Precision Agriculture (2020) 21:1121–1155
https://doi.org/10.1007/s11119-020-09711-9

Remote sensing and machine learning for crop water stress determination in various crops: a critical review

Shyamal S. Virnodkar[1] · **Vinod K. Pachghare**[1] · **V. C. Patil**[2] · **Sunil Kumar Jha**[2]

Published online: 17 February 2020

Abstract

The remote sensing (RS) technique is less cost- and labour- intensive than ground-based surveys for diverse applications in agriculture. Machine learning (ML), a branch of artificial intelligence (AI), provides an effective approach to construct a model for regression and classification of a multivariate and non-linear system. Without being explicitly programmed, machine learning models learn from training data, i.e., past experience. Machine learning, when applied to remotely sensed data, has the potential to evolve a real-time farm-specific management system to reinforce farmers' ability to make appropriate decisions. Recently, the use of machine learning techniques combined with RS data has reshaped precision agriculture in many ways, such as crop identification, yield prediction and crop water stress assessment, with better accuracy than conventional RS methods. As agriculture accounts for approximately 70% of the worldwide water withdrawals, it must be used in the most efficient way to obtain maximum yields and food production. The use of water management and irrigation based on plant water stress have been demonstrated to not only save water but also increase yield. To date, RS and ML-based results have encouraged farmers and decision-makers to adopt this technology to meet global food demands. This phenomenon has led to the much-needed interest of researchers in using ML to improve agriculture outcomes. However, the use of ML for the potential evaluation of water stress continues to be unexplored and the existing methods can still be greatly improved. This study aims to present an overall review of the widely used methods for crop water stress monitoring using remote sensing and machine learning and focuses on future directions for researchers.

Keywords Remote sensing · Machine learning · Crop water stress · Crops

Abbreviations

AHS	Airborne hyperspectral scanner
ANFIS	Adaptive neuro-fuzzy inference system
ANN	Artificial neural network

✉ Shyamal S. Virnodkar
shyamal@somaiya.edu

[1] Department of Computer Engineering and IT, College of Engineering Pune, Savitribai Phule Pune University, Pune 411005, India

[2] K. J. Somaiya Institute of Applied Agricultural Research, Saidapur, Karnataka, India

BRNN	Bayesian regularized neural network
CAR	Chlorophyll absorption ratio
CFS	Correlation feature selection
CIT	Conditional inference tree
CWSI	Crop water stress index
DEM	Digital elevation model
DI	Deficit irrigation
DMSV	Digital multispectral video
DT	Decision tree
DVI	Difference vegetation index
ELM	Extreme learning machine
EM	Electromagnetic
EO	Earth observation
ET	Evapotranspiration
ETM+	Enhanced thematic mapper plus
FCBF	Fast correlation based filter
FDR	Frequency domain reflectometry
GA	Genetic algorithm
Gc	Canopy conductance
GI	Green index
GIS	Geographical information system
GLM	Generalized regression model
GNDVI	Green normalized difference vegetation index
IRECI	Inverted red edge chlorophyll index
Kc	Crop coefficient
LASSO	Least absolute shrinkage and selection operator
LDA	Linear discrimination
LSE	Land surface emissivity
LSSVM	Least squares support vector machine
LST	Land surface temperature
LWP	Leaf water potential
MAE	Mean absolute error
MARS	Multivariate adaptive regression splines
MCARI	Modified chlorophyll absorption in reflectance index
MDA	Mean decrease in accuracy
MDG	Mean decrease in Gini
MDWI	Maximum difference water index
mNDVI	Modified NDVI
MPa	Megapascal pressure unit
MSAVI	Modified soil adjusted vegetation index
MSI	Moisture stress index
MSR	Modified simple ratio
MTVI	Modified triangular vegetation index
N	Nitrogen
ND	Novelty detection
NDGI	Normalized difference greenness vegetation index
NDII	Normalized difference infrared index
NDRE	Normalized difference red edge index
NDRI	Normalized difference ratio index

NDVI	Normalized difference vegetation index
NDWI	Normalized difference water index
NIR	Near infrared
NRI	Normalized reflectance index
OSAVI	Optimal soil adjusted vegetation index
OSH	Optimal separation hyperplane
PA	Precision agriculture
PCA	Principal component analysis
PLS	Partial least squares
PRD	Partial root drying
PRI	Photochemical reflectance index
PRI norm	Normalized photochemical index
RBF	Radial basis function
RDVI	Renormalized difference vegetation index
RE	Relative error
REIP	Red edge inflection point
RF	Random forest
RGRI	Red green ratio index
RI	Ratio index
RMSE	Root mean square error
RWC	Relative water content
SADFAT	Spatio-temporal adaptive data fusion algorithm for temperature
SAM	Spectral angle mapper
SEBAL	Surface energy balance algorithm for land
SIPI	Structure insensitive pigment index
SRI	Simple ratio index
SRTM	Shuttle radar topography mission
SRWI	Simple ratio water index
SSCM	Site specific crop management
ST	Stress time index
SVM	Support vector machine
SWIR	Short wave infrared
TCARI	Transformed chlorophyll absorption reflectance index
TDR	Time domain reflectometry
TIR	Thermal infrared
UAV	Unmanned aerial vehicle
VI	Vegetation indices
VIS	Visible
VIT	Vegetation index temperature
VPD	Vapour pressure deficit
VPG	Vapour pressure gradient
WABI-1	Water balance indices–1
WABI-2	Water balance indices–2
WABI-3	Water balance indices–3
WI	Water index
WNDVI	Weighted NDVI
WP	Water productivity
XGBoost	Extreme gradient boosting

List of symbols

$\sum$dRE Sum of red edge first derivatives

g_L Leaf conductance

Introduction

Agriculture plays a key role in the economy of many countries, especially in developing countries such as India and Brazil. Soil health, climate change, humidity, water supply, pollution, rainfall, pests and weeds are all factors that impact whether a high agricultural yield can be achieved. Precision agriculture (PA), also known as site-specific crop management (SSCM), is an approach to farm management that uses information technology to ensure that crops and soil receive exactly what they need for good health and productivity. PA is based on observing, measuring and responding to inter- and intra-field spatial variability in crops and soils. The aim of PA is to ensure sustainability, profitability and protection of the environment. The approach includes accessing real-time data about, *inter alia*, the conditions of crops, soil and evapotranspiration. The benefits of PA are improving crop productivity and farm profitability, improving precise hybrid selection and the matching of fertilizer application and decreasing chemical bills and fuel costs. Mulla (2013) mentioned key advances in remote sensing applications in PA and identified the knowledge gaps. PA applications initially worked with ground sensors for soil organic matter and diversified to include vehicle-, aircraft- and satellite-mounted sensors. Real-time crop health monitoring that does not affect the environment or crop health is possible when remote sensing is used. Mulla (2013) suggested a need for developing precision farming approaches that can provide customized management of farm inputs for an individual plant. Further suggestions by Mulla (2013) include working on chemometric or spectral decomposition methods of analysis, developing sensors to estimate nutrient deficiencies, developing additional spectral indices and integrating historical archives of satellite data with real-time data. Remote sensing is a means of obtaining and analysing data about an object or phenomenon without contact with the object or phenomenon that is under investigation. Remote sensing systems are categorized into sensor-based systems and platform-based systems. Active sensors (backscatter-based measurements) and passive sensors (reflectance-based measurements) (Mulyono, et al. 2016) are two types of sensors that capture the reflectance in the electromagnetic (EM) spectrum. Specific remote sensing platforms such as ground vehicles, airplanes, satellites and handheld gadgets are utilized to mount the sensor. The data acquired by the sensors rely upon four resolutions, spatial, temporal, radiometric and spectral and these data are analysed and prepared to utilize in assorted applications. Machine learning has the ability to process large amounts of information in a non-linear framework. As remote sensing creates much information, ML algorithms are suitable analysis methods. Various machine learning algorithms, such as decision trees (DTs), support vector machines (SVMs), artificial neural networks (ANNs), genetic algorithms (GAs) and ensemble learning, have been used effectively on remotely sensed information in farming with high precision. Another serious issue of remote sensing in agribusiness is the acquisition of additional ground truth samples; however, this problem is overcome by SVM algorithms without influencing the exactness of the results as a result of the ability of SVMs to prepare models while utilizing few samples (Mountrakis et al. 2011).

Remote sensing applications in farming vary from crop classification, harvest arrangement, crop yield forecast, disease detection and management, evaluation of crop wellbeing and crop water stress detection. Detection of crop water stress in different growing seasons is necessary to predict yield conditions and plan irrigation scheduling. Different methodologies have been investigated to distinguish crop water stress. These methods are based on soil water measurements, plant responses and remote sensing. The main aim of this study was to review the crop water stress detection approaches for various crops worldwide that utilize different remote sensing methods and machine learning algorithms. The results of the research are additionally incorporated in the present study, which reveal that diverse methodologies are effectively utilized for specific crops. This review is based on a detailed study of the literature published in the main remote sensing journals.

Crop water stress detection methods

Water is a key contributing component to the quality and amount of developed yields. Water stress is a physiological response of plants when water availability is diminished. Harvest water pressure is a lack of water, which is distinguished as a reduction in the soil water content or from the physiological reactions of the crop to water shortage (Ihuoma and Madramootoo 2017). Crop water stress reduces photosynthesis and transpiration in plants. In areas with insufficient rainfall, a proper quantity of water to be fed to crops is essential to maintain crop yields and soil conditions. Supplying more water than necessary to the field also leads to soil erosion, loss of nutrients and damages the health of crops and soil. Water scarcity is another serious problem in arid and semi-arid areas. Proper water management is therefore essential in such regions where irrigation is a key factor to attain the desired crop yield, crop quality and water utilization. To control irrigation management and scheduling, one should know the quantity and timings of the water supply, which can be determined with a proper spatial evaluation of plant water stress. A comparative analysis of conventional and modern crop water stress assessment methods is provided in Table 1 and briefly discussed below.

Field measurement-based methods

Methods based on soil water measurements

The traditional methods for crop water stress detection are based on in situ soil moisture measurements and meteorological variables to assess water loss from a soil–plant system (Gonzalez-Dugo et al. 2006). Soil samples are collected from a few points of the entire field with assumptions of uniform water holding capacity, the same soil structure and the same evapo-transpiration rate, which are sometimes deceptive in reality. These methods provide point information that does not reflect the entire area and are laborious. Other soil-based methods to detect crop water stress include gravimetric soil water measurements (Tanriverdi et al. 2016; Sharma et al. 2018), soil moisture sensor measurements (Enciso et al. 2007) and soil water balance calculations (Ihuoma and Madramootoo 2017).

Methods based on plant responses

Later, plant-based approaches were adopted that were more sensitive than the soil moisture-based approaches, which included stomatal conductance, leaf water potential, relative water content, stem and fruit diameter and sap flow measurements (Fernandez 2017;

Table 1 Comparative analysis of crop water detection methods

Approaches	Methods	Description	Advantages	Disadvantages	References
Soil water measurement-based indirect approach	Time domain reflectometry (TDR)	An electromagnetic technique based on the principle of the dielectric constant of water and other substances such as soil	✓ Less destructive and less time-consuming than gravimetric methods ✓ Saves labour costs ✓ Measurements can be taken at the required soil depth ✓ Not sensitive to soil textures	✓ Expensive equipment ✓ Calibration curves depend on soil texture ✓ Environmentally sensitive ✓ Requires soil texture-specific calibration curves	Sharma et al. (2018); Tanriverdi et al. (2016)
	Frequency domain reflectometry (FDR)	Based on dielectric properties of soil	✓ Non-destructive ✓ Accurate measurements	✓ Sensitive to the environment	Sharma et al. (2018)
	Neutron probe	Measures volumetric water content in soil	✓ High accuracy ✓ Allows measurements at many depths ✓ Relatively easy	✓ Expensive equipment ✓ Requires licensing and monitoring ✓ Time intensive	Sharma et al. (2018); Enciso et al. (2007)
	Tensiometers	Soil water potential	✓ Easy to install and operate ✓ Inexpensive ✓ Good accuracy ✓ Suitable for high-frequency irrigation schedules or sampling	✓ Destructive ✓ Requires close contact with the soil for continuous reading	Sharma et al. (2018); Enciso et al. (2007)
	Gamma attenuation	Volumetric water content	✓ Non-destructive ✓ Fast	✓ Different soil gives different calibration curves ✓ Provides low accuracy	Sharma et al. (2018)
	Gypsum sensor (resistive)	Based on electrical resistance	✓ Simple ✓ Inexpensive ✓ Measurement of soil moisture in the same location as the field over an extended period	✓ Requires block-wise calibration, which changes with time ✓ Gives inaccurate measurements	Enciso et al. (2007)

Table 1 (continued)

Approaches	Methods	Description	Advantages	Disadvantages	References
	Hygrometric techniques	Soil water potential	✓ Requires low maintenance ✓ Automated measurements and suitable for irrigation system control	✓ Cannot be used practically due to an expensive and complex system	Sharma et al. (2018)
Soil water measurement-based direct approach	Gravimetric technique	A direct method, including phases of weighing a moist sample, oven drying, reweighing and calculating a mass of water lost as a percentage of the dried soil mass	✓ Highly accurate ✓ Reliable technique ✓ Little scope for instrumental error ✓ Not affected by soil type or salinity	✓ Time-consuming ✓ Based on measurement mass ✓ Destructive ✓ Labour intensive	Sharma et al. (2018); Enciso et al. (2007); Tanriverdi et al. (2016)
Soil water balance-based approach	Soil water balance	Based on soil water balance calculations	✓ Easy to use ✓ Suitable for irrigation water indication	✓ Not very accurate ✓ Depends on evapotranspiration, precipitation and irrigation schedules	Hassan-Esfahani et al. (2015); Ihuoma and Madramootoo (2017)
Plant-based approach	Stomatal conductance	A measure of the degree of stomata opening	✓ Widely used indicator of water stress in many studies	✓ Labour intensive ✓ Not suitable for precision irrigation ✓ Measurement depends on the time of day and crop canopy structure	Fernandez (2017)
	Relative water content	Measures water status in terms of the physiological consequence of cellular water deficit	✓ A direct and useful indicator of the state of the water balance of a plant	✓ Destructive and time-consuming	Krishna et al. (2019)
	Leaf water potential	A direct and good indicator of leaf water status	✓ Widely used reference method	✓ Destructive ✓ Slow ✓ Shows different behaviour for isohydric and anisohydric species	Cohen et al. (2005); Rallo et al. (2014); Zarco-Tejada et al. (2013); Rapaport et al. (2015)

Table 1 (continued)

Approaches	Methods	Description	Advantages	Disadvantages	References
	Pre-dawn leaf water potential	Based on quasi-equilibrium between water potential of soil and plant It is measured before sunrise	✓ Considers the influence of both soil and plant ✓ Reliable indicator for vine water stress	✓ Labour intensive ✓ Time-consuming	Poccas et al. (2017)
	Stem water potential	A direct good indicator of leaf water status	✓ Used as a benchmark for many research studies ✓ Type of leaf, nutritional status and size and shape of leaf do not have an influence on it	✓ Destructive ✓ Time-consuming ✓ Weather conditions, soil dryness and root health are affecting parameters	Poblete et al. (2017); Romero et al. (2018); Loggenberg et al. (2018); Moller et al. (2006); Baluja et al. (2012); Osroosh et al. (2015)
	Sap flow measurement	Measures through heat pulse and heat balance methods Indicator of plant water status. A direct indicator of stem water status	✓ Not labour intensive ✓ Direct and non-destructive measurement of transpiration ✓ Robust	✓ Complex instruments required	Fernandez (2017)
Remote sensing-based approach	Vegetation indices (VIs)	Vegetation indices are used to highlight the characteristics of vegetation	✓ Non-destructive with high temporal and spectral resolution	✓ Requisite image analysis is still a challenging task ✓ Precision reduces to leaf scale to canopy scale	Baluja et al. (2012); Zarco-Tejada et al. (2013); Romero et al. (2018); Poblete et al. (2017)

Table 1 (continued)

Approaches	Methods	Description	Advantages	Disadvantages	References
	Xanthophyll indices	Measures changes in photosynthetic pigment cells occurring due to water stress PRI, PRI_{norm}, normalized reflectance index (NRI) are sensitive to xanthophyll changes	✓ A good indicator of water stress in small fields	✓ Not useful for large commercial agricultural fields due to size, the structure of canopy and spatial variability of pigment concentration that affects PRI use ✓ No simulation methods for the daily development of xanthophylls cycle pigments and effects of chlorophyll, carotenoid, etc.	Zarco-Tejada et al. (2013); Poblete et al. (2017); Rapaport et al. (2015)
	Water indices	Measures the reflectance through the SWIR and near-infrared region used to represent canopy water content. Commonly used indices are WI, SRWI, NDWI, MSI	✓ Non-destructive measure of leaf water content ✓ Very good direct indicators of water stress	✓ The problem of scaling up to the canopy level	Zarco-Tejada et al. (2003); Gao (1996); Rapaport et al. (2015)
	Water balance indices	Monitors alterations in chlorophyll fluorescence and in leaf water contents through green and SWIR spectral ranges. Indices computed are WABI, WABI-1, WABI-2	✓ Performed very well at the leaf and canopy level	✓ High cost single spectral instrument is required ✓ Tree canopy physiological variations need to be considered ✓ Issue of SWIR band penetrability through thick atmospheric layers	Rapaport et al. (2015)

Table 1 (continued)

Approaches	Methods	Description	Advantages	Disadvantages	References
	RS-based ET estimation	ET is estimated as a residual of the surface energy balance equation LE = Rn-G-H Latent Energy (LE) Rn = net radiation from sky G = Heat to ground H = Heat to air	✓ A single, high-resolution thermal band is adequate and essential ✓ METRIC and SEBAL have good accuracy and consistency	✓ Deriving ET is a complex process ✓ ET is not directly measured ✓ High-resolution thermal imaging is critical	Bastiaanssen et al. (1998); Allen et al. (2007)
	Infrared thermometry based CWSI	CWSI is measured using canopy temperature and its reduction relative to ambient air temperature	✓ Depends on VPD ✓ Direct method	✓ Different baselines to be calculated for different crops ✓ Time-consuming ✓ Many parameters are required to assess CWSI	Idso et al. (1981); Jackson et al. (1981)
	LST-based CWSI	Using LST along with hot and cold pixels method to quantify CWSI	✓ Purely remote sensing technique ✓ Labour and time- non-intensive	✓ Computation of LST is cumbersome and varies according to the method followed to compute LST	Veysi et al. (2017)

Ihuoma and Madramootoo 2017). As one of the most accurate in situ methods, stem water potential (Ψ_{stem}) is used to assess water stress and can be measured by using a pressure chamber (Turner 1988).

Such methods are reliable; however, the assessment of plant water stress with in situ measurements is time-consuming and labour intensive, as they are assessed for each and every crop. Moreover, this method provides an inaccurate indication of the whole field due to heterogeneity in soil and crops.

Remote sensing-based methods

Spectral indices-based methods

With the advent of remote sensing, it is possible to cover a large field with non-invasive and productive techniques (Romero et al. 2018) to detect the spatial variability in plant water status with high temporal resolution. Remote sensing methods based on spectral vegetation indices and infrared thermometry (Ihuoma and Madramootoo 2017) are widely used for crop water stress detection because they are non-destructive and not labour- or time-intensive. The remote sensing method is extensively used in vegetation studies that make use of the spectral reflectance of crops. Spectral reflectance is a measure of the wavelength of the electromagnetic energy collected from objects on Earth. The biochemical and biophysical properties of plants, such as biomass, crop evapotranspiration and canopy water content, are related to spectral properties that are used for spectral reflectance. Mathematical combinations of two or more spectral bands are referred to as spectral indices that are applied to detect water stress in crops. Among copious spectral water and vegetation indices, the water index (WI) (Zarco-Tejada et al. 2003), normalized difference water index (NDWI) (Zarco-Tejada et al. 2003; Rapaport et al. 2015), photo-chemical reflectance index (PRI) (Zarco-Tejada et al. 2013), modified soil adjusted vegetation index (MSAVI) (Rozenstein et al. 2018), optimal soil adjusted vegetation index (OSAVI) (Romero et al. 2018; Baluja et al. 2012), normalized difference vegetation index (NDVI) (Baluja et al. 2012; Rapaport et al. 2015) and normalized difference greenness vegetation index (NDGI) (Romero et al. 2018), to name a few, have been extensively adopted to detect water stress in crops.

Infrared thermometry and CWSI-based methods

Infrared thermometry is an effective method to assess plant water stress at a local scale and is used to schedule irrigation in various crops. This method focuses on measuring the canopy temperature, which was originally suggested by Jackson et al. (1977). The variability in canopy temperature (Gonzalez-Dugo et al. 2006) and spectral indices derived using canopy temperature (Osroosh et al. 2015) have been used to indicate water stress.

The crop water stress index (CWSI), one of the most adopted indicators of plant water stress, is computed from canopy temperature. Canopy temperature is inversely related to leaf stomatal closure and transpiration. Stomatal closure is a consequence of water stress in crops, which, in turn, diminishes the transpiration rate in plants. A low transpiration rate decreases the cooling of plants; hence, canopy temperature increases, which is treated as an indicator of water stress. This concept forms the basis to develop the CWSI, which was first introduced by Jackson et al. (1977), (1981) and Idso et al. (1981). This index is based on the vapour pressure deficit (VPD) and the difference between air and canopy temperature. The CWSI based on canopy temperature and meteorological terms following Idso et al.

(1981) were used by Rud et al. (2014). The empirical CWSI equation uses upper and lower baselines. The empirical CWSI (Idso et al. 1981; Jackson et al. 1981; Veysi et al. 2017) is given in Eq. (1)

$$CWSI = \frac{\left(dT - dT_{ll}\right)}{\left(dT_{ul} - dT_{ll}\right)} \tag{1}$$

where dT is given by (Tc-Ta), which is the difference between canopy temperature (Tc) and air temperature (Ta); dT_{ll} is the lower baseline of fully watered crops; and dT_{ul} is the upper baseline of water-stressed crops. dT_{ll} and dT_{ul} are computed from the atmospheric VPD and vapour pressure gradient (VPG), respectively (Veysi et al. 2017). The upper baseline provides the difference between air and canopy temperature, which is much less in water-stressed crops, concluding that the crop lacks water. Relative humidity in air inversely affects transpiration in non-water-stressed crops. The lower baseline describes the situation for non-water-stressed crops where more transpiration takes place that lowers the canopy temperature. The lower baseline depends on the VPD, whereas the upper baseline does not.

The drawback of this approach is the necessity of knowing the non-water stress baseline, which varies from crop to crop and local climatic zones (Berni et al. 2009a). To eliminate the problem of knowing the non-water stress baseline, Jones (2013) modified the CWSI and defined a new normalized CWSI, which is described as follows.

$$CWSI = \frac{T_{canopy} - T_{wet}}{T_{dry} - T_{wet}} \tag{2}$$

where T_{canopy} is the canopy temperature captured using unmanned aerial vehicle (UAV)-borne thermal infrared (TIR), T_{wet} gives the fully transpiring canopy temperature and T_{dry} represents the water-stressed canopy temperature. T_{wet} and T_{dry} are equivalent to T_{base} and T_{max} in the original formula for the CWSI derived by Idso et al. (1981). However, normalization of the CWSI is a more complex process with changing atmospheric conditions than using VPD alone. Cohen et al. (2005) also indicated two drawbacks of using the CWSI based on canopy temperature:1. It is difficult to accurately separate canopy temperature from the soil background due to the lack of spatial resolution of handheld or airborne sensors, 2. Varying atmospheric conditions complicate the normalization of the CWSI.

LST-based CWSI

Unlike the method discussed above, where CWSI computation was performed using calculations on the data collected from ground measurements and canopy temperature, Veysi et al. (2017) determined the CWSI using only satellite image data using the following equation:

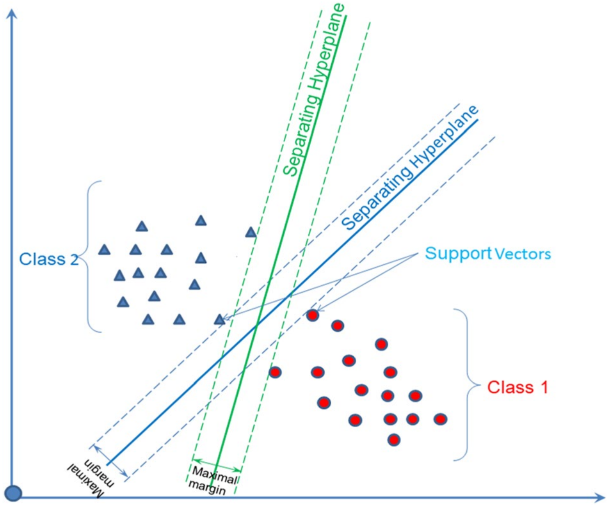
$$CWSI = \frac{T_s - T_{cold}}{T_{hot} - T_{cold}} \tag{3}$$

where T_s is the land surface temperature (LST) derived from a satellite image that gives canopy temperature, T_{cold} is the temperature of cold pixels and T_{hot} is the temperature of hot pixels. Cold pixels are those covered by fully watered crops and hot pixels represent water-stressed crops. Bastiaanssen et al. (1998) described evapo-transpiration using the surface energy balance algorithm for land (SEBAL) for the selection of hot and cold pixels, which was followed by Veysi et al. (2017) for cold pixel selection, with little changes suggested for hot pixel selection. Hot pixels are selected from the area with maximum water stress. LST is a key parameter in the biophysical processes of evapo-transpiration, water and surface energy balance (Li et al. 2013; Bai et al. 2015). LST is retrieved from the thermal infrared data of satellite imagery but is not calculated directly. LST measurements require cloud removal, radiometric calibration, emissivity and atmospheric corrections, which are challenging tasks. A remarkable study (Li et al. 2013) provided a review of the progress in LST estimations from thermal infrared data primarily captured by polar orbiting satellites. The study provides a theoretical basis to extract LST and listed the difficulties, which were related to LST, land surface emissivity (LSE), atmosphere coupling, the physical meaning of the satellite-derived LST and satellite-derived LST validation problems. They categorized the algorithms into single-channel methods, multi-channel methods and multi-angled methods with known LSEs. The methods without a priori known LSEs were categorized into a stepwise retrieval method, simultaneous retrieval of LSEs, LST with known atmospheric information and simultaneous retrieval with unknown atmospheric information. Validation of the retrieved LST can be undertaken using temperature-based methods, radiance-based methods and cross validation. The existing earth observations (EO) do not provide TIR images with detailed temporal and spatial resolution simultaneously (Bai et al. 2015). Bai et al. (2015) used Landsat enhanced thematic mapper plus (ETM+) TIR and MODIS images to retrieve the LST to overcome the problem of obtaining TIR images at a detailed spatial and temporal resolution from the available satellites. They developed a novel fusion method by combining image fusion and spatio-temporal fusion techniques to derive LST. First, an extreme machine learning algorithm was applied to enhance the spatial resolution of Landsat ETM+TIR data. After that, MODIS LST and thermal sharpened Landsat data were fused using the spatio-temporal adaptive data fusion algorithm for temperature mapping (SADFAT) (Weng et al. 2014) to derive synthetic data with high temporal resolution.

Evapotranspiration-based methods

Land surface evapotranspiration (ET) measures the amount of water lost to the atmosphere through soil evaporation and plant transpiration. ET influences water resources, water rights management and the hydrological cycle at local and regional scales. Penman (1948) established a framework for relating evapotranspiration to meteorological factors (Allen et al. 1998). Since then, considerable advances have been made in evapo-transpiration processes with energy exchanges. Conventionally, ET estimation requires meteorological data for model simulations or empirical equations. However, these techniques are not viable to effectively estimate ET at a regional scale because of the diversity in land covers or temporal changes in the landscape (Zhang and Lemeur 1995). The most frequently used method for estimating ET at present is the Penman–Monteith equation. The point-based approach makes this technique limited to the local scale and therefore is not suitable for large heterogeneous areas. There was a need to introduce the RS technique to evaluate ET at local and regional scales. Large area coverage with high-resolution imagery in an instantaneous view is possible through RS and the data can be utilized to retrieve parameters such as radiometric surface temperature, VI and albedo (Choudhury 1989); therefore, RS data are more suitable for the estimation of ET using energy balance techniques. The energy balance concept and net radiation are used as the principal parameters in most remote sensing methods used to estimate ET. There are two widely used satellite-based models for ET

Fig. 1 Support vector machine example



estimation, SEBAL (Bastiaanssen et al. 1998), which is based on visible and thermal infrared spectral radiances of dry and wetland surfaces and the mapping evapotranspiration at high resolution with internalized calibration (METRIC) (Allen et al. 2007), which is based on short wave and long wave thermal images that provide better accuracy and consistency in results. Other remotely sensed ET models include Penman–Monteith, Priestley-Taylor, surface temperature and vegetation index space (Zhang et al. 2016). However, the predictive accuracy of these methods depends on the retrieval of vegetation indices and meteorological variables obtained from remote sensing techniques (Glenn et al. 2010; Verstraeten et al. 2008).

In summary, applications of machine learning algorithms to RS data, i.e., spectral bands, parameters retrieved through LST, VI and albedo, can greatly contribute to the determination of plant water stress. Before discussing the application of ML, first, it is essential to review the machine learning algorithms widely used in crop water stress assessments.

Overview of widely adopted machine learning algorithms in agriculture

Support vector machine

Support vector machine is a statistical learning approach to classify heterogeneous data with higher accuracy than traditional statistical classifiers, without assuming a specific data distribution. This method is a supervised, non-parametric learner (Pal and Mather 2005) that can also be used for regression. The SVM classifier separates the given labelled data samples into predefined classes in a multidimensional space (Fig. 1). The SVM learner has the intention of achieving the optimal separation hyperplane (OSH), which is a decision boundary between classes that minimizes classification error in training by having the maximum margin and later generalizing to unseen data. The margin is referred to as the distance between data samples from classes. The margin of the classifier is maximized with the help of support vectors. Support vectors are data points that lie closer to the margin, mainly contributing to fitting the hyperplane. Other data points do not contribute much to the position and orientation of the hyperplane and hence are discarded. Research has shown that remotely sensed data can be accurately classified by an SVM classifier (Foody and Mathur 2004a). SVM is basically designed for binary classification but can be extended for classification of multiple classes using pair-wise coupling techniques (Khobragade et al. 2015), one-against-all, one against-others, directed acyclic graph (Mountrakis et al. 2011) and many other methods suggested by Hsu and Lin (2002) and Melgani and Bruzzone (2004).

The classification accuracy of any classifier depends on the number and selection of training samples (Khobragade et al. 2015). The collection of ground truth data is a very cumbersome and labour and cost-intensive process in remote sensing applications. For that reason, the capability of SVM to work successfully on a small number of training samples (Foody and Mathur 2004b) without compromising the classification accuracy compared to conventional methods makes this method more promising in the remote sensing domain. Overfitting in machine learning represents a model that exactly models the training data. Overfitting negatively affects model performance as it learns noise in the data. Overfitting is also called capacity control or bias-variance trade-off, which is efficiently dealt with by an SVM even with small training samples (Mountrakis et al. 2011). Ghoggali et al. (2009) combined a genetic algorithm and SVM to classify RS data with limited training samples by designating unlabelled samples using a multi-objective genetic optimization framework.

Random forest classifier

Random forest (RF) or random decision forest is an ensemble learner (Breiman 2001) that is built by constructing many weak decision trees for classification and regression. RF is a non-parametric machine learning algorithm. Bootstrap (training) samples are randomly selected from an original dataset to construct multitudinous trees with the replacement of samples. There are chances of not selecting any sample at all or selecting any sample more than once. The trees are grown in the best possible ways, i.e., pruning is not applied. The original dataset is divided into in–bag samples (two-thirds of the original data) for training the trees and out-of-bag samples (the remaining one-third of the original data) for internal cross validation to estimate the learning process error, which is termed an out-of-bag error. Each tree is built independently without pruning based on the two user-defined (hyper parameters) attributes, forming the forest. The first attribute is the number of trees (Ntree) and the other is the number of features used to split each node while creating the tree (Mtry). The forest is grown to its maximum size until each node becomes pure. The majority vote of predictions of all the trees decides the ensemble's final decision. To test new data, it runs through all the produced trees and each tree votes for a class. The class that receives the maximum votes will be the final selected class. Figure 2 depicts the training and testing phases of the random forest algorithm. One of the best advantages of RF is that it is used for both classification and regression. The classifier also produces low generalization error (Breiman 2001). As RF is an effective tool for prediction, it does not overfit because of the law of large numbers. Randomness lies in bagging and the selection of random features. Adam et al. (2017) used RF's capability to handle interactions and non-linearities among other numerical and categorical features. Mtry and Ntree values have been well investigated by many studies. Belgiu and Draguct (2016) decided on 500 as the value for Ntree for two reasons: stabilization of error and the number available in R software to train the model. The other Ntree values that were investigated were 5000, 1000

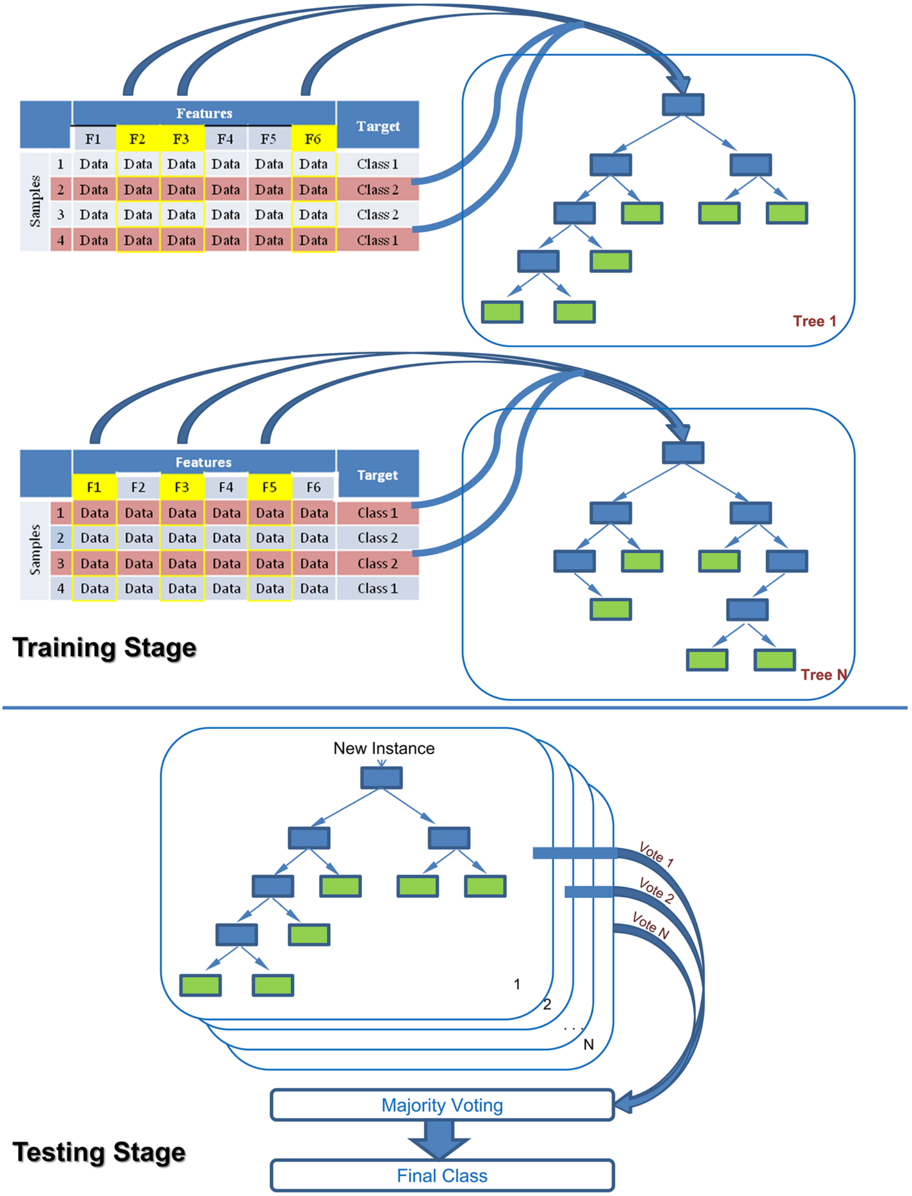


Fig. 2 Training and testing stages in the random forest algorithm

and 100 (Belgiu and Druaguct 2016). An other parameter, Mtry, can take any value up to the number of variables in the original dataset but is normally assigned as the square root of the number of features (Gislason et al. 2006). The curse of dimensionality, also known as the Hughes effect or Hughes phenomenon, says that an increase in the dimensions of the dataset increases the classification accuracy, but at some point, the accuracy begins to decrease due to the limitation of training samples (Alonso et al. 2011). More dimensions may not always necessarily produce good results. The calculation of optimal values for training samples and dimensions computed by variable importance together prove to be time- and cost-effective solutions with good classification accuracy. Variable importance is referred to as the statistical significance of every feature with respect to its contribution to the developed model and has achieved great significance, especially in high-dimensional datasets.

In RF, using the mean decrease in accuracy (MDA) (Abdel-Rahman et al. 2014) and mean decrease in Gini (MDG) (Breiman 2001; Pedergnana et al. 2013), vegetation importance is calculated. R software (R Development-Core-Team 2005; Liaw et al. 2002) is found to be a widely used tool to implement RF over Weka, Scikit-learn, MATLAB, etc. (Belgiu and Draguct 2016).

eXtreme gradient boosting (XGBoost)

The idea of boosting is enhancing a weak learner to become a better learner. In gradient boosting, lower accuracies of produced pruned trees are combined to obtain an accurate model (Loggenberg et al. 2018). Gradient boosting is implemented using the XGBoost classifier, which was designed for speed and better performance (Breiman 2001). XGBoost (Chen and Guestrin 2016) uses the information provided as feedback from the previously grown trees to build further trees and attempts to lower the error in the next iterations.

Rotation forest

Another tree-based ensemble approach is rotation forest, which differs from RF only in considering different subsets of features in the training trees (Poona et al. 2016). Feature extraction is carried out on a newly created rotated feature space using principal component analysis (PCA) (Rodriguez et al. 2006). The final decision is made, which is similar to RF. Rotation forest can be implemented in R software, Python and MATLAB.

Oblique random forest

Oblique random forest (Breiman 2001) creates trees using bagging and selects random variables for node splitting. Linear discriminant analysis (LDA), PCA, ridge regression, partial least squares (PLS) and SVM are used to split the node (Poona et al. 2016). Unlike RF, oblique RF learns the optimal split direction by using all the selected variables. R software can be used to implement oblique RF.

Artificial neural network

Remote sensing generates a very large amount of data and many sensors capture minute changes within plants. This type of non-linear problem can be analysed by applying the ANN model because of its capability to model a linear and highly non-linear relationship between input and output datasets. ANN basically consists of one input layer, one output layer and zero or more hidden layer(s), which are used to solve complex problems, as shown in Fig. 3. The ANN model learns itself by selecting appropriate values for weights (Samborska et al. 2014). The ANN model has become promising in agriculture for numerous applications, such as modelling thermal information to assess water stress (King and

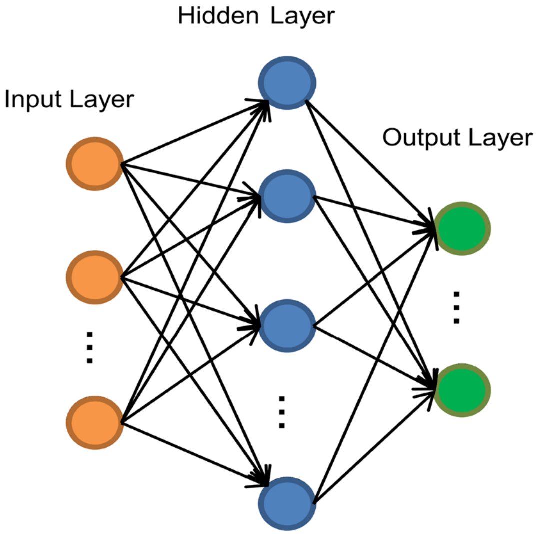


Fig. 3 Layers in an artificial neural network

Shellie 2016), vegetation mapping (Carpenter et al. 1999), yield prediction (Jiang et al. 2004; Khairunniza-Bejo et al. 2014) and prediction of nitrogen stress (Goel et al. 2003).

Integration of RS and ML for crop water stress detection

Water stress detection using VI

Among the abundant available spectral indices, many indices have been evaluated by researchers to assess water stress in different crops. Due to different platforms, spectral band combinations, instrumentation and spatial resolutions, it was difficult to reach a mathematical formula that expresses all vegetation indices. Hence, the visible (VIS) band for vegetation and the non-visible band for vegetation surface-based mathematical formulae have been developed according to the applications (Xue and Su 2017). They listed more than 100 vegetation indices with their applicability, advantages and disadvantages in their review study. Several vegetation indices in the visible spectrum domain found a good correlation with plant water status (Romero et al. 2018) in vineyard management. For grapevines, information on VIS and shortwave infrared (SWIR) (Rapaport et al. 2015), near-infrared (NIR) and SWIR (Rallo et al. 2014), VIS, green, red edge and NIR (Poccas et al. 2017) has been suggested to be good indicators of water stress. There are many spectral indices that are direct or indirect indicators of water stress. Table 2 presents the EO-based indices that have a direct relationship with the water content of crops, whereas Table 3 includes a list of the numerous indirect indices used as indicators of water status in crops. The CWSI was found to be the best indicator, with the WI and NDWI being good direct indicators of water status (Fig. 4).

The most commonly used VI is the NDVI, which uses the NIR and red bands of the EM spectrum to evaluate the health of the crop. Different studies revealed that NDVI was one of the best indirect indicators of water stress in crops (Baluja et al. 2012).

Table 2 EO-based spectral indices indicating direct water stress in plants

VI	Formula	Cultivars	R^2	References
NDWI	$(R_{860} - R_{1240})/(R_{860} + R_{1240})$	Vineyard	0.01–0.99	Zarco-Tejada et al. (2003); Gao (1996)
		Grapevines	0.04	Rapaport et al. (2015)
SRWI	R_{858}/R_{1240}	Vineyard	0.7	Zarco-Tejada et al. (2003)
	R_{680}/R_{1240}	Olive groves	0.41	Rallo et al. (2014)
WABI-1	$(R_{1490} - R_{531})/(R_{1490} + R_{531})$	Grapevines	0.72	Rapaport et al. (2015)
WABI-2	$(R_{1500} - R_{538})/(R_{1500} + R_{538})$	Grapevines	0.89	Rapaport et al. (2015)
WABI-3	$(R_{1485} - R_{550})/(R_{1485} + R_{550})$	Grapevines	0.61	Rapaport et al. (2015)
WI	R_{900}/R_{970}	Vineyard	0.95	Zarco-Tejada et al. (2003); Serrano et al. (2010)
		Grapevines	0.12	Rapaport et al. (2015)
MDWI	$\frac{\max R[1500-1750]-\min R[1500-1750]}{\max R[1500-1750]+\min R[1500-1750]}$	Grapevines	0.34	Rapaport et al. (2015)
MSI	R_{1600}/R_{820}	Grapevines	0.31	Rapaport et al. (2015)
		Olive Groves	0.48	Rallo et al. (2014)

(R stands for reflectance)

Figure 5 shows various vegetation indices that indirectly indicate the water status in crops that were investigated in different studies. Rahman et al. (2004) used NDVI for the identification of sugarcane areas and assessment of crop conditions. Sugarcane leaf water content along with various other parameters, such as nitrogen deficiency, pigments, foliar nutrients and agronomic parameters, influence the spectral response of the crop. The infrared/red ratio from the Landsat TM NIR radiometer, SWIR bands and the digital multispectral video (DMSV) sensor were useful for detecting the water content in sugarcane crops, as per the study by Abdel-Rahman and Ahmed (2008). Katsoulas et al. (2016) presented a review of crop water stress and nutrient detection through crop reflectance measurement approaches and sensors in a greenhouse. They found ground-based sensor data indices to be efficient for water stress detection but were influenced by leaf age, leaf thickness, soil background, canopy structure, etc. Water stress can be captured by a change in the canopy due to a reduction in the photosynthesis process. At the canopy level, VIS, red edge and NIR regions have been proven to be the best to detect crop water stress (Berni et al. 2009b). With different exposures and slopes, Brunini and Turco (2016) aimed to determine sugarcane water stress indices in irrigated areas. They evaluated the daily water stress index and soil water potential for sugarcane and found that the water stress index varies according to the exposure and the slope. The water stress index derived from infrared thermometry was used to determine the ideal time when sugarcane crops needed to be irrigated. They experimented with different growing phases of sugarcane (tillering, growth and maturation) surfaces with slopes from 0 to 40% and solar exposures. They noticed that the ideal time for irrigation varied with the phases of sugarcane and ranged from 2.0 to 5.0 °C. Bajwa and Vories (2006) monitored canopy temperature and reflectance-based VIs, including NDVI, green normalized difference vegetation index (GNDVI), stress time index (ST), CWSI and canopy temperature-based indices to assess the response of cotton to water stress. Rozenstein et al. (2018) estimated cotton water consumption using the crop coefficient (Kc) and 22 VIs from spectral bands. The NDVI was found to be strongly correlated with Kc and water stress (DeTar et al. 2006) for cotton crops. Jackson et al. (1977) combined remotely

Table 3 EO-based spectral indices indicating plant water stress indirectly

VI	Formula	Cultivars	R^2	References
GI	R_{550}/R_{670}	*Vitis vinifera* L. cv tempranillo	0.54	Baluja et al. (2012); Zarco-Tejada et al. (2005)
		Grapevines	0.06	Poblete et al. (2017)
			0.14	Romero et al. (2018); Zarco-Tejada et al. (2005)
DVI	$NIR - RED$	Grapevines	0.23	Romero et al. (2018)
		Grapevines	0.8959	Rozenstein et al. (2018)
RDVI	$\frac{R_{800}-R_{670}}{\sqrt{R_{800}+R_{670}}}$	Grapevines	0.05	Romero et al. (2018); Roujean and Breon (1995)
			0.00	Poblete et al. (2017); Roujean and Breon (1995)
			0.10	Baluja et al. (2012); Roujean and Breon, (1995)
		Cotton	0.49	Zarco-Tejada et al. (2013)
NDVI	$(R_{800} - R_{670})/(R_{800} + R_{670})$	*Vitis vinifera* L. cv tempranillo	0.68	Baluja et al. (2012); Rouse et al. (1974)
	$R_{800} - R_{675}/R_{800} + R_{675}$	Grapevines	0.03	Rapaport et al. (2015); Rouse et al. (1974)
	$(R_{800} - R_{670})/(R_{800} + R_{670})$		0.20	Rapaport et al. (2015); Rouse et al. (1974)
			0.35	Rapaport et al. (2015); Rouse et al. (1974)
mNDVI	$R_{750} - R_{705}/R_{750} + R_{705}$	Grapevines	0.28	Rapaport et al. (2015)
Weighted NDVI	$(NIR - (w * RED))/(NIR + (w * RED))$	Cotton	0.883	DeTar et al. (2006)
GNDVI	$NIR - GREEN/NIR + GREEN$	Cotton	0.9484	Rozenstein et al. (2018)
NDRE	$(R_{790} - R_{720})/(R_{790} + R_{720})$	Grapevines	0.05	Romero et al. (2018); Barnes et al. (2000)
NDII	$(R_{820} - R_{1650})/(R_{820} + R_{1650})$	Grapevines	0.24	Rapaport et al. (2015)
SIPI	$R_{800} - R_{445}/R_{800} + R_{680}$	Grapevines	0.17	Rapaport et al. (2015)
OSAVI	$\frac{NIR-RED}{NIR+RED+0.16}$	Grapevines	0.42	Romero et al. (2018); Rondeaux et al. (1996); Haboudane et al. (2002)
		Vineyard	0.35	Baluja et al. (2012)
MSAVI	$\frac{\left(2\times R_{800}+1-\sqrt{(2\times R_{800}+1)^2-8\times(R_{800}-R_{670})}\right)}{2}$	Grapevines	0.11	Baluja et al. (2012)
			0.00	Poblete et al. (2017)
			0.25	Romero et al. (2018); Qi et al. (1994)

Table 3 (continued)

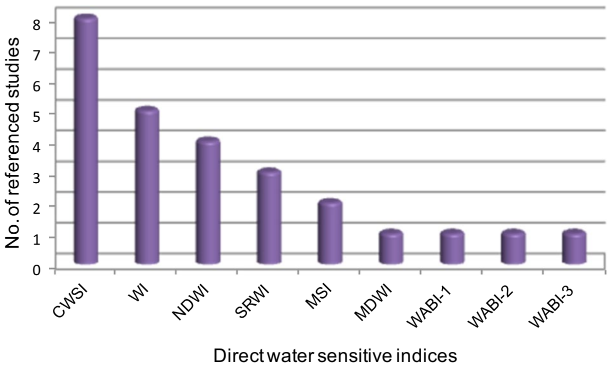
VI	Formula	Cultivars	R^2	References
RGRI	RED/GREEN	Grapevines	0.20	Romero et al. (2018); Gamon and Surfus, (1999)
SRI	NIR/RED	Grapevines	0.27	Romero et al. (2018); Birth and McVey (1968)
	R_{800}/R_{550}		0.06	Poblete et al. (2017)
		Vitis vinifera L. cv tempranillo	0.64	Baluja et al. (2012); Jordan (1969)
PRI	$(R_{570} - R_{531})/(R_{570} + R_{531})$	*Vitis vinifera* L. cvThompson seedless	0.53	Zarco-Tejada et al. (2013); Gamon et al. (1992)
	$(R_{531} - R_{550})/(R_{531} + R_{550})$	Grapevines	0.19	Rapaport et al. (2015); Gamon et al. (1992)
			0.09	Poblete et al. (2017)
	$(R_{530} - R_{570})/(R_{530} + R_{570})$	*Vitis vinifera* L cv tempranillo	0.25	Baluja et al. (2012); Gamon and Surfus (1999)
PRI norm	$PRI_{570}/((RDVI \times R_{700})/R_{670})$	Grapevines	0.82	Zarco-Tejada et al. (2013)
REIP	Max dR(680–780)	Grapevines	0.34	Rapaport et al. (2015)
$\sum$dRE	$\sum$dR(680–780)	Grapevines	0.14	Rapaport et al. (2015)
TCARI	$3 \times \left[(R_{700} - R_{670}) - 0.2 \times (R_{700} - R_{550}) \times \left(\frac{R_{700}}{R_{670}}\right)\right]$	Corn	0.81	Haboudane et al. (2002)
		Vineyard	0.45	Baluja et al. (2012)
		Grapevines	0.04	Rapaport et al. (2015)
TCARI/OSAVI	$\frac{3\times[(R_{700}-R_{670})-0.2\times(R_{700}-R_{550})\times(R_{700}/R_{670})]}{(1+0.16)(R_{800}-R_{670})/(R_{800}+R_{670})+0.16}$	Corn	0.98	Haboudane et al. (2002)
		Vitis vinifera L. cvThompson seedless	0.01	Zarco-Tejada et al. (2013)
		Grapevines	0.58	Baluja et al. (2012)
		Grapevines	0.09	Poblete et al. (2017)
	R_{680}/R_{1240}	Olive groves	0.41	Rallo et al. (2014)
MCARI	$[(R_{700} - R_{670}) - 0.2 \times (R_{700} - R_{550})] \times \left(\frac{R_{700}}{R_{670}}\right)$	Grapevines	0.01	Baluja et al. (2012)
			0.02	Poblete et al. (2017)
MCARI1	$1.2 \times [2.5 \times (R_{800} - R_{670}) - 1.3 \times (R_{800} - R_{550})]$	Grapevines	0.21	Baluja et al. (2012)
			0.03	Poblete et al. (2017)

Table 3 (continued)

VI	Formula	Cultivars	R^2	References
MCARI2	$\frac{1.2\times[2.5\times(R_{800}-R_{670})-1.3\times(R_{800}-R_{550})]}{\left(\sqrt{(2\times R_{800}+1)^2}\right)\times 6\times(R_{800}\times 5\times R_{670})\times 0.5}$	Grapevines	0.00	Baluja et al. (2012)
			< 0.01	Poblete et al. (2017)
MSR	$\left(\left(\frac{R_{800}}{R_{670}}\right)-1\right)/\left(\sqrt{\left(\frac{R_{800}}{R_{670}}\right)+1}\right)$	Grapevines	0.66	Baluja et al. (2012)
			0.34	Poblete et al. (2017)
MTVI3	$1.2\times\left[1.2\times\left(R_{800}-R_{550}\right)-2.5\times\left(R_{670}-R_{550}\right)\right]$	Grapevines	0.01	Baluja et al. (2012)
			0.03	Poblete et al. (2017)
		Olive Groves	0.48	Rallo et al. (2014)
CAR	$\frac{((R_{700}-R_{500})*670+R_{670}+(R_{550}-((R_{700}-R_{500})*670)*550))}{\sqrt{((R_{700}-R_{500})*670)^2}}$	Vineyard	0.33	Baluja et al. (2012)
CAR1	$CAR*\left(R_{700}/R_{670}\right)$	Vineyard	< 0.00	Baluja et al. (2012)
RDVI	$\left(R_{800}-R_{670}/\sqrt{\left(R_{800}+R_{670}\right)}\right)$	Vineyard	0.10	Baluja et al. (2012)

(R stands for reflectance)

Fig. 4 Direct indicators of water stress used in studies



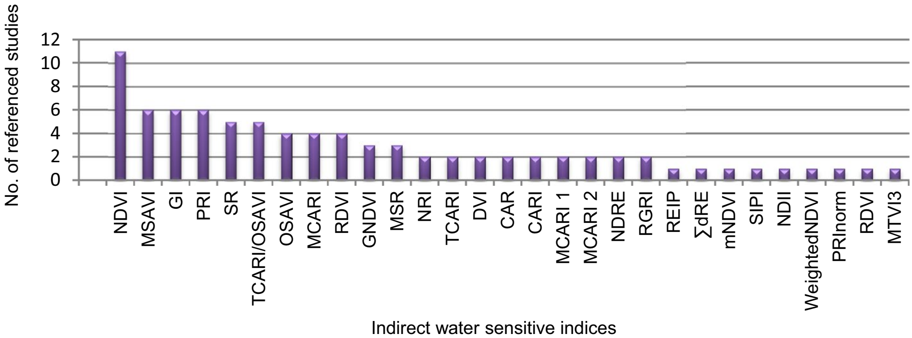


Fig. 5 Indirect indicators of water stress used in studies

measured canopy temperature with ground-based air temperature, which became a practical tool for assessing the water requirements of wheat.

Water stress detection using the CWSI

Crop water deficit or water status monitoring is essential for irrigation scheduling (Xu et al. 2018). The CWSI is capable of quantifying crop water stress 24–48 h prior to stress detection by visual observation (Kacira et al. 2002). After Idso et al. (1981), CWSI has successfully been applied to many different plants, such as wheat (Yuan et al. 2004), cotton (O'Shaughnessy and Evett. 2010), maize (Romano et al. 2011), potato (Ramirez et al. 2016), bean (Erdem et al. 2006), some vegetables (Cremona et al. 2004; Rud et al. 2014) and fruits (Paltineanu et al. 2009).

The capability of an empirical CWSI under varying irrigation systems, such as surface and sub-surface drip systems, was evaluated by Colak et al. (2015) to determine the effect of water stress on the yield and water use efficiency of yield, along with the effects of deficient irrigation (DI) and partial root drying (PRD), on the yield and water relations in eggplant. Colak et al. (2015) concluded that eggplants should be irrigated at CWSI values between 0.18 and 0.20 and parameters of irrigation significantly affecting the yield were the growth area, irrigation method, irrigation intervals and irrigation levels. An empirical formulation of the CWSI was also utilized by Cohen et al. (2005) in addition to canopy temperature derived from thermal images to predict the leaf water potential (LWP) using a regression model in cotton plants under a range of irrigation regimes. Cohen et al. (2005) found a good relationship between LWP and CWSI that was stronger than that between LWP and canopy temperature. Furthermore, Cohen et al. (2005) focused on developing a procedure for water stress mapping that combined the LWP estimation model with spatial structure analysis.

The CWSI based on the RS technique is more stable and feasible for irrigation management (Rud et al. 2014; Bai et al. 2015) at local and regional levels than an empirical CWSI. Taghvaeian et al. (2012) developed water-stressed and non-water-stressed baselines in a region of Colorado USA for irrigated maize by using infrared thermometry with few weather parameters. Furthermore, they estimated a remote sensing-based CWSI, which revealed that the data collection time was a key parameter in utilizing the CWSI approach. Their major contribution was in identifying irrigation timing and estimating irrigation requirements. Veysi et al. (2017) proposed a new procedure to calculate the CWSI from satellite data using hot and cold pixels without considering the ground ancillary data for irrigation scheduling during the sugarcane growing season (May–September) and they found that this procedure outperformed the other two approaches with a good coefficient of determination. Veysi et al. (2017) further noticed that VWC was negatively related to the CWSI with R^2 values of 0.42–0.78. Eight Landsat 8 satellite images were captured along with ground truth data, which were collected using in situ measurements of canopy temperature and VWC to validate the results of the new approach. The CWSI derived from a UAV airborne hyperspectral scanner (AHS) and in situ measurements in olive orchards were mapped with spatially distributed canopy conductance by Berni et al. (2009a). The correlation between field-measured leaf stomatal conductance and AHS imagery was biased in the radiometric calibration or atmospheric correction, with an R^2 value of 0.59. They found a good relationship between the estimated CWSI from UAV thermal imagery, with LWP having an R^2 of 0.82 and canopy conductance having an R^2 of 0.91. Berni et al. (2009a) validated the model against ground thermal sensors and used airborne remote sensing thermal imagery, concluding that, for heterogeneous olive orchards, energy balance equations and the theoretical formulation of the CWSI can be combined to compute canopy conductance (Gc) and the CWSI, which can be used to obtain actual evapotranspiration and schedule irrigation. Moller et al. (2006) worked on CWSI determination for grapevines by fusing thermal and visible imagery. Leaf conductance (g_L), stem water potential and leaf area index along with meteorological parameters were considered to calculate the CWSI. Although excess water supply or water stress negatively impacts crop yield and quality, it is useful to have slight to moderate water deficits to ensure optimal quality in the cultivation of grapevines. Moller et al. (2006) aimed to compare a thermal-based CWSI with plant water status, test various reference surfaces and determine the relationship between stem water potential and stomatal conductance with thermal visible images. Their results concluded that a strong correlation existed between the CWSI and leaf conductance compared to the correlation between the CWSI and stem water potential. Hyperspectral, multispectral and thermal data were explored to measure nitrogen (N) and water stress in wheat (Tilling et al. 2007). Thermal images were used to quantify water stress for the full canopy and the 2D CWSI and vegetation index temperature (VIT) trapezoid method was used for partially covered vegetation fields. Their findings state that irrigated fields are consistently less stressed than rain-fed fields.

To date, the CWSI has been assessed using thermal, UAV, hyperspectral and multispectral data and it has been found that the CWSI produces better results than other conventional and RF-based VI methods.

Water stress detection using ML

Over the past few decades, machine learning techniques have been progressively used in diverse applications of remote sensing. Using the GA and ML techniques, a model was developed by Hassan-Esfahani et al. (2015) from Landsat images, local weather data and field measurements and this model reported field conditions using a soil balance approach. This model comprises two modules:

- Water allocation optimization
- Soil water balance model components forecasting

Optimal crop water application rates based on the crop type, sensitivity to water stress and growth stage have been identified in the optimization module by employing GA. The output of this module is given to the forecasting module, which allocates water across the area covered by the centre pivot irrigation system. The model was evaluated on alfalfa and oats, resulting in 20% less water use. Sun et al. (2017) designed a crop water stress system across two platforms, a multi-core high-performance computing platform (SPARTAN) and a cloud platform (NeCTAR), to support parallelism of the analysis of thermal images. These thermal images were captured by UAV and underwent the process of first detecting edges, then building a Gaussian mixture model for each crop species and finally calculating the water stress index according to the mean value from the Gaussian model.

Other well-known ML techniques, SVM and RF, are regularly considered classical data-driven techniques and are popular in many remote sensing applications, mainly including crop classification (Yang et al. 2011; Saini and Ghosh 2018) regression (Kaheil et al. 2008) and LULC mapping (Warner and Nerry 2009; Huang et al. 2008). The popularity of SVM is due to its several promising characteristics, such as the kernel trick and structural risk minimization principle (Vapnik 1999). The selection of the kernel trick influences the generalization ability of SVM in many remote sensing applications (Mountrakis et al., 2011). The popularity of RF is because of its ability to address data overfitting. Nevertheless, very few studies have been carried out on the applicability of SVM and RF in determining crop water stress. For instance, Poccas et al. (2017) selected three hyperspectral reflectance vegetation indices (NIR, WI and D1) and the day of the year predictors for the inclusion in RF and SVM predictive machine learning models to model predawn leaf water potential for assessing water stress in grapevines. Moshou et al. (2014) attempted to discriminate healthy and water-stressed wheat canopies grown in a greenhouse environment. They made use of a spectrograph and a fluorimeter for their study, but remote or vehicle-mounted sensing could also be used. They developed a hybrid classification technique with a multisensory fusion system and least squares support vector machine (LSSVM), which was able to detect and discriminate between two stress factors, namely the onset of Septoria tritici disease and water stress in winter wheat. LSSVM displayed 99% performance in their investigation. AlSuwaidi et al. (2018) designed an innovative classification framework to analyse hyperspectral data to detect plant diseases, crop stress conditions and crop type classification. Their framework comprised spectral profile extraction, significant wavelength selection, novelty detection classifier construction and ensemble learning. Leaf pixel values were considered to provide spectral profiles. ReliefF, chi-square, Gini index, information gain, fast correlation-based filter (FCBF) and correlation feature selection (CFS) algorithms were employed for optimal feature selection. Novelty scores using novelty detection (ND) SVM were used to detect novelty and the final decision was made using ensemble majority voting. Loggenberg et al. (2018) combined terrestrial hyperspectral remote sensing with machine learning to model water stress in vineyards. They applied RF and XGBoost to discriminate stressed and non-stressed Shiraz vines. They compared the results with in-field stem water potential. Moreover, the utility of the spectral subset of wavebands derived using the gains from RF MDA and XGBoost was evaluated. Key parameters of XGBoost were established as follows: max_depth = 6, subsample = 1, eta = 0.3, nrounds = 100–1000, gamma = 0, min_child_weight = 1 and colsample_bytree = 1. Loggenberg et al. (2018) expressed their willingness to further investigate the development of the framework's robustness and operational capabilities. The achieved results were quite noticeable; for all wavebands (p = 176), the RF test accuracy was 83.3% (KHAT = 0.67) and the XGBoost test accuracy was 78.3% (KHAT = 0.6). For the subset of wavebands (p = 18), the RF test accuracy was 83.3% (KHAT = 0.67) and the XGBoost test accuracy was 80.0% (KHAT = 0.6). However, RF and SVM algorithms are rarely applied for determining water status, unlike ANN. ANN is a widely utilized ML technique in water stress detection and other studies in agriculture and is good at tackling agricultural issues where deterministic models are inaccessible. Romero et al. (2018) observed aerial multispectral imagery for various vegetation indices, such as the difference vegetation index (DVI), green index (GI), MSAVI, NDVI, NDGI, NDRE, OSAVI, red green ratio index (RGRI), renormalized difference vegetation index (RDVI) and simple ratio index (SRI). These indices have been applied as inputs to the model. Then, correlations between midday stem water potential (Ψ_{stem}) and VIs were estimated and evaluated using statistical methods and machine learning algorithms for vineyard studies. The research focused on the building of two models. The first model was built using ANN with (Ψ_{stem}) and VIs and showed high correlation between water potential, which was estimated through ANN and Ψ_{stem} measured by in situ measurements. Another model was a pattern recognition ANN model for irrigation scheduling with Ψ_{stem} as the input, providing severe, moderate and no water stress as outputs. They measured Ψ_{stem} using two Scholander pressure bomb techniques for ground truth data on ninety vine plots, which was further applied to other twenty-three plots, which revealed high correlation values between the Ψ_{stem} modelled with ANN and observed Ψ_{stem}. The use of plant water stress characterized by water potential to schedule irrigation in vineyards, nut trees and almond trees was investigated by Poblete et al. (2017). They built an artificial neural network model to predict the spatial variability in Ψ_{stem} in a drip-irrigated Carmenere vineyard in Talca, Maule region, Chile. They worked on UAV multispectral imagery and fed bands as inputs to ANN. The stem water potential measured using a pressure chamber was used to validate the results. The coefficient of determination between ANN outputs and ground truth measurements of Ψ_{stem} was obtained in the range of 0.56 to 0.87. They found the best performance for the bands 550, 570, 670, 700 and 800 nm. Their results showed that the Ψ_{stem} estimated using the ANN model had a mean absolute error (MAE) of 0.1 MPa, root mean square error (RMSE) of 0.12 MPa and relative error (RE) of −9.1%, drawing the conclusion that ANN performed well to estimate Ψ_{stem}. Another water stress indicator based on plant response, the relative water content (RWC), was also predicted under the water deficit stress status of rice genotypes by Krishna et al. (2019) through spectral indices, multivariate techniques and neural network techniques. Krishna et al. (2019) utilized existing water band indices and proposed new water band indices, namely ratio index (RI) and normalized difference ratio index (NDRI) for the prediction of the RWC. From Fig. 6, it is observed that ANN is heavily utilized for the determination of water stress and in other areas of agriculture as well. ML is also increasingly used to estimate hydrological and renewable energy variables. To gauge reference ET and evaporation, a number of studies have recommended that machine learning

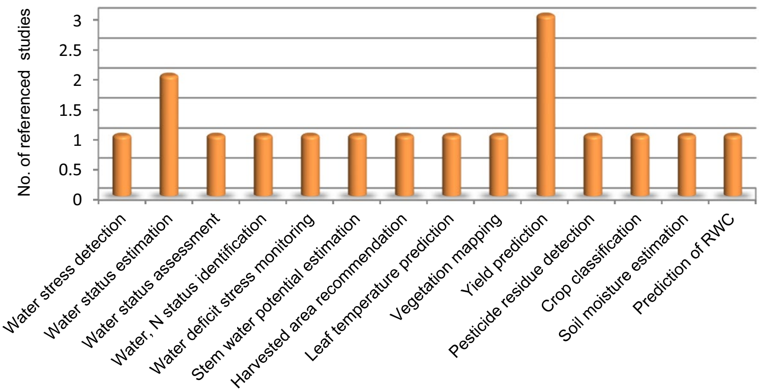


Fig. 6 Use of ANN in different studies

methods can give preferable estimates over experimental conditions depending on various driving meteorological variables.

However, much consideration has been paid to the estimation of ET in earthbound biological systems utilizing machine learning modelling approaches, alluding mostly to ANN and SVM (Dou and Yang 2018). ANN and SVM were developed to simulate and predict daily ET by Dou and Yang (2018) with the extreme learning machine (ELM) and adaptive neuro-fuzzy inference system (ANFIS) algorithms. These are two state-of-the-art machine learning algorithms that have been extensively used in hydrological time series modelling and forecasting (Gocic et al. 2016; Alizadeh et al. 2017). Dou and Yang (2018) investigated the feasibility and effectiveness of using ELM and ANFIS to model and estimate daily ET with flux tower observations in different types of ecosystems. They found that these approaches provided a novel perspective for scaling up ET from the ecosystem to a regional or global scale with remote sensing data.

Exceptionally constrained research was conducted on improving the CWSI with the utilization of ML algorithms. For example, a 1-km resolution monthly mean T_a dataset over the Tibetan Plateau was developed by Xu et al. (2018) using remote sensing, ML and auxiliary data, as they faced the issue of limited T_a observations due to an uneven distribution of stations and low density. Eleven environmental variables were extracted from MODIS, topographic index data and shuttle radar topography mission (SRTM) digital elevation model (DEM) data. Using these variables, an optimal model was built for T_a estimation with the contribution of ten ML algorithms, namely, Bayesian regularized neural network (BRNN), SVM with radial basis function (RBF) kernel, least absolute shrinkage and selection operator (LASSO), ridge regression, generalized linear model (GLM), multivariate adaptive regression splines (MARS), conditional inference tree (CIT), RF, eXtreme gradient boosting and cubist, among which the cubist algorithm was found to be the best model with the lowest precision error. This was the first attempt to develop a spatio-temporally resolved monthly T_a dataset over the region using RS and ML. Although this dataset is useful for climate change and environmental studies, T_a estimation by ML methods is helpful in improving CWSI calculations. Canopy temperature is another parameter in CWSI calculations that relies upon environmental conditions and plant reactions. A study by Andrade et al. (2018) endeavoured to forecast canopy temperature acquired by a remote system of IRTs mounted on three-range variable rate irrigation centre pivot systems for irrigated corn

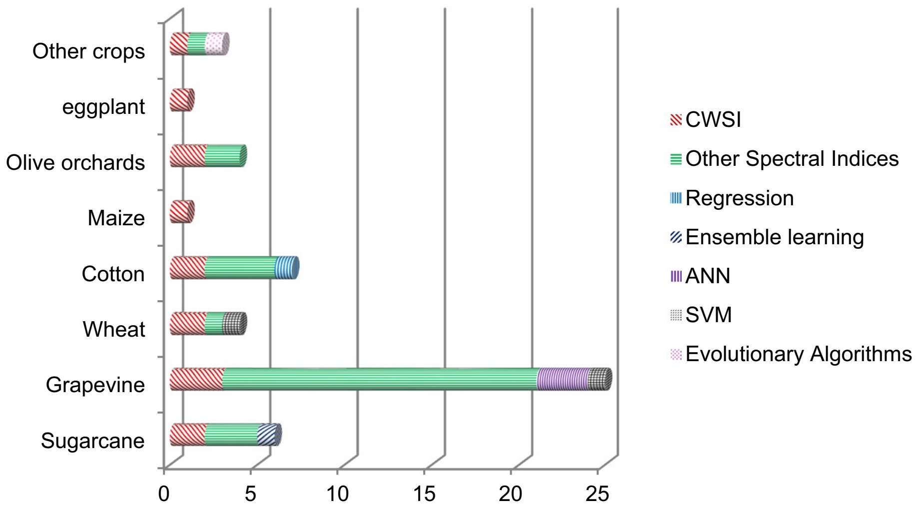


Fig. 7 Summary of the work done for crop water stress detection in different crops

crops. This system of IRTs is an irrigation scheduling supervisory control and data acquisition system (ISSCADAS) that gathered information from climate detecting frameworks, soil and plants and provided it to computerized irrigation scheduling algorithms, which dealt with the generation of site-specific plant water stress prescription maps. The expansion of ML capabilities in the ISSCADAS would help clients when poor perceivability conditions prevent the accurate estimation of canopy temperatures. Other parameters of CWSI, including the upper baseline and lower baseline, are crop-specific and vary with crop cultivation regions. Regions with dry climate have different upper and lower stress and non-stress temperature thresholds than those in normal condition regions. Only a few research studies have been carried out to date that have investigated the role of state-of-the-art machine learning techniques to estimate the upper and lower stress and well-watered threshold of temperature required by CWSI calculations. Therefore, there is much scope for ML algorithms in estimating these parameters. There is also an interesting direction towards improving these parameters using machine learning that can also account for plant response and environmental conditions.

Figure 7 presents a summary of the work done for the determination of crop water stress in different crops using different RS and ML methods.

Future directions

The future of farming depends largely on the adoption of cutting-edge technology such as ML, RS, geographical information system (GIS), UAV and cloud computing capabilities. However, these technologies are yet to make a dent in the agriculture sector in India. Fast degrading land, water resources and climate change effects make it necessary to use modern technologies to overcome these problems and achieve rational and efficient water use during crop cultivation. These technologies include real-time data analysis and real-time detection of plant water stress using advanced techniques. There is a strong correlation between plant water stress and water productivity (WP), which provides an opportunity to study the causes of the differences in water use to produce a unit of a specific crop using ML to pin-point areas where these differences occur and strategize approaches for increasing water productivity. Several studies have indicated that it will be highly significant to address plant water stress using machine learning, which will help farmers improve water and cropland management practices in the low WP areas, which will substantially enhance the food security of the expanding population without having to increase (a) crop sowing areas and (b) irrigation water allocations. Another important limitation is the high cost of different cognitive solutions available in the farming market. The solutions need to become more affordable to ensure that technology reaches masses. An open-source platform would make the solutions more affordable, resulting in rapid adoption and increased understanding among farmers.

Another problem in implementing ML algorithms is the requirement of high computational power. Advances in ML algorithms that reduce computational time for processing the data will significantly improve the use of ML in remote sensing.

Conclusions

Conventional techniques such as soil moisture measurement techniques have limitations in terms of sensor costs, their installations and trouble in acquiring estimations, particularly for heterogeneous crops and soil. These techniques provide point information and therefore inaccurately represent large fields. Plant-based estimations are reliable and more accurate but are not sophisticated and require a great deal of time. As seen in the literature, critical connections exist between remotely sensed features such as PRI and NDVI with LWP, stomatal conductance, crop coefficient and stem water potential. However, this kind of accuracy is insufficient to permit the utilization of single parameter measurements for the estimation of plant water status. Many researchers have investigated the capability of using the CWSI for different crops. EO-based CWSI was suggested to be the best indicator of water stress in agricultural crops, in contrast with other VIs and WIs at local and regional scales. Furthermore, studies discovered an infrared thermometer that could be used to estimate canopy temperature, which was reasonable to identify crop water stress and has moved towards becoming a benchmark technique for ground truth information. Among many other in situ measurements, such as LWP and canopy temperature, midday stem water potential is the most utilized technique to validate the outcomes acquired from remotely detecting systems.

The remarkable results of ML on the agricultural sector enhance the existing RS techniques, especially when ML is combined with RS data. A powerful ML technique, ANN, provides an effective tool to mine UAV multispectral data and assesses the contribution of each feature to the target. Non-influencing indices are adjusted by the weights of the ANN. Two other ML classifiers, SVM and RF, were shown to be powerful for the classification and prediction of RS data; however, they were not explored to their fullest potential in crop water stress determination using remote sensing data. Original RF classifiers were improved with oblique and rotation RF classification. Oblique RF method worked well on different datasets with discrete factorial features. The first application of oblique RF in remote sensing was implemented for multiclass land cover and land use mapping using World View 2 images. Oblique RF was also employed for the classification of symptomatic stress in *Pinus radiata* seedlings. This variant of RF could be evaluated for crop water determination. This assertion needs to be further investigated. Another novel approach to improve the RF classifier for remote sensing is the rotation RF that concatenates different rotation feature spaces into a higher space at the training stage. Studies have reported the superior performance of rotation RF in classification over RF, SVM and k-NN techniques. To date, rotation RF has not been evaluated for water stress determination using remote sensing data.

Machine learning has the capability of organizing data from systematic ground observations, ground sensors, meteorological and remote sensing (satellites, airborne, drone) sources. The availability of these data and other related data is paving the way for the deployment of ML in agriculture. To date, ML techniques have been used for identification, yield prediction and crop condition determination, but crop water stress assessments are essential for irrigation management and therefore require attention from the research community.

Acknowledgements The authors of the present study are grateful to the reviewers for providing their comments and suggestions that greatly improved this article.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflicts of interest.

References

Abdel-Rahman, E. M., & Ahmed, F. B. (2008). The application of remote sensing techniques to sugarcane (*Saccharum* spp. hybrid) production: A review of the literature. *International Journal of Remote Sensing, 29*(13), 3753–3767.

Abdel-Rahman, E. M., Mutanga, O., Adam, E., & Ismail, R. (2014). Detecting Sirex noctilio grey-attacked and lightning-struck pine trees using airborne hyperspectral data, random forest and support vector machines classifiers. *ISPRS Journal of Photogrammetry and Remote Sensing, 88*, 48–59.

Adam, E., Deng, H., Odindi, J., Abdel-Rahman, E. M., & Mutanga, O. (2017). Detecting the early stage of phaeosphaeria leaf spot infestations in maize crop using in situ hyperspectral data and guided regularized random forest algorithm. *Journal of Spectroscopy*. https://doi.org/10.1155/2017/6961387.

Alizadeh, M. J., Kavianpour, M. R., Kisi, O., & Nourani, V. (2017). A new approach for simulating and forecasting the rainfall-runoff process within the next two months. *Journal of Hydrology, 548*, 588–597.

Allen, R. G., Pereira, L. S., Raes, D., & Smith, M. (1998). Crop evapotranspiration—Guidelines for computing crop water requirements—FAO Irrigation and drainage paper 56. *FAO, Rome, 300*(9), D05109.

Allen, R. G., Tasumi, M., & Trezza, R. (2007). Satellite-based energy balance for mapping evapotranspiration with internalized calibration (METRIC)—Model. *Journal of Irrigation and Drainage Engineering, 133*(4), 380–394.

Alonso, M. C., Malpica, J. A., & de Agirre, A. M. (2011). Consequences of the Hughes phenomenon on some classification techniques. In *Proceedings of the American Society for Photogrammetry and Remote Sensing (ASPRS) 2011 annual conference* (pp. 1–5). Milwaukee, Wisconsin: ASPRS.

AlSuwaidi, A., Grieve, B., & Yin, H. (2018). Feature-ensemble-based novelty detection for analyzing plant hyperspectral datasets. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 11*, 1041–1055.

Andrade, M. A., Evett, S. R., & O'Shaughnessy, S. A. (2018). Machine learning algorithms applied to the forecasting of crop water stress indicators. In *Proceeding in technical irrigation show*. California, USA: Irrigation Association.

Bai, Y., Wong, M. S., Shi, W.-Z., Wu, L.-X., & Qin, K. (2015). Advancing of land surface temperature retrieval using extreme learning machine and spatio-temporal adaptive data fusion algorithm. *Remote Sensing, 7*(4), 4424–4441.

Bajwa, S. G., & Vories, E. D. (2006). Spectral response of cotton canopy to water stress. Paper number 061064. St Joseph, MI, USA: ASABE.

Baluja, J., Diago, M. P., Balda, P., Zorer, R., Meggio, F., Morales, F., et al. (2012). Assessment of vineyard water status variability by thermal and multispectral imagery using an unmanned aerial vehicle (UAV). *Irrigation Science, 30*(6), 511–522.

Barnes, E. M., Clarke, T. R., Richards, S. E., Colaizzi, P. D., Haberland, J., Kostrzewski, M., et al. (2000). Coincident detection of crop water stress, nitrogen status and canopy density using ground based multispectral data. In *Proceedings of the fifth international conference on precision agriculture* (Vol. 1619). Madison, WI, USA: ASA/SSSA/CSSA.

Bastiaanssen, W. G. M., Menenti, M., Feddes, R. A., & Holtslag, A. A. M. (1998). A remote sensing surface energy balance algorithm for land (SEBAL). 1. Formulation. *Journal of Hydrology, 212*, 198–212.

Belgiu, M., & Druaguct, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing, 114*, 24–31.

Berni, J. A. J., Zarco-Tejada, P. J., Sepulcre-Cantó, G., Fereres, E., & Villalobos, F. (2009a). Mapping canopy conductance and CWSI in olive orchards using high resolution thermal remote sensing imagery. *Remote Sensing of Environment, 113*(11), 2380–2388.

Berni, J. A. J., Zarco-Tejada, P. J., Suárez Barranco, M. D., & Fereres Castiel, E. (2009b). Thermal and narrow-band multispectral remote sensing for vegetation monitoring from an unmanned aerial vehicle. *IEEE Transactions on Geoscience and Remote Sensing, 47*(3), 722–738.

Birth, G. S., & McVey, G. R. (1968). Measuring the color of growing turf with a reflectance spectrophotometer 1. *Agronomy Journal, 60*(6), 640–643.

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32.

Brunini, R. G., & Turco, J. E. P. (2016). Water stress indices for the sugarcane crop on different irrigated surfaces. *Revista Brasileira de Engenharia Agricola e Ambiental, 20*(10), 925–929.

Carpenter, G. A., Gopal, S., Macomber, S., Martens, S., Woodcock, C. E., & Franklin, J. (1999). A neural network method for efficient vegetation mapping. *Remote Sensing of Environment, 70*(3), 326–338.

Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 785–794). New York, NY, United States: Association for Computing Machinery.

Choudhury, B. L. (1989). Estimating evaporation and carbon assimilation using infrared temperature data vistas in modeling . In *Theory and applications of optical remote sensing* (pp. 628–690). New York, United States: Wiley.

Cohen, Y., Alchanatis, V., Meron, M., Saranga, Y., & Tsipris, J. (2005). Estimation of leaf water potential by thermal imagery and spatial analysis. *Journal of Experimental Botany, 56*(417), 1843–1852.

Colak, Y. B., Yazar, A., Colak, I., Akca, H., & Duraktekin, G. (2015). Evaluation of crop water stress index (CWSI) for eggplant under varying irrigation regimes using surface and subsurface drip systems. *Agriculture and Agricultural Science Procedia, 4*, 372–382.

Cremona, M. V., Stützel, H., & Kage, H. (2004). Irrigation scheduling of kohlrabi (*Brassica oleracea* var. gongylodes) using crop water stress index. *HortScience, 39*(2), 276–279.

DeTar, W. R., Penner, J. V., & Funk, H. A. (2006). Airborne remote sensing to detect plant water stress in full canopy cotton. *Transactions of the ASABE, 49*(3), 655–665.

Dou, X., & Yang, Y. (2018). Evapotranspiration estimation using four different machine learning approaches in different terrestrial ecosystems. *Computers and Electronics in Agriculture, 148*, 95–106.

Enciso, J., Porter, D., Peries, X., et al. (2007). Irrigation monitoring with soil water sensors (Spanish). Fact sheet B-6194. College Station, Texas, USA: Texas AgriLife Extension Service, Texas A&M System.

Erdem, Y., Sehirali, S., Erdem, T., & Kenar, D. (2006). Determination of crop water stress index for irrigation scheduling of bean (*Phaseolus vulgaris* L.). *Turkish Journal of Agriculture and Forestry, 30*(3), 195–202.

Fernandez, J. (2017). Plant-based methods for irrigation scheduling of woody crops. *Horticulturae, 3*(2), 35.

Foody, G. M., & Mathur, A. (2004a). A relative evaluation of multiclass image classification by support vector machines. *IEEE Transactions on Geoscience and Remote Sensing, 42*(6), 1335–1343.

Foody, G. M., & Mathur, A. (2004b). Toward intelligent training of supervised image classifications: Directing training data acquisition for SVM classification. *Remote Sensing of Environment, 93*(1–2), 107–117.

Gamon, J. A., & Surfus, J. S. (1999). Assessing leaf pigment content and activity with a reflectometer. *New Phytologist, 143*(1), 105–117.

Gamon, J. A., Penuelas, J., & Field, C. B. (1992). A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency. *Remote Sensing of Environment, 41*(1), 35–44.

Gao, B.-C. (1996). NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment, 58*(3), 257–266.

Ghoggali, N., Melgani, F., & Bazi, Y. (2009). A multiobjective genetic SVM approach for classification problems with limited training samples. *IEEE Transactions on Geoscience and Remote Sensing, 47*(6), 1707–1718.

Gislason, P. O., Benediktsson, J. A., & Sveinsson, J. R. (2006). Random forests for land cover classification. *Pattern Recognition Letters, 27*(4), 294–300.

Glenn, E. P., Nagler, P. L., & Huete, A. R. (2010). Vegetation index methods for estimating evapotranspiration by remote sensing. *Surveys in Geophysics, 31*(6), 531–555.

Gocic, M., Petković, D., Shamshirband, S., & Kamsin, A. (2016). Comparative analysis of reference evapotranspiration equations modelling by extreme learning machine. *Computers and Electronics in Agriculture, 127*, 56–63.

Goel, P. K., Prasher, S. O., Patel, R. M., Landry, J.-A., Bonnell, R. B., & Viau, A. A. (2003). Classification of hyperspectral data by decision trees and artificial neural networks to identify weed stress and nitrogen status of corn. *Computers and Electronics in Agriculture, 39*(2), 67–93.

Gonzalez-Dugo, M. P., Moran, M. S., Mateos, L., & Bryant, R. (2006). Canopy temperature variability as an indicator of crop water stress severity. *Irrigation Science, 24*(4), 233.

Haboudane, D., Miller, J. R., Tremblay, N., Zarco-Tejada, P. J., & Dextraze, L. (2002). Integrated narrowband vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote Sensing of Environment, 81*(2–3), 416–426.

Hassan-Esfahani, L., Torres-Rua, A., & McKee, M. (2015). Assessment of optimal irrigation water allocation for pressurized irrigation system using water balance approach, learning machines, and remotely sensed data. *Agricultural Water Management, 153*, 42–50.

Hsu, C.-W., & Lin, C.-J. (2002). A comparison of methods for multiclass support vector machines. *IEEE Transactions on Neural Networks, 13*(2), 415–425.

Huang, H., Gong, P., Clinton, N., & Hui, F. (2008). Reduction of atmospheric and topographic effect on Landsat TM data for forest classification. *International Journal of Remote Sensing, 29*(19), 5623–5642.

Idso, S. B., Jackson, R. D., Pinter, P. J., Jr., Reginato, R. J., & Hatfield, J. L. (1981). Normalizing the stress-degree-day parameter for environmental variability. *Agricultural Meteorology, 24*, 45–55.

Ihuoma, S. O., & Madramootoo, C. A. (2017). Recent advances in crop water stress detection. *Computers and Electronics in Agriculture, 141*, 267–275.

Jackson, R. D., Reginato, R. J., & Idso, S. B. (1977). Wheat canopy temperature: A practical tool for evaluating water requirements. *Water Resources Research, 13*(3), 651–656.

Jackson, R. D., Idso, S. B., Reginato, R. J., & Pinter, P. J., Jr. (1981). Canopy temperature as a crop water stress indicator. *Water Resources Research, 17*(4), 1133–1138.

Jiang, D., Yang, X., Clinton, N., & Wang, N. (2004). An artificial neural network model for estimating crop yields using remotely sensed information. *International Journal of Remote Sensing, 25*(9), 1723–1732.

Jones, H. G. (2013). *Plants and microclimate: A quantitative approach to environmental plant physiology*. Cambridge, United Kingdom: Cambridge University Press.

Jordan, C. F. (1969). Derivation of leaf-area index from quality of light on the forest floor. *Ecology, 50*(4), 663–666.

Kacira, M., Ling, P. P., & Short, T. H. (2002). Establishing Crop Water Stress Index (CWSI) threshold values for early, non-contact detection of plant water stress. *Transactions of the ASAE, 45*(3), 775.

Kaheil, Y. H., Rosero, E., Gill, M. K., McKee, M., & Bastidas, L. A. (2008). Downscaling and forecasting of evapotranspiration using a synthetic model of wavelets and support vector machines. *IEEE Transactions on Geoscience and Remote Sensing, 46*(9), 2692–2707.

Katsoulas, N., Elvanidi, A., Ferentinos, K. P., Kacira, M., Bartzanas, T., & Kittas, C. (2016). Crop reflectance monitoring as a tool for water stress detection in greenhouses: A review. *Biosystems Engineering, 151*, 374–398.

Khairunniza-Bejo, S., Mustaffha, S., & Ismail, W. I. W. (2014). Application of artificial neural network in predicting crop yield: A review. *Journal of Food Science and Engineering, 4*(1), 1.

Khobragade, A., Athawale, P., & Raguwanshi, M. (2015). Optimization of statistical learning algorithm for crop discrimination using remote sensing data. In *Advance computing conference (IACC), 2015 IEEE international* (pp. 570–574).

King, B. A., & Shellie, K. C. (2016). Evaluation of neural network modeling to predict non-water-stressed leaf temperature in wine grape for calculation of crop water stress index. *Agricultural Water Management, 167*, 38–52.

Krishna, G., Sahoo, R. N., Singh, P., Bajpai, V., Patra, H., Kumar, S., et al. (2019). Comparison of various modelling approaches for water deficit stress monitoring in rice crop through hyperspectral remote sensing. *Agricultural Water Management, 213*, 231–244.

Li, Z.-L., Tang, B. H., Wu, H., Ren, H., Yan, G., Wan, Z., et al. (2013). Satellite-derived land surface temperature: Current status and perspectives. *Remote Sensing of Environment, 131*, 14–37.
Liaw, A., Wiener, M., et al. (2002). Classification and regression by randomForest. *R News, 2*(3), 18–22.
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.
Melgani, F., & Bruzzone, L. (2004). Classification of hyperspectral remote sensing images with support vector machines. *IEEE Transactions on Geoscience and Remote Sensing, 42*(8), 1778–1790.
Moller, M., Alchanatis, V., Cohen, Y., Meron, M., Tsipris, J., Naor, A., et al. (2006). Use of thermal and visible imagery for estimating crop water status of irrigated grapevine. *Journal of Experimental Botany, 58*(4), 827–838.
Moshou, D., Pantazi, X.-E., Kateris, D., & Gravalos, I. (2014). Water stress detection based on optical multisensor fusion with a least squares support vector machine classifier. *Biosystems Engineering, 117*, 15–22.
Mountrakis, G., Im, J., & Ogole, C. (2011). Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing, 66*(3), 247–259.
Mulla, D. J. (2013). Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. *Biosystems Engineering, 114*(4), 358–371.
Mulyono, S., et al. (2016). Identifying sugarcane plantation using LANDSAT-8 images with support vector machines. In *Institute of Physics (IOP) conference series: Earth and environmental science*. IOP Publishing, 47, 012008.
O'Shaughnessy, S. A., & Evett, S. R. (2010). Canopy temperature based system effectively schedules and controls center pivot irrigation of cotton. *Agricultural Water Management, 97*(9), 1310–1316.
Osroosh, Y., Peters, R. T., Campbell, C. S., & Zhang, Q. (2015). Automatic irrigation scheduling of apple trees using theoretical crop water stress index with an innovative dynamic threshold. *Computers and Electronics in Agriculture, 118*, 193–203.
Pal, M., & Mather, P. M. (2005). Support vector machines for classification in remote sensing. *International Journal of Remote Sensing, 26*(5), 1007–1011.
Paltineanu, C., Chitu, E., & Tanasescu, N. (2009). Correlation between the crop water stress index and soil moisture content for apple in a loamy soil: A case study in southern Romania. In *VI International symposium on irrigation of horticultural crops 889* (pp. 257–264). International Society for Horticultural Science (ISHS) Acta Horticulturae 889.
Pedergnana, M., Marpu, P. R., Dalla Mura, M., Benediktsson, J. A., & Bruzzone, L. (2013). A novel technique for optimal feature selection in attribute profiles based on genetic algorithms. *IEEE Transactions on Geoscience and Remote Sensing, 51*(6), 3514–3528.
Penman, H. L. (1948). Natural evaporation from open water, bare soil and grass. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, *193*(1032), 120–145.
Poblete, T., Ortega-Farias, S., Moreno, M. A., & 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.
Poccas, I., Gonccalves, J., Costa, P. M., Gonccalves, I., Pereira, L. S., & Cunha, M. (2017). Hyperspectral-based predictive modelling of grapevine water status in the portuguese douro wine region. *International Journal of Applied Earth Observation and Geoinformation, 58*, 177–190.
Poona, N., Van Niekerk, A., & Ismail, R. (2016). Investigating the utility of oblique tree-based ensembles for the classification of hyperspectral data. *Sensors, 16*(11), 1918.
Qi, J., Chehbouni, A., Huete, A. R., Kerr, Y. H., & Sorooshian, S. (1994). A modified soil adjusted vegetation index. *Remote Sensing of Environment, 48*(2), 119–126.
R Development-Core-Team. (2005). *A language and environment for statistical computing*. ISBN 3-900051-07-O. Vienna, Austria: R foundation for Statistical Computing 2013. https://www.r-project.org/.
Rahman, M. R., Islam, A., & Rahman, M. A. (2004). NDVI derived sugarcane area identification and crop condition assessment. *Plan Plus, 1*(2), 1–12.
Rallo, G., Minacapilli, M., Ciraolo, G., & Provenzano, G. (2014). Detecting crop water status in mature olive groves using vegetation spectral measurements. *Biosystems Engineering, 128*, 52–68.
Ramirez, D. A., Yactayo, W., Rens, L. R., Rolando, J. L., Palacios, S., De Mendiburu, F., et al. (2016). Defining biological thresholds associated to plant water status for monitoring water restriction effects: Stomatal conductance and photosynthesis recovery as key indicators in potato. *Agricultural Water Management, 177*, 369–378.
Rapaport, T., Hochberg, U., Shoshany, M., Karnieli, A., & Rachmilevitch, S. (2015). Combining leaf physiology, hyperspectral imaging and partial least squares-regression (PLS-R) for grapevine water status assessment. *ISPRS Journal of Photogrammetry and Remote Sensing, 109*, 88–97.

Rodriguez, J. J., Kuncheva, L. I., & Alonso, C. J. (2006). Rotation forest: A new classifier ensemble method. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 28*(10), 1619–1630.

Romano, G., Zia, S., Spreer, W., Sanchez, C., Cairns, J., Araus, J. L., et al. (2011). Use of thermography for high throughput phenotyping of tropical maize adaptation in water stress. *Computers and Electronics in Agriculture, 79*(1), 67–74.

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. *Computers and Electronics in Agriculture, 147*, 109–117.

Rondeaux, G., Steven, M., & Baret, F. (1996). Optimization of soil-adjusted vegetation indices. *Remote Sensing of Environment, 55*(2), 95–107.

Roujean, J.-L., & Breon, F.-M. (1995). Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. *Remote Sensing of Environment, 51*(3), 375–384.

Rouse, J. W., Jr., Haas, R. H., Schell, J. A., & Deering, D. W. (1974). Monitoring vegetation systems in the Great Plains with ERTS. *NASA Special Publication, 351*, 309.

Rozenstein, O., Haymann, N., Kaplan, G., & Tanny, J. (2018). Estimating cotton water consumption using a time series of Sentinel-2 imagery. *Agricultural Water Management, 207*, 44–52.

Rud, R., Cohen, Y., Alchanatis, V., Levi, A., Brikman, R., Shenderey, C., et al. (2014). Crop water stress index derived from multi-year ground and aerial thermal images as an indicator of potato water status. *Precision Agriculture, 15*(3), 273–289.

Saini, R., & Ghosh, S. K. (2018). Crop classification on single date sentinel-2 imagery using random forest and support vector machine. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences, XLII-5*, 683–688.

Samborska, I. A., Alexandrov, V., Sieczko, L., & Kornatowska, B. (2014). Artificial neural networks and their application in biological and agricultural research. *Journal of NanoPhotoBioSciences, 2*, 14–30.

Serrano, L., Gonzalez-Flor, C., & Gorchs, G. (2010). Assessing vineyard water status using the reflectance based water index. *Agriculture, Ecosystems & Environment, 139*(4), 490–499.

Sharma, P. K., Kumar, D., Srivastava, H. S., & Patel, P. (2018). Assessment of different methods for soil moisture estimation: A review. *Journal of Remote Sensing and GIS, 9*(1), 57–73.

Sun, G., Xie, H., & Sinnott, R. O. (2017). A crop water stress monitoring system utilising a hybrid e-infrastructure. In *Proceedings of the10th international conference on utility and cloud computing* (pp. 161–170). New York, United States: Association for Computing Machinery.

Taghvaeian, S., Chávez, J. L., & Hansen, N. C. (2012). Infrared thermometry to estimate crop water stress index and water use of irrigated maize in Northeastern Colorado. *Remote Sensing, 4*(11), 3619–3637.

Tanriverdi, C., Degirmenci, H., Gonen, E., & Boyaci, S. (2016). A comparison of the gravimetric and TDR methods in terms of determining the soil water content of the corn plant. *Scientific Papers-Series A-Agronomy, 59*, 153–158.

Tilling, A. K., O'Leary, G. J., Ferwerda, J. G., Jones, S. D., Fitzgerald, G. J., Rodriguez, D., et al. (2007). Remote sensing of nitrogen and water stress in wheat. *Field Crops Research, 104*(1–3), 77–85.

Turner, N. C. (1988). Measurement of plant water status by the pressure chamber technique. *Irrigation Science, 9*(4), 289–308.

Vapnik, V. N. (1999). An overview of statistical learning theory. *IEEE Transactions on Neural Networks, 10*(5), 988–999.

Verstraeten, W., Veroustraete, F., & Feyen, J. (2008). Assessment of evapotranspiration and soil moisture content across different scales of observation. *Sensors, 8*(1), 70–117.

Veysi, S., Naseri, A. A., Hamzeh, S., & Bartholomeus, H. (2017). A satellite based crop water stress index for irrigation scheduling in sugarcane fields. *Agricultural Water Management, 189*, 70–86.

Warner, T. A., & Nerry, F. (2009). Does single broadband or multispectral thermal data add information for classification of visible, near-and shortwave infrared imagery of urban areas? *International Journal of Remote Sensing, 30*(9), 2155–2171.

Weng, Q., Fu, P., & Gao, F. (2014). Generating daily land surface temperature at Landsat resolution by fusing Landsat and MODIS data. *Remote Sensing of Environment, 145*, 55–67.

Xu, Y., Knudby, A., Shen, Y., & Liu, Y. (2018). Mapping monthly air temperature in the Tibetan Plateau from MODIS data based on machine learning methods. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 11*(2), 345–354.

Xue, J., & Su, B. (2017). Significant remote sensing vegetation indices: A review of developments and applications. *Journal of Sensors, 2017*, 1–17.

Yang, C., Everitt, J. H., & Murden, D. (2011). Evaluating high resolution SPOT 5 satellite imagery for crop identification. *Computers and Electronics in Agriculture, 75*(2), 347–354.

Yuan, G., Luo, Y., Sun, X., & Tang, D. (2004). Evaluation of a crop water stress index for detecting water stress in winter wheat in the North China Plain. *Agricultural Water Management, 64*(1), 29–40.

Zarco-Tejada, P. J., Rueda, C. A., & Ustin, S. L. (2003). Water content estimation in vegetation with MODIS reflectance data and model inversion methods. *Remote Sensing of Environment, 85*(1), 109–124.

Zarco-Tejada, P. J., Berjon, A., Lopez-Lozano, R., Miller, J. R., Martin, P., Cachorro, V., et al. (2005). Assessing vineyard condition with hyperspectral indices: Leaf and canopy reflectance simulation in a row-structured discontinuous canopy. *Remote Sensing of Environment, 99*(3), 271–287.

Zarco-Tejada, P. J., Gonzalez-Dugo, V., Williams, L. E., Suarez, L., Berni, J. A. J., Goldhamer, D., et al. (2013). A PRI-based water stress index combining structural and chlorophyll effects: Assessment using diurnal narrow-band airborne imagery and the CWSI thermal index. *Remote Sensing of Environment, 138*, 38–50.

Zhang, L., & Lemeur, R. (1995). Evaluation of daily evapotranspiration estimates from instantaneous measurements. *Agricultural and Forest Meteorology, 74*(1–2), 139–154.

Zhang, K., Kimball, J. S., & Running, S. W. (2016). A review of remote sensing based actual evapotranspiration estimation. *Wiley Interdisciplinary Reviews: Water, 3*(6), 834–853.