. Five measurements were collected at each sampling location from 0.3 to 1.5 m, in increments of 0.3 m. On each soil moisture measuring day, a "standard count" was performed on the neutron probe to calibrate readings for current environmental conditions.

At each sample location, midday stem water potential was measured using a pressure chamber filled with nitrogen gas (PMS Instrument Company Model 615, Albany, OR). To obtain midday stem water potential, a leaf was first covered for a minimum of fifteen minutes using a mylar foil bag (Fulton et al. 2001; Fulton 2018), which removes the

effect of solar radiation on the stem water potential reading. Per established "best practices" for SWP measurement in California, only well-shaded leaves occurring in the lower canopy were selected for measurement (Fulton 2018). All stem water potential measurements were collected between the hours of 11 am and 2:30 pm, in an effort to center the time range of sample collection around solar noon. The leaf was then cut from the tree and inserted into the pressure chamber with the end of the severed stem protruding from the end of a small gasket. Then pressure was slowly added to the chamber using the nitrogen gas until water began to bubble out of the end of the stem. The pressure at which this occurred was considered the tree's stem water potential (Dreschler et al. 2019). For each sample location, one leaf was measured from the focus tree, and one from the next tree immediately to the south. These two measurements were averaged and considered as the representative stem water potential for the location. Occasionally nitrogen gas would run low, in which case only one leaf would be measured at each location. While greater repetition of measurements would have been ideal, we chose to focus on capturing spatial variability over repetition. It has also been established that lower shaded leaves equilibrate more readily with the tree's water conducting system, requiring less repetition for accurate measurement (Fulton 2018). For eight of the 2022 collection dates, ten extra measurements were obtained from each validation location for the purpose of spatial validation of SWP maps.

Environmental and soil data

Reference evapotranspiration (ET_o) data for each date was obtained from California's CIMIS weather network to help improve the scalability of the research. The data consisted of daily ET_o (mm/day) and was derived from the CIMIS station near Williams, CA, located approximately 16 km north of the study area. Daily ET_o values were considered representative of the entire orchard. ET_o values were not multiplied by a crop coefficient, as variations in base ET_o values were considered sufficient for guiding model predictions. Additionally, maps of apparent soil electrical conductivity (EC_a) at depths of 0.75 and 1.5 m for the entire orchard were collected using an EM-38 MK11 (Geonics Limited, Ontario Canada).

Remote sensing data collection and analysis

Multispectral imagery of the tree canopies was collected on each measurement date using a DJI Matrice 100 quadcopter drone (Shenzhen DJI Sciences and Technologies Ltd., Shenzhen, Guangdong0 equipped with a Micasense RedEdge multispectral camera (AgEagle Aeriel Systems, Wichita, KS) [Fig. 2]. The six spectral bands measured were centered on the 475, 560, 668, 717, and 842 nm wavelengths. Flights were conducted as close to solar noon as logistics allowed, between the hours of 11:30 am and 2:30 pm local time. Flight altitude was set at 80 m, utilizing an 85% front and side overlap for the images obtained. Directly before and after each flight, multispectral images of a reflectance calibration panel were captured in order to account for current light conditions.

During the 2021 season, multispectral images were collected at too sparse a scale, so creation of a complete orthomosaic of canopy reflectance values for each date was not possible. Instead, corresponding images for each sample location were selected based on coordinates embedded in their metadata, and canopy reflectance values were extracted using R Studio. Reflectance values were calibrated using a reflectance panel to account for specific ambient light conditions of each date. For all 2022 collection dates, multispectral data was processed using Agisoft Metashape. By stitching together many overlapping images, an

Fig. 2 Left to Right: DJI Matrice 100, Micasense Reflectance Calibration Panel, and Micasense RedEdge Multispectral Camera

orthomosaic covering the entire area of the study site was produced. Using raster transformation functions within the software, each pixel was set to display its resultant NDRE value, based on the pixel's reflectance values for the red edge and near-infrared spectral bands. Raster files displaying NDRE were then exported to ArcMap software, where they were all clipped to uniform boundaries and resampled to a pixel size of 0.2 m to decrease model processing time. Using GPS coordinates of each sample location, NDRE values were extracted from each raster.

Statistical metrics utilized

Coefficient of correlation (R^2), Root Mean Squared Error (RMSE), Normalized Root Mean Squared Error (nRMSE), and Index of Agreement (IOA) were used to evaluate model performance in this study. Normalization of RMSE can be performed in several ways; in this case, it was calculated by dividing RMSE by the range of predicted values from the corresponding dataset (Otto 2019). This is helpful when comparing datasets with significantly different spreads of values, such as the comparison of raw and baseline adjusted SWP predictions. The index of Agreement represents the ratio of the mean squared error and potential error and is effective at detecting differences in observed and predicted variances and means (Krause 2005).

Artificial neutral network (ANN) modeling

Following data collection in 2021, MATLAB's Deep Learning Toolbox was used to set up and train artificial neural networks. In this case, baseline corrections were applied to all stem water potential values, with the intention of better accounting for environmental conditions at the time of measurement. SWP baseline corrections are based on relative humidity and air temperature and can be obtained from tables available online (Fulton 2019). The baseline value is subtracted from the "raw" stem water potential value, giving baseline-adjusted stem water potential. Initially, eight neural network models were trained with different combinations of input parameters and compared based on their performance. The best model utilized soil moisture at 0.3 to 1.5 m depths, ET_o, and NDRE as predictive variables. Basic sensitivity analysis was performed by retraining the model numerous times, with a different predictive variable excluded each time. The most important variable was determined to be soil moisture, followed by NDRE, and ET_o. Sensitivity analysis of individual soil moisture depths was not performed at this stage.

Following the collection and processing of all data for the 2022 growing season, four more neural network models were trained. In order to simplify future model applications, soil moisture data for 1 m through 1.5 m depth was omitted.

The first new ANN to be trained utilized soil moisture from 0.3 to 0.6 m, daily ET_o, and NDRE as inputs. The second utilized the same predictors but was trained to predict raw SWP values rather than baseline-adjusted SWP. The third included elevation data as an additional predictor, and the fourth utilized soil moisture at 0.3, 0.6, and 1.5 m depths, ET_o, and NDRE. This was performed to mirror the exact parameters used to train random forest models and determine if this parameter combination also produced better results in an ANN model. The elevation data was obtained in 2019 for a previous study at the same site (Dreschler and Kisekka 2022), and was considered representative of current conditions. The third model produced less accurate results than the first two based on resultant R^2 and RMSE values. Thus, the first and second were both selected for the generation of SWP maps in order to gain further insight into the effect of predicting baseline-adjusted versus non-baseline SWP values.

To generate maps of SWP using the selected ANN models, raster layers were first created for each input parameter using ArcMap. Maps were only generated for dates on which extra validation SWP measurements had been collected, which were May 17, May 27, June 1, June 14, July 5, July 29, August 3, and August 18 (2022). To do so, kriging was performed using the coordinates and matching data of each sample location, using ArcGIS 10's default measurement variation of 100%. The kriging was fit to the boundaries of the previously generated NDRE layers and sampled to the same pixel size (Fig. 3). Each matching set of input layers was then exported to R Studio, where their pixel values were extracted and transformed into a singular array in which each column represented a single input parameter. The array was then exported to MATLAB, where it was fed into the ANN model, resulting in an output array of predicted SWP values. The output array was then exported to R Studio, where it was then transformed back into a raster layer with the same boundaries and resolution as before. Each map was then displayed in ArcMap (Fig. 4), where predicted SWP values were extracted from all sample and validation locations for further analysis.

Random forest (RF) modeling

To train random forest models for SWP prediction, an R package called CAST was used which allows for spatial–temporal modeling using machine learning (Meyers 2018). Initially, a "forward feature selection" (FFS) was performed on the training dataset with all predictive variables included (Fig. 5). The analysis showed that by starting with two predictive variables and iteratively increasing the number of predictors, the R^2 value of the resulting model's predictions stopped improving after five predictors had been added. The FFS reported soil moisture at

Fig. 3 Examples of kriging based spatial input layers used for stem water potential mapping in an almond orchard at Nickels Soil Lab near Arbuckle California. Left: Soil Moisture (0.3 m), Center: Soil Moisture (0.6 m), Right: Remotely sensed Normalized Difference Red Edge Index
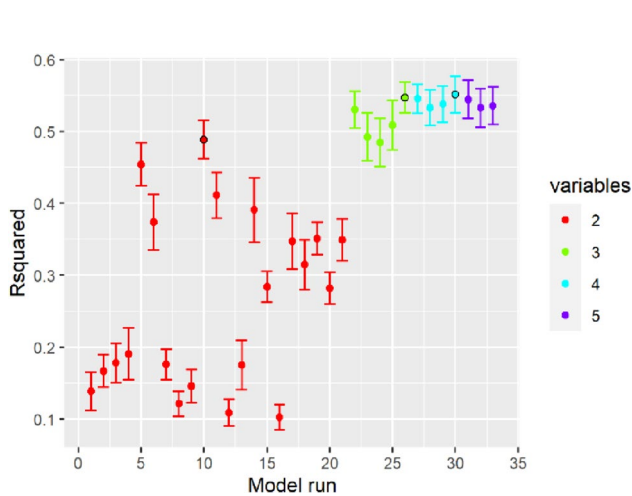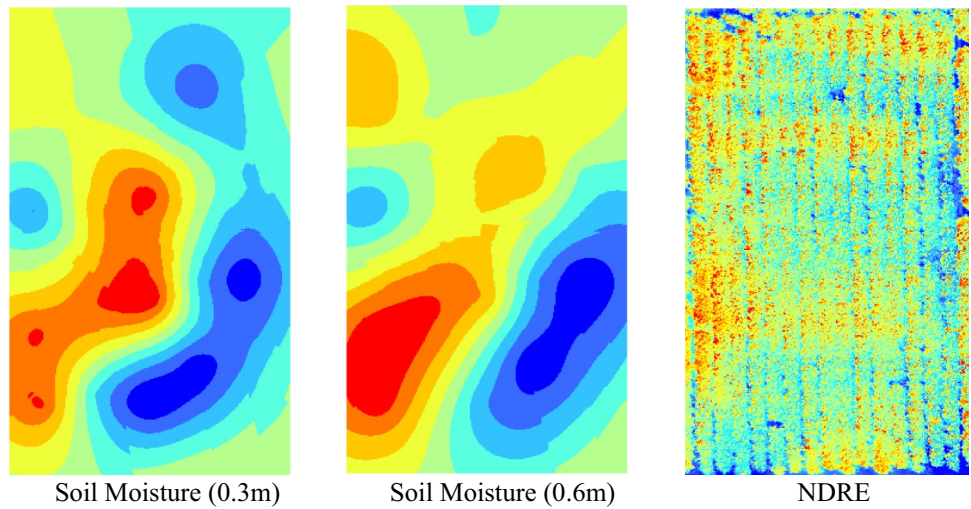

Soil Moisture (0.3m)   Soil Moisture (0.6m)   NDRE

Fig. 4 ANN based maps displaying predicted raw stem water potential values for an almond orchard at Nickels Soil Lab near Arbuckle California
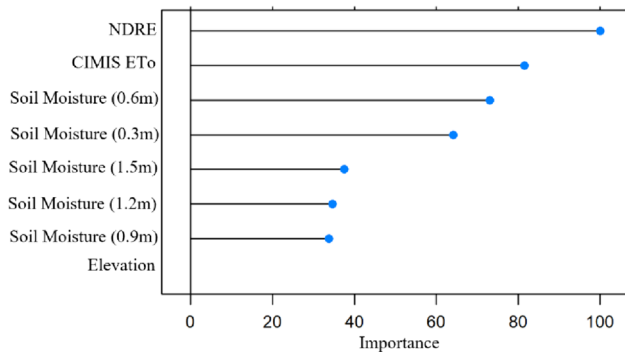
Fig. 5 Results of random forest forward feature election analysis. As seen in the chart legend, varying colors represent different numbers of predictive variable
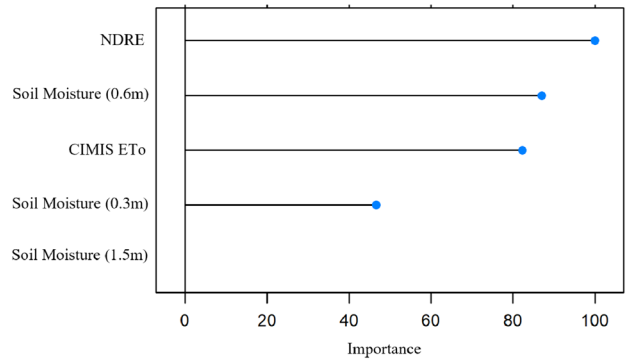

Fig. 6 Results of variable importance analysis with all potential predictors

0.3 and 0.6 m, NDRE, and ET_o to be the strongest predictive variables. Additionally, an initial random forest model was trained with all input variables and a "variable importance" function was applied, which consistently ranked 1.5 m soil moisture as the fifth most important variable (Fig. 6). It is not completely known why 1.5 m moisture had a significant impact on the model, but it is suspected that a layer of hardpan may be located at this depth in the soil, which when blocking percolation of moisture could indirectly impact the trees' SWP. Thus, the final combination of predictive variables for random forest modeling was selected as 0.3, 0.2, and 1.5 m soil moisture, ET_o, and NDRE.

Model validation

Using the designated input parameters, two random forest models were trained with 500 "trees" and an "mtry" value of two, based on the recommendation of using a value of p/3 in which p represents the number of predictive variables. A K-fold function with ten repetitions was used for model

validation. One model was trained to predict baseline adjusted SWP while the other predicted non-baseline values. In addition to regular model training and validation, a target-oriented validation was also performed. Rather than random folds, this method utilized specified folds in which the entire time series data of a different sampling location was omitted, using the CAST package's "CreateSpaceTimeFolds" function. This type of validation provided insight into the model's ability to predict stem water potential spatially. The same variable importance analysis function was also applied to the results of the target-oriented validation in order to understand the importance of each variable as it pertains to spatial prediction specifically (Fig. 7).

Once sufficient random forest models had been trained and validated for both baseline adjusted and non-baseline SWP, maps were generated with the same input raster layers used for ANN map generation. For each date, the corresponding raster layers were imported to R Studio, where they were input to the model, which then returned a map of the predicted SWP.

Following SWP mapping using the random forest models described above, one more model was trained with the additional predictors of apparent soil electrical conductivity (EC_a) at 0.75 and 1.5 m depths. The purpose of this analysis was to observe whether additional soil variables may help improve the model's predictions, either spatially or simply with point data. In this case, the model was trained to predict raw SWP values only. A variable importance analysis was conducted in R to determine the importance of 0.75 and 1.5 m EC_a relative to the other selected variables. Results were identical to the analysis performed for the previous RF model (see Fig. 6), with the addition of 1.5 m EC_a and 0.75 m EC_a being ranked second to last and last, respectively. Despite both EC_a
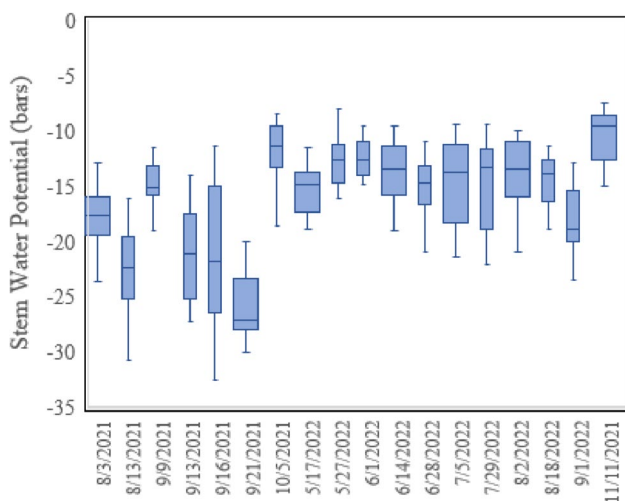
variables being ranked as least important, they were still included in the model along with the other five variables in order to observe whether their presence could improve the spatial performance of SWP maps. For this model, "Mtry" was set to a value of three, due to there being seven predictive variables. Once the model had been trained, maps were once again generated in R studio using the same input raster layers as the previous models.

Results and discussion

Measured stem water potential

Figure 8 illustrates the spread of measured SWP values used as model target values. Over the entire study period, observed water stress levels ranged from "minimal" to "severe". Categorization provided by the University of California Agricultural Extension defines minimal stress as −6 to −10 bars, mild as −10 to −14 bars, moderate as −14 to −18 bars, high as −18 to −22 bars, very high as −22 to −30 bars, and below −30 bars as severe (Fulton 2019). Conditions ranging from mild to high or very high were observed on five out of seven measurement dates in 2021, and on eight out of ten measurement dates in 2022. At least moderate water stress was observed on every measurement date throughout the entire study period. The wide variations in water stress observed at the study site, often within the same several hour period, reinforces the applicability of SWP as an irrigation scheduling guide.

ANN based stem water potential modeling

Following the 2021 growing season, nine preliminary ANN models were trained with varying predictive variables. For all models, the coefficient of correlation between observed and predicted SWP ranged from 0.65 to 0.92. The best performing model utilized soil moisture, ET_o, and NDRE as predictive variables. Following collection of all 2022 data, four additional ANN models were trained with two full seasons of data (Table 1). For these models, coefficient of correlation between observed and predicted values ranged from 0.74 to 0.8 for training data, and from 0.75 to 0.88 for testing data (Table 1). Compared with ANN 1, the normalized RMSE of ANN 2 increased by about 13 percent for both the training and testing datasets. Model 3 included the same data as Model 2, with the addition of elevation as a predictor. No significant increase in model performance was observed, indicating that the addition of elevation data did not increase predictive power. Model 4 was trained to reflect the performance of ANN modeling with the best combination of parameters from random forest modeling. All goodness-of-fit statistics showed a decrease

Fig. 7 Results of variable importance analysis on target-oriented validation

Fig. 8 Range of measured stem water potential values observed on each data collection date an almond orchard at Nickels Soil Lab near Arbuckle California. Values shown in this figure do not include baseline adjustments
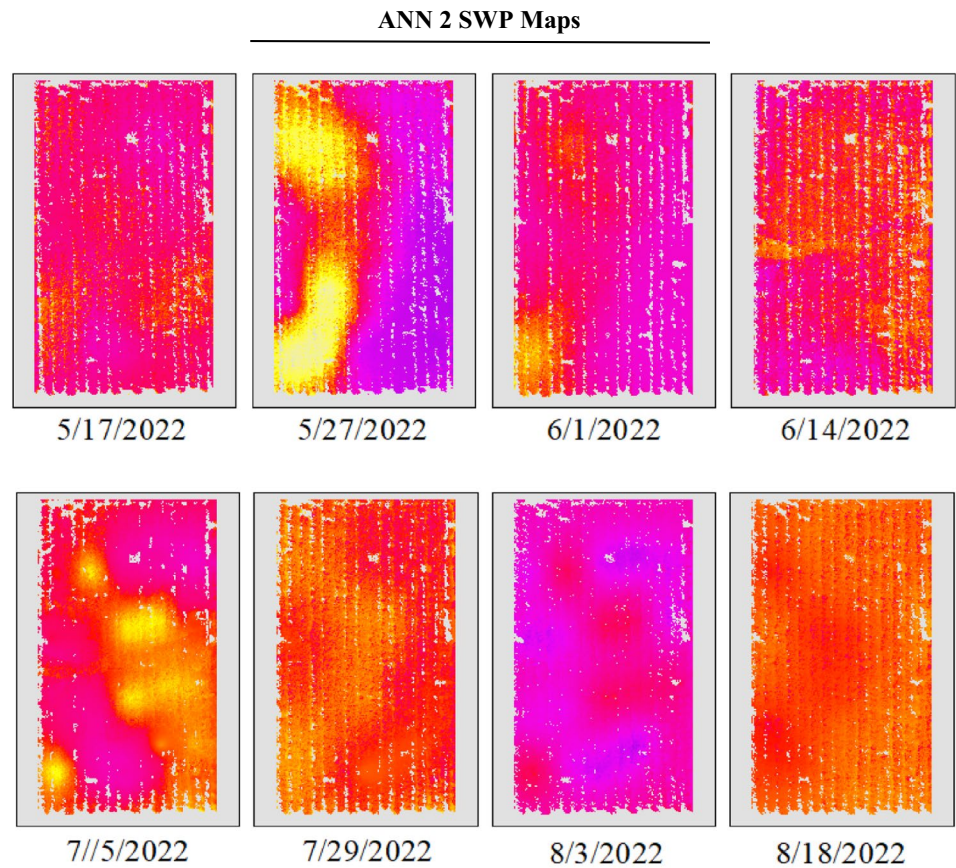
ANN 2 SWP Maps

5/17/2022   5/27/2022   6/1/2022   6/14/2022

7//5/2022   7/29/2022   8/3/2022   8/18/2022

Stem Water Potential (bars)

-10

-20

Table 1 Goodness-of-fit statistics for different machine learning models used to predict stem water potential in an almond orchard at Nickels Soil Lab near Arbuckle California

Model	Training data performance				Testing data performance			
	R^2	RMSE	nRMSE	IOA	R^2	RMSE	nRMSE	IOA
ANN 1	0.8	3.08	0.12	0.7	0.88	2.96	0.12	0.74
ANN 2	0.74	3.39	0.14	0.66	0.81	3.24	0.14	0.72
ANN 3	0.75	3.26	.14	0.67	0.79	3.84	0.16	0.71
ANN 4	0.75	3.34	0.13	0.67	0.75	3.56	0.16	0.67
RF 1	na	na	na	na	0.6	3.32	0.13	na
RF 2	na	na	na	na	0.57	3.36	0.14	na
RF 3	na	na	na	na	0.57	3.39	0.14	na

IOA Index of Agreement

in model performance from model 2, indicating that the best parameter combination for random forest SWP modeling is not necessarily the best for ANN SWP modeling.

Following generation of SWP maps for designated dates using the selected ANN models, predicted SWP values were extracted from all sampling locations as well as extra

validation locations. These values were plotted against corresponding observed values, allowing for insight into the models' abilities to predict SWP spatially. For spatially predicted values at sampling locations, the coefficient of correlation with observed values was much lower than for non-spatial point data, ranging from 0.2 to 0.3. RMSE and normalized RMSE both increased significantly as well. For extra validation locations, coefficient of correlation decreased slightly more, and RMSE and normalized RMSE both increased slightly. This makes sense, as model uncertainty would increase further the more you move away from data collection locations.

Random forest based stem water potential modeling

The variable importance analysis performed in R Studio on the random forest training dataset (Fig. 7) revealed NDRE, CIMIS ET_o, 0.6 m soil moisture, 0.3 m soil moisture, and 1.5 m soil moisture to be the most important predictors, in that order. It is worth noting that the 0.9 and 1.2 m soil moisture data were reported to have little to no significance in regard to the model's prediction this is probably due to a clay restricting layer underlying the alluvial Arbuckle soil series as reported in previous studies (Andreu et al. 1997; Drechsler and Kisekka 2022; Vanella et al. 2022). Elevation was also had no significance whatsoever. While 1.5 m moisture was reported to have only slightly more significance than 0.9 and 1.2 m, it was included as a predictor due to the recommendation of five variables from the feed forward selection function. The same analysis applied to the target-oriented validation produced nearly the same result, with the exception of 0.6 m soil moisture being listed as more important than CIMIS ET_o. This was interpreted as an indication that while 0.6 m moisture may not be as significant as ET_o in prediction based solely on point data, it becomes more useful than ET_o when the model is extrapolated spatially between point data locations. The results of the forward feature selection function applied to all predictive data are shown in Fig. 5. The analysis showed that as models were successively trained with more predictive variables, the coefficient of correlation for the resulting model stopped improving after a fifth variable was added. This finding agrees with the variable importance analysis, which indicated five parameters that were significant to model prediction.

Random forest stem water potential mapping

For the baseline-adjusted and non-baseline random forest models, coefficients of correlation were 0.57 and 0.6, respectively, and nRMSE values were 0.14 and 0.13, respectively. These results indicate an increase in predictive performance compared with ANN models. Due to the nature of the R Studio package used, it was not possible to estimate the index of agreement for random forest models. A slight increase in performance was observed when comparing baseline-adjusted with non-baseline model results. A roughly 5 percent increase in the coefficient of correlation was observed, while nRMSE decreased by approximately 0.7 percent.

The performance of SWP maps generated by random forest models generally increased compared with that of ANN. For non-baseline SWP maps, the coefficient of correlation at validation locations was 0.4, while nRMSE and index of agreement were 0.19 and 0.56, respectively. Contrary to model training results, the performance of baseline adjusted SWP maps was lower by comparison ($R^2 = 0.17$, nRMSE = 0.24, IOA = 0.52). The best spatial mapping results were produced by non-baseline adjusted random forest maps, while the best model training results occurred with baseline-adjusted random forest models. The reason behind this contradiction is not entirely known. Given that baseline SWP values were calculated from measurements collected 16 km from the site, it is possible that baseline values helped to guide point-model predictions, but fell short in accounting for true variation throughout the orchard, leading to a slight increase in model uncertainty.

Comparison of RF and ANN models

No significant difference was observed in model performance for ANN models versus RF. RF based maps produced noticeably more accurate predictions at data sampling locations than ANN based maps, but prediction accuracy at validation measurement locations was essentially the same for both map types (Table 2). For all models, a significant decrease in performance was observed when comparing respective map performance to model performance. Generally, coefficients of correlation decreased by roughly fifty percent, and nRMSE values roughly doubled. In models for which index of agreement could be calculated, an approximate twenty five percent decrease in value was observed. It is also worth noting that for all models at all sampling locations, the x–y distribution of observed and predicted values had a slope must steeper than the line y = x. This means that for potentially large observed values, models were often underestimating the magnitude of predicted values.

Figure 9 illustrates the ability of each model type to predict SWP at sampling (focus tree) locations, at which corresponding soil moisture values were also collected, while Fig. 10 illustrates the models' abilities to predict SWP at spatially randomized validation locations (see Fig. 1). At sampling locations, models generally tended

Table 2 Summary of goodness-of-fit for stem water potential mapping using different machine learning models in an almond orchard at Nickels Soil Lab near Arbuckle California

Model	Sampling locations					Validation locations				
	R^2	RMSE	nRMSE	IOA	Bias (MSD)	R^2	RMSE	nRMSE	IOA	Bias (MSD)
ANN 1	0.3	3.06	0.21	0.52	0.25	0.29	3.64	0.23	0.54	−0.42
ANN 2	0.2	3.11	0.2	0.46	0.73	0.18	3.81	0.2	0.5	0.18
RF 1	0.6	2.74	0.19	0.59	−0.32	0.17	3.84	0.24	0.52	−0.61
RF 2	0.47	2.71	0.18	0.56	0.04	0.4	3.6	0.19	0.56	−1.05
RF 3	0.57	2.61	0.17	0.57	0.07	0.28	3.8	0.19	0.54	−0.66

*IOA Index of Agreement; *MSD* Mean Signed Difference

to underestimate the magnitude of SWP values, while at validation locations they tended to overestimate (since SWP values are negative, negative bias indicates an overestimation of SWP magnitude). At validation locations, the methods of interpolation used to create the "input" layers shown in Fig. 3 increase uncertainty and variation across all predictors, which in turn appears to cause overestimation in the models' predictions. For RF models, map performance at sampling locations remained mostly the same as model performance, while performance at validation locations was noticeably worse (Table 2, Figs. 9 and 10). For ANN models, map performance at sampling locations was significantly worse than model performance but remained mostly constant when considering performance at validation points. Thus, no significant difference in map performance at validation locations was observed between ANN and RF-based maps, but RF-based maps performed significantly better than ANN maps at sampling locations. The addition of EC_a data to the RF model produced results largely similar to those of the previous RF models (Table 2, Fig. 11). Model training with point data resulted in a coefficient of correlation of $R^2 = 0.57$ and an nRMSE of 0.14. Resultant SWP maps predicted SWP at sampling locations with $R^2 = 0.57$ and nRMSE of 0.17, and at validation locations with $R^2 = 0.28$ and nRMSE $= 0.19$.

Discussion

Results from this study have demonstrated the potential for machine learning models to predict spatial SWP in almonds, although methods will need to be greatly refined in order to obtain any sort of commercial viability. The plausible indirect link between level-level vegetative indices (e.g., NDRE) estimated from remote sensing data and stem water potential could be that leaf-level indices are related to leaf water potential, and leaf water potential is related to stem water potential. In the same species leaf water potential and stem water potential are normally directly correlated.

Both ANN and RF models failed to accurately characterize spatial SWP variability throughout the orchard.

Across all models, an average coefficient of correlation of $R^2 = 0.73$ and average normalized RMSE of 0.14 were observed between measured and predicted values, which is largely in agreement with the findings of recent similar studies. Poblete et al. (2017) and Pocas et al. (2017) used ML frameworks to predict grapevine SWP with $R^2 = 0.58 - 0.87$ and $R^2 = 0.77$, respectively. Meyers et al. (2019) predicted $T_{diff\ dry}$, (novel plant water status indicator) with $R^2 = 0.78$ (average). Based on sensitivity analysis from both RF and ANN models, it seems that spectral reflectance data was the strongest predictor of water stress, followed closely by root zone soil moisture. This confirms the findings of Zhang et al. (2019) that Normalized Difference Red Edge Index can vary significantly in response to changes in canopy water content. The same study also found the red edge spectral band (integral to NDRE calculation) to correlate well with canopy water content ($R^2 = 0.78$). Romero et al. (2018) used an ANN model to predict vineyard SWP using only remotely sensed vegetation indices ($R^2 = 0.72$). Both the literature and this study's results indicate the potential for development of future almond SWP models based entirely on remotely sensed multispectral data, complementing in-situ sensors and minimizing labor requirements. Going forward, this could help to increase the commercial viability of similar workflows for the assessment of water stress. Another benefit is the framework proposed in this paper does not require thermal data, which may lower cost and allow greater flexibility in the development of future workflows.

In both RF and ANN models, map prediction accuracy was far less accurate than that of the point models themselves. It seems that a higher spatial sampling density for input variables is necessary in order to account for the true spatial variation encountered in an orchard. Additionally, this performance gap might be improved upon by the inclusion of more remotely sensed parameters, such as leaf canopy temperature, which can be precisely measured on a pixel-by-pixel scale.

When comparing baseline-adjusted SWP predictions to raw SWP predictions, little difference was observed in model performance with point data. This agrees with results published by Gutierrez et al. (2018), who used thermal indices as predictors in ANN models to predict
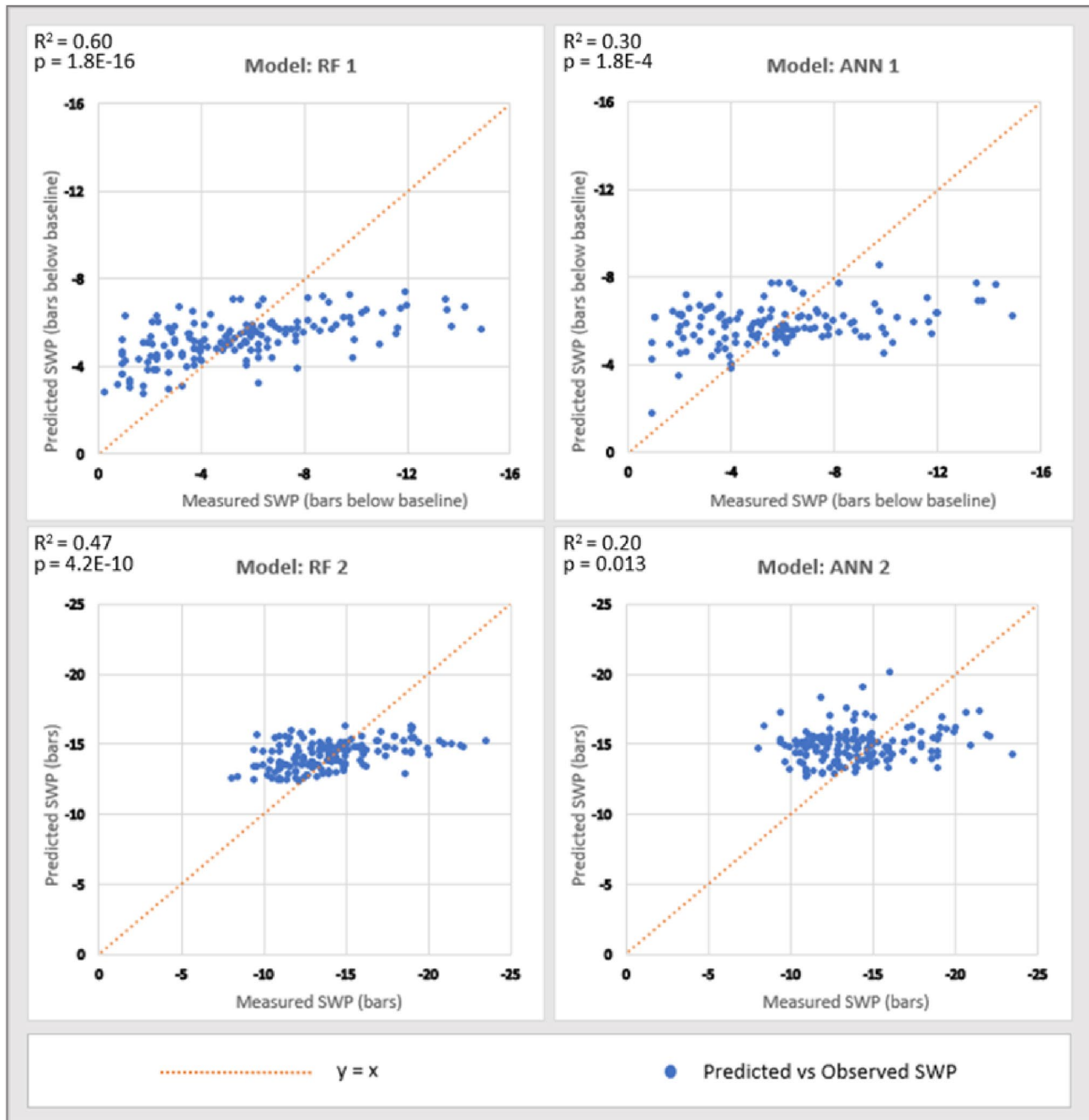
Fig. 9 Model performance (R^2 and p-value) in predicting stem water potential at focus tree locations in an almond orchard at Nickels Soil Lab near Arbuckle California

grapevine SWP. It was found that whether or not baseline temperatures were included in thermal indices, the resulting models performed equally well. However, raw SWP showed a slight increase in accuracy for spatial map predictions. As previously mentioned, the single hourly value parameters used in baseline calculation likely do not account for variations occurring within the orchard. Furthermore, baseline calculations are by nature an approximation,

which introduces further uncertainty to SWP models. It is possible that training models with larger datasets could reveal these uncertainties to a greater degree. It is also worth noting that baseline calculations may still play a role in making irrigation decisions based on predicted SWP values. Baselines reflecting daily conditions may influence the exact "threshold" of stress at which growers whether to apply irrigation.
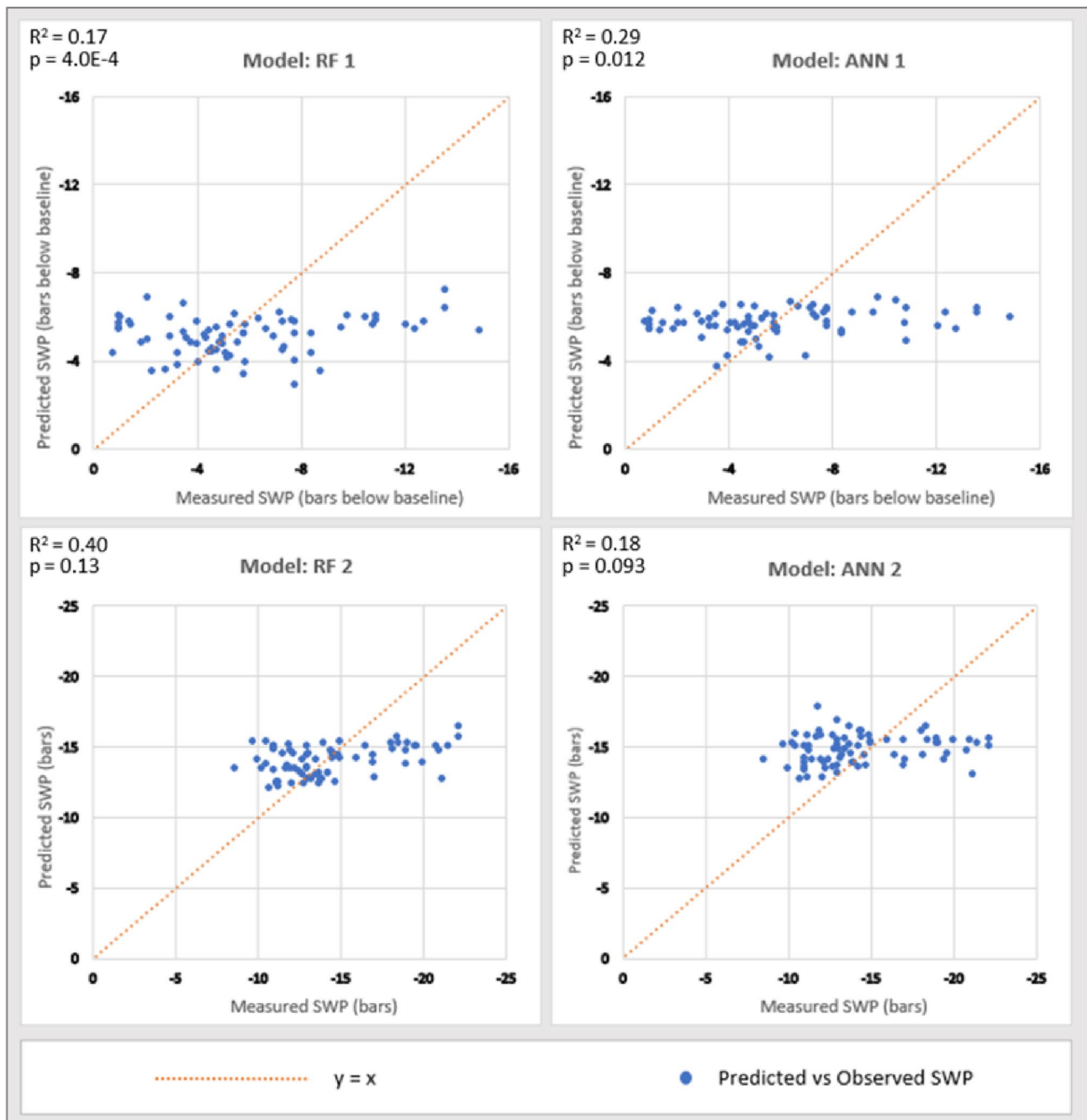
Fig. 10 Model performance (R^2 and p-value) at validation locations in an almond orchard at Nickels Soil Lab near Arbuckle California

Both RF and ANN based models performed similarly, however, RF based maps showed more accurate predictions at data sampling locations. In ANN based maps, prediction accuracy was roughly the same at both data sampling and validation locations. Both model types appeared to frequently underestimate the magnitude of predicted SWP values, when the target value would in fact be quite large. At sampling locations, the RF framework seemed to maintain a degree of consistency between model and map predictions

that the ANN framework did not. Nonetheless, based on this study's results it does not seem reasonable to suggest that either model type is better suited than the other for water stress prediction in almonds. Before that can be conclusively determined, spatial model performance must be drastically improved through other means, such as more robust training datasets. In all likelihood, the type of machine learning model is far less important than variable choice, sampling frequency, and spatial sampling distribution.
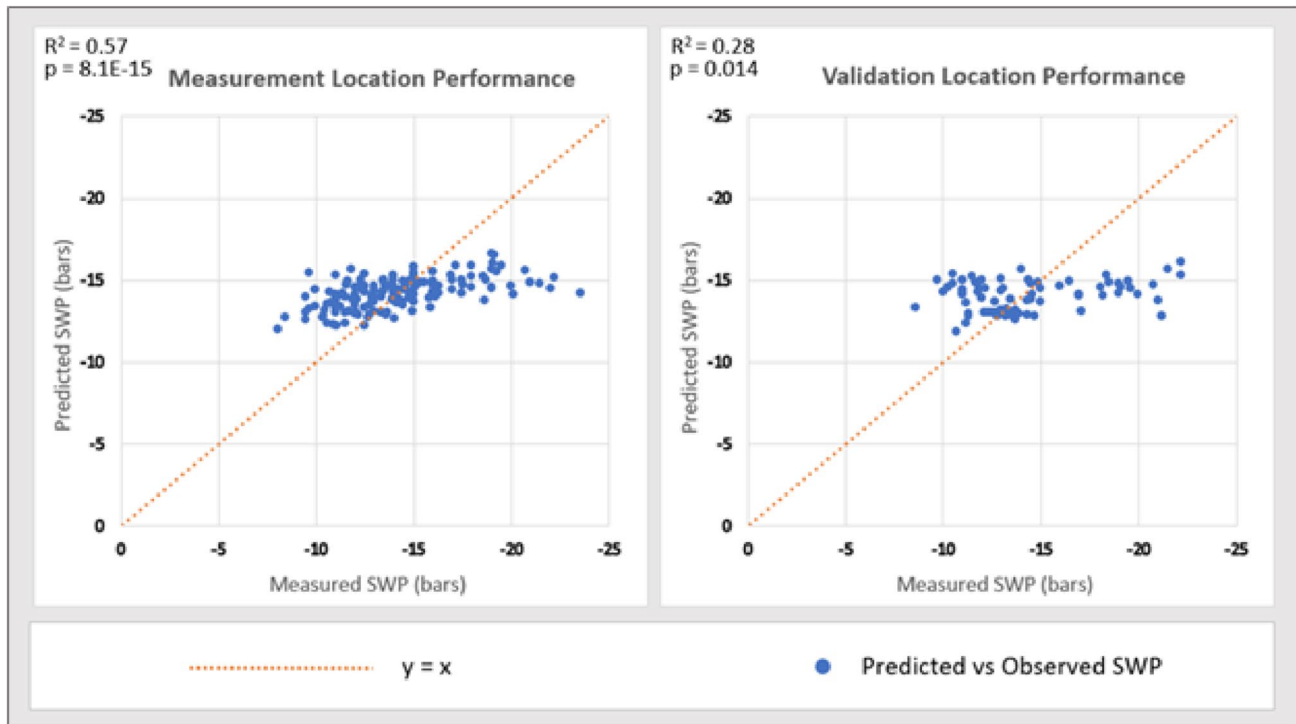
Fig. 11 Model performance (R^2 and p-value) at validation and focus tree locations for Random Forest Model 3 (random forest model with electrical conductivity data added) in an almond orchard at Nickels Soil Lab near Arbuckle California

Table 3 Descriptions of predictive variables names used for stem water prediction for all models in an almond orchard at Nickels Soil Lab near Arbuckle California

Model name	Predictor variable	SWP prediction
ANN 1	SM (0.3–0.6 m), ET, NDRE	Baseline adjusted
ANN 2	SM (0.3–0.6 m), ET, NDRE	Raw
ANN 3	SM (0.3–0.6 m), ET, NDRE, Elevation	Raw
ANN 4	SM (0.3, 0.6, 1.5 m), ET, NDRE	Raw
RF 1	SM (0.3, 0.6, 1.5 m), ET, NDRE	Baseline adjusted
RF 2	SM (0.3, 0.6, 1.5 m), ET, NDRE	Raw
RF 3	Same as RF 2, plus soil EC (0.5 and 1 m)	Raw

The addition of soil EC_a data to the RF model did not appear to have a significant effect on model or map prediction accuracy. As seen in Tables 2 and 3, RF3 (the model with EC_a data) performed nearly identically to RF2 (the RF model predicting raw SWP values). Maps produced by the RF3 were found to have nearly the same accuracy of prediction as other RF maps, at both sampling and validation locations. One may note that EC_a data appears to have had some effect on spatial prediction, as the SWP maps in Fig. 16 have a slightly blocky appearance, mirroring the shape of the EC_a raster layers used as inputs (see Fig. 12 in supplemental

materials section for EC_a raster layer examples). It is possible that a more spatially dense collection of EC_a data could have a greater impact on model results, but based on the results of RF3 EC_a does not appear to be a strong indicator of water stress in almonds (See Supplemental Figs. 13, 14 and 15).

The study's methodology contains numerous limitations which may have hindered the accuracy and robustness of model predictions. Ideally, more than two repetitions of SWP measurement would have been performed at each focus tree, however, measurement quantity was often limited by the volume of nitrogen gas that was able to be transported to the field in one visit. The intense but varying heat of summer days at the site also made it difficult to predict the correct amount of gas to bring. Additionally, frequency and timing of data collection was significantly limited by external logistical factors. In an ideal scenario, site visits would have been structured to capture a variety of environmental conditions more methodically. As seen in Fig. 8, the range of SWP values varied significantly throughout the season, often containing apparent outliers. While a significant portion of this variation is likely due to field conditions, it is also possible that some measurement errors occurred while using the pressure bomb. As it is an analog device that must be operated and read manually, a certain amount of measurement error is to be expected.

The results of this study have demonstrated the potential for SWP prediction through machine learning frameworks but fall short of providing a truly promising basis for development of commercially viable irrigation decision tools. Even if these models had produced strong spatial predictions, the framework utilized for data collection and modeling must be greatly simplified to become practical in industry settings. Although the goal of this study was to find a replacement for labor intensive pressure bomb measurements, the collection of data and modeling processes described require just as much if not more labor. Due to its complexity, it certainly cannot be performed quickly enough to detect water stress in time to be properly addressed by a grower. What results have demonstrated though is that certain multispectral reflectance values can be strong predictors of water status and can quickly be gathered over a very large area. Using this knowledge, it may be possible to develop a greatly simplified modeling approach which allows SWP predictions to be generated on a same-day basis, perhaps in combination with strategically placed soil moisture sensors. This study does however inform several recommendations for future work. Whenever possible, studies should utilize larger training datasets for model training. If possible, increasing the frequency and regularity of data collection may allow researchers to better capture a meaningful variety of water stress levels. More robust data would likely help to eliminate some of the ambiguities encountered while analyzing model performance and likely would increase model and map performance in all aspects. Additionally, large datasets collected from a variety of regional geographies could help to increase model applicability. Increasing spatial sampling density and including more "high spatial density" variables, such as remotely sensed canopy temperature may also improve results. Alternate spatial interpolation methods may also be investigated in order to improve the spatial representation of predictive variables. Lastly, future studies should investigate the utility of SWP maps as an irrigation scheduling decision guide, in which irrigation events for different management zones are triggered once predicted SWP reaches a predetermined value.

Conclusions

This study sought to investigate the viability of using ANN and RF machine learning models and multispectral remote sensing to predict stem water potential in almond orchards. Its findings indicate overall that workflows combining machine learning and remote sensing can produce acceptable values of SWP at locations of trees whose data was used in the training of the models.

However, the performance of the machine learning models was unacceptable when extrapolated to non-sampled locations. Given the water use challenges facing the California almond industry, irrigation decision support tools based on this study's workflow may prove to be an integral component in developing more robust irrigation scheduling using tree water status as feedback which has been hampered in the past by low adoption of the pressure chamber. However, the accuracy and turn-around time of spatial predictions must be greatly improved (e.g., using more data in the training of machine learning models). At the very least, the data collection and image processing, and machine learning modeling must be automated or greatly simplified to truly reduce labor costs and enhance adoption. Larger training datasets with a high spatiotemporal resolution of the explanatory variables will also be necessary to achieve acceptable levels of spatial prediction of stem water potential. Nonetheless, this study has demonstrated the potential of predicting stem water potential using machine learning and multispectral remote sensing at least at tree locations whose data was used in model training which can provide a low-cost feedback for irrigation scheduling in the absence of stem water sensors or pressure chamber.

Declarations

References

Almond Board of California. (2019). Supplemental File 1 Irrigation.

Andreu J, Hopmans JW, Schwankl LJ (1997) Spatial and temporal distribution of soil water balance for a drip-irrigated almond tree. Agric Water Manag 35(issue 1–2):123–146

Belgiu M, Druaguct L (2016) Random forest in remote sensing: a review of applications and future directions. ISPRS J Photogramm Remote Sens 114:24–31

Berni J, Zarco-Tejada PJ, Suarez L, Fereres E (2009) Thermal and narrowband multispectral remote sensing for vegetation monitoring from an unmanned aerial vehicle. IEEE Trans Geosci Remote Sens 47(3):722–738. https://doi.org/10.1109/tgrs.2008.2010457

Blaya-Ros PJ, Blanco V, Domingo R, Soto-Valles F, Torres-Sánchez R (2020) Feasibility of low-cost thermal imaging for monitoring water stress in young and mature sweet cherry trees. Appl Sci 10:5461. https://doi.org/10.3390/app10165461

California Irrigation Management Information System (CIMIS). (2024). https://cimis.water.ca.gov/

Campbell GS, Campbell MD (1982) Irrigation scheduling using soil moisture measurements: theory and practice. In: Hillel DJ (ed) Advances in Irrigation, vol 1. Academic Press, New York, pp 25–42

Carter GA (1991) Primary and secondary effects of water content on the spectral reflectance of leaves. Am J Bot 78(7):916–924

Colomina I, Molina P (2014) Unmanned aerial systems for photogrammetry and remote sensing: a review. ISPRS J Photogramm Remote Sens 92:79–97

David Goldhamer & Robert Beede (2004) Regulated deficit irrigation effects on yield, nut quality and water-use efficiency of mature pistachio trees. J Hortic Sci Biotechnol 79(4):538–545. https://doi.org/10.1080/14620316.2004.11511802

DeJonge KC, Taghvaeian S, Trout TJ, Comas LH (2015) Comparison of canopy temperature-based water stress indices for maize. Agric Water Mgmt 156:51–62. https://doi.org/10.1016/j.agwat.2015.03.023

Dhillon R, Rojo F, Upadhyaya SK, Roach J, Coates R, Delwiche M (2018) Prediction of plant water status in almond and walnut trees using a continuous leaf monitoring system. Precision Agric 20(4):723–745. https://doi.org/10.1007/s11119-018-9607-0

Drechsler K, Kisekka I (2022) Variety specific irrigation of almonds during hull split, effects on yield and quality. Agric Water Manag 271:107770. https://doi.org/10.1016/j.agwat.2022.107770

Drechsler K, Kisekka I, Upadhyaya S (2019) A comprehensive stress indicator for evaluating plant water status in almond trees. Agric Water Manag 216:214–223. https://doi.org/10.1016/j.agwat.2019.02.003

Durigon A, Lier QD (2013) Canopy temperature versus soil water pressure head for the prediction of crop water stress. Agric Water Manag 127:1–6. https://doi.org/10.1016/j.agwat.2013.05.014

Espinoza CZ, Khot LR, Sankaran S, Jacoby PW (2017) High resolution multispectral and thermal remote sensing-based water stress assessment in subsurface irrigated grapevines. Remote Sens 9:961

Fulton A, Buchner R, Gilles C, Olson B, Bertagna N, Walton J, Schwankl L, Shackel K (2001) Rapid equilibration of leaf and stem water potential under field conditions in almonds, walnuts, and prunes. HortTechnology Horttech 11(4):609–615

Fulton, A. (2018, September 3). *Good pressure chamber field measurement technique*. Sacramento valley orchard source. https://www.sacvalleyorchards.com/manuals/stem-water-potential/good-pressure-chamber-technique/

Fulton, A. (2019, March 5). *Advanced SWP interpretation in almond*. Sacramento valley orchard source. Retrieved January 17, 2023, from http://www.sacvalleyorchards.com/manuals/stem-water-potential/advanced-swp-interpretation-in-almond

Gislason PO, Benediktsson JA, Sveinsson JR (2006) Random forests for land cover classification. Pattern Recogn Lett 27(4):294–300

Goldhamer DA (2005) Tree water requirements and regulated deficit irrigation. In: Ferguson L (ed) Pistachio production manual, 4th edn. Fruit and Nut Research and Information Center, University of California, Davis, pp 103–116

Gutierrez S, Diago MP, Fernandez- Novales J, Tardaguila J (2018) Vineyard water status assessment using on-the-go thermal imaging and machine learning. PLoS ONE 13(2):e0192037. https://doi.org/10.1371/journal.pone.0192037

Hawkins, E. (2017, January 12). Using soil electrical conductivity (EC) to delineate field variation. Ohioline. https://ohioline.osu.edu/factsheet/fabe-565

Hsu KL, Gupta HV, Sorooshian S (1995) Artificial neural network modeling of the rainfall-runoff process. Water Resour Res 31:2517–2530

Idso SB, Jackson RD, Reginato RJ (1977) Remote-sensing of crop yields. Science 196:19–25

Idso S, Jackson R, Pinter P, Reginato R, Hatfeld J (1981) Normalizing the stress-degree-day parameter for environmental variability. Agric Meteorol 24:45–55

Jackson RD, Idso SB, Reginato RJ, Pinter PJ (1981) Canopy temperature as a crop water stress indicator. Water Resour Res 17(4):1133–1138. https://doi.org/10.1029/wr017i004p01133

Jones HG (1999) Use of infrared thermometry for estimation of stomatal conductance as a possible aid to irrigation scheduling. Agric for Meteorol 95(3):139–149. https://doi.org/10.1016/S0168-1923(99)00030-1

King B, Shellie K (2016) Evaluation of neural network modeling to predict non-water-stressed leaf temperature in wine grape for calculation of crop water stress index. Agric Water Manag 167:38–52

Krause P, Boyle DP, Bäse F (2005) Comparison of different efficiency criteria for hydrological model assessment. Adv Geosci 5:89–97. https://doi.org/10.5194/adgeo-5-89-2005

Loggenberg K, Strever A, Greyling B, Poona N (2018) Modelling water stress in a shiraz vineyard using hyperspectral imaging and machine learning. Remote Sensing 10(2):202. https://doi.org/10.3390/rs10020202

Martí P, Gasque M, González-Altozano P (2013) An artificial neural network approach to the estimation of stem water potential from frequency domain reflectometry soil moisture measurements and meteorological data. Comput Electron Agric 91:75–86. https://doi.org/10.1016/j.compag.2012.12.001

Meyers JN, Kisekka I, Upadhyaya SK, Michelon GK (2019) Development of an artificial neural network approach for predicting plant water status in almonds. Trans ASABE 62(1):19–32. https://doi.org/10.13031/trans.12970

Meyers, H (2018). CAST [source code] https://cran.r-project.org/web/packages/CAST/index.html

Möller M, Alchanatis V, Cohen Y, Meron M, Tsipris J, Naor A, Ostrovsky V, Sprintsin M, Cohen S (2007) Use of thermal and visible imagery for estimating crop water status of irrigated grapevine. J Exp Bot 58:827–838

Naor A (2008) Water stress assessment for irrigation scheduling of deciduous trees. Acta Hort 792(792):467–481

Otto, S.A. (2019, Jan.,7). How to normalize the RMSE. Retrieved from https://www.marinedatascience.co/2019/01/07/normalizing-the-rmse/

Parker TA, Palkovic A, Gepts P (2020) Determining the genetic control of common bean early-growth rate using unmanned aerial vehicles. Remote Sensing 12(11):1748

Poblete T, Ortega-Farías S, Moreno M, Bardeen M (2017) Artificial neural network to predict vine water status spatial variability using multispectral information obtained from an unmanned aerial vehicle (UAV). Sensors 17(11):2488. https://doi.org/10.3390/s17112488

Pôças I, Gonçalves J, Costa PM, Gonçalves I, Pereira LS, Cunha M (2017) Hyperspectral-based predictive modelling of grapevine water status in the Portuguese douro wine region. Int J Appl Earth Obs Geoinf 58:177–190. https://doi.org/10.1016/j.jag.2017.02.013

Romero M, Luo Y, Su B, Fuentes S (2018) Vineyard water status estimation using multispectral imagery from an UAV platform and machine learning algorithms for irrigation scheduling management. Comput Electron Agric 147:109–117. https://doi.org/10.1016/j.compag.2018.02.013

Santesteban LG, Miranda C, Marín D, Sesma B, Intrigliolo DS, Mirás-Avalos JM, Escalona JM, Montoro A, de Herralde F, Baeza P, Romero P, Yuste J, Uriarte D, Martínez-Gascueña J, Cancela JJ, Pinillos V, Loidi M, Urrestarazu J, Royo JB (2019) Discrimination ability of leaf and stem water potential at different times of the day through a meta-analysis in grapevine (Vitis vinifera L.). Agric Water Manag 221:202–210

Starr GC (2005) Assessing temporal stability and spatial variability of soil water patterns with implications for precision water management. Agric Water Manag 72(3):223–243

Sumner, D., William Matthews, Medellin-Azuara, J., & Bradley, A. (2014). (rep.). The Economic Impacts of the California Almond Industry.

Vanella D, Peddinti RS, Kisekka I (2022) Unravelling soil water dynamics in almond orchards characterized by soil-heterogeneity using electrical resistivity tomography. Agric Water Manag 269:107652. https://doi.org/10.1016/j.agwat.2022.107652

Virnodkar S, Pachghare V, Patil VC, Jha S (2020) Remote sensing and machine learning for crop water stress determination in various crops: a critical review. Precision Agric 21:1121–1155. https://doi.org/10.1007/s11119-020-09711-9

Yang M, Gao P, Zhou P, Xie J, Sun D, Han X, Wang W (2021) Simulating canopy temperature using a random forest model to calculate the crop water stress index of Chinese brassica. Agronomy 11(11):2244. https://doi.org/10.3390/agronomy11112244

Zhang F, Zhou G (2019) Estimation of vegetation water content using hyperspectral vegetation indices: a comparison of crop water indicators in response to water stress treatments for summer maize. BMC Ecol 19(1):18. https://doi.org/10.1186/s12898-019-0233-0

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.