Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

A Low-cost Soil Moisture Monitoring Method by Using Walabot and Machine Learning Algorithms

Haoyu Niu * **Yanan Wang** ** **Tiebiao Zhao** ***
YangQuan Chen ***

* *Department of Electrical Engineering and Computer Science, University of California, Merced, CA 95340 USA (e-mail: hniu2@ucmerced.edu).*
** *Power Electronics and Mechanical Control Laboratory, Beijing Institute of Technology, Beijing, China 100081 (e-mail:wangyn_bit@126.com).*
*** *Department of Mechanical Engineering, University of California, Merced, CA 95340 USA, (e-mail: tzhao3@ucmerced.edu; ychen53@ucmerced.edu).*

Abstract: Soil moisture plays an important role in agricultural processes, which has a significant effect on crop evapotranspiration, the exchange of water, and energy fluxes. Recently, soil moisture can be measured by remote sensing or proximate sensing techniques, such as thermal, optical, and microwave measurements. However, there are limitations to the applications of these methods, such as low spatial resolution, limited surface penetration, and vegetation. In this study, it proposed a new low-cost soil moisture monitoring method by using a Walabot sensor and machine learning algorithms. Walabot is a pocket-sized device cutting-edge technology for Radio Frequency tridimensional sensing. Unlike the remote sensing tools such as unmanned aerial vehicles (UAVs) limited by cloud cover or payload capability, the Walabot can be used flexibly in the field and provide data information more promptly and accurately than UAVs or satellite. By putting different moisture levels of soil on the Walabot, the Walabot can collect radio frequency reflectance data from different levels of soil moisture. Then, machine learning algorithms, such as principal component analysis (PCA), linear discriminant analysis (LDA), have been applied for data processing. Results showed that Walabot has a state-of-art performance in estimating soil moisture.

Keywords: Soil moisture, proximate sensing, Walabot, machine learning.

1. INTRODUCTION

In Engman (1991), Engman et al. defined soil moisture as the temporary storage of precipitation within a shallow layer of the earth. The soil moisture plays an important role in hydrological applications, such as agriculture Zhao et al. (2018), climate change Entekhabi et al. (1994), and meteorology Fast and McCorcle (1991). For example, the data analysis from soil moisture monitoring can be used for crop yield estimation, irrigation treatment inference Niu et al. (2019c), and warning of drought Engman (1991). Soil moisture monitoring can also be applied for pest management Powell et al. (2007) and evapotranspiration estimation Niu et al. (2019b). Therefore, it is important to monitor the soil moisture accurately. Typically, there are two types of methods for monitoring the soil moisture, proximate sensing and remote sensing.

Proximate sensing methods for soil moisture are currently restricted to point-specific measurements Wang and Qu (2009). For example, researchers usually put the soil moisture probes in the test field for monitoring. However, these discrete measurements can not represent the spatial and temporal soil moisture distribution for the whole field.

With the development of remote sensing technology, the satellite has been widely used for soil moisture remote sensing Engman (1990). Many researchers have proved that optical and thermal remote sensing can be used for soil moisture measurements. For example, in Wang and Qu (2007), Wang et al. proposed the normalized multiband drought index (NMDI) for remotely sensing the soil based on the soil spectral characteristic. Since variations of soil moisture have a significant influence on soil surface temperature Friedl and Davis (1994), thermal infrared remote sensing is also used for measuring the soil temperature to correlate it with soil moisture. Active and passive microwave remote sensing techniques are also commonly used for soil moisture measurements Walker (1999). For passive microwave sensors, they can measure the intensity of microwave emission from the soil, which is proportional to the brightness temperature, a product of the surface temperature and emissivity Wang and Qu (2009); Wigneron et al. (2003). However, there are disadvantages to these methods. Limited surface penetration

Copyright lies with the authors

can be a problem both for optical and thermal remote sensing. Cloud contamination can be another issue Wang and Qu (2009). The data acquired from the microwave has a low spatial resolution.

Therefore, in this study, the authors proposed a new low-cost (less than $ 1000) soil moisture monitoring method by using a Walabot sensor and machine learning algorithms. Walabot is a pocket-sized device and cutting-edge technology for Radio Frequency tridimensional sensing, which has already been used in many research topics, such as nematodes detection Niu et al. (2020), and battery voltage detection Wang et al. (2019). It can work flexibly in the field and provide data information accurately than remote sensing methods with machine learning algorithms. First, the sensor was used to collect radio frequency reflectance of sampling soil, as shown in Fig. 1, which can detect the physical structure of the soil moisture. Second, the collected data were pre-processed by data enhancement or a wavelet transform. Third, processed data was used by PCA Wold et al. (1987) and LDA Balakrishnama and Ganapathiraju (1998) for analysis. Results showed that the Walabot successfully classified the different levels of soil moisture with a state-of-art performance. Moreover, with the development of wireless technology and micro-electromechanical systems, and computer vision, we might even use the Walabot to recognize real-time soil moisture monitoring in future research.

The rest of the paper is organized as follows. In Section 2, we give a more detailed introduction for the Walabot sensor, data collection, and analysis. Results and discussions are presented in Section 3. In Section 4, concluding remarks are presented for using Walabot to detect soil moisture.

2. MATERIAL AND METHODS

2.1 Study site

This research was conducted at Mechatronics, Embedded Systems and Automation (MESA) Lab in Atwater, California, USA ($37°22'30.6''N$, $120°34'40.9''W$).

2.2 Walabot

The sensor being used is Walabot Developer (Vayyar Imaging Ltd), as shown in Fig. 2. The Walabot Developer is a programmable 3D sensor that uses radio frequency to see through the soil and creates a reflectance image within one second. The frequency range is 3.3 - 10 GHz (US/FCC model) and 6.3 - 8 GHz (EU/CE model). The average transmit power of both models is below 41dBm/MHz and do not have any health problem. In principle, the Walabot uses an antenna array to illuminate the area in front of it, and captures the returning signals. The signals are produced and recorded by VYYR2401 A3 System-on-Chip integrated circuits.

Based on the technical specs Walabot (2018), the Walabot can sense the environment by transmitting, receiving, and recording signals from multiple antennas. Multiple transmit-receive antenna pairs' recordings are analyzed to build a 3D image of the environment. Then, researchers

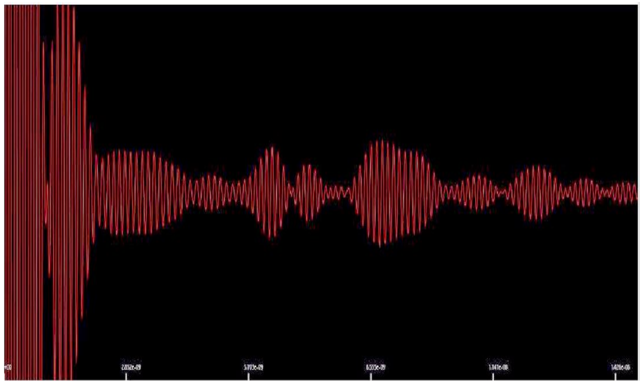

Fig. 1. Radio Frequency reflectance

can detect changes in the environment by analyzing the sequences of images. The sensor is also capable of short-range imaging into dielectric environments, such as drywall and concrete. Therefore, it can be used in many study areas as follows:

1. In-Room / Wall imaging
2. Object detection, location and tracking
3. Speed measurement and motion sensing
4. Dielectric properties of materials sensing

2.3 Experiment setup

In this study, the authors used Walabot to detect different levels of soil moisture. The experiment was conducted in the MESA Lab. The soil was sampled in an almond field near the lab and was divided into 3 cups, as shown in Fig. 2. All the soil samplings are from the same spot in the almond field to make sure they are homogeneous. The soil was dried out to make sure all the 3 cups of soil are at the same lowest moisture level. The weights of three cups of dry soil are 632 grams, 630 grams, and 634 grams. 6 g or 8 g water was added in every cup each time (10 times in total) to increase the soil moisture until the soil moisture is saturated, as shown in Table 1.

Table 1. Soil samplings

Soil condition	Soil sample 1	Soil sample 2	Soil sample 3
Dry	632g	630g	634g
1	640g	638g	642g
2	648g	646g	648g
3	654g	652g	656g
4	660g	660g	664g
5	668g	668g	670g
6	676g	674g	678g
7	682g	682g	686g
8	690g	690g	692g
9	696g	696g	700g
10 (Saturated)	704g	704g	708g

2.4 Data collection and processing

The Walabot was used to measure the soil moisture every time after the water was added. Each measurement by the Walabot was repeated ten times to reduce the likelihood of errors or anomalous results so that it could increase the confidence interval.

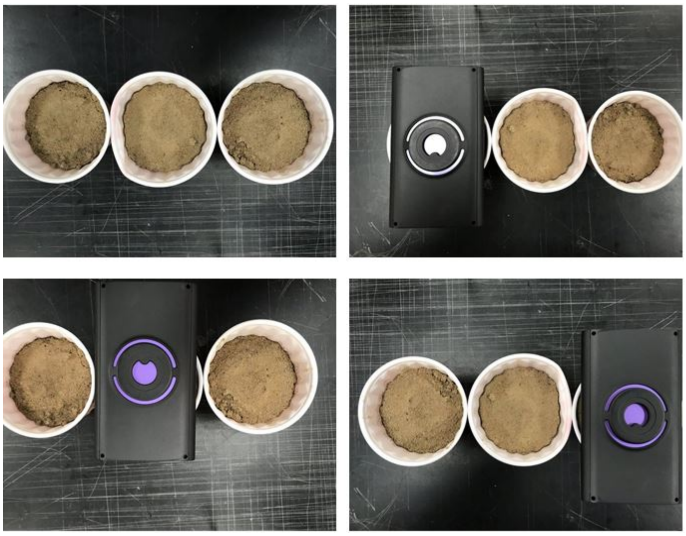

Fig. 2. Walabot data collection

For image processing, the authors used two different machine learning methods, the Principle Component Analysis (PCA) and Linear Discriminant Analysis (LDA). Both of them can reduce the dimensionality of the datasets and increase the classification accuracy.

Principle Component Analysis. Principle Component Analysis (PCA) is a fast and flexible unsupervised method for data dimensionality reduction Jolliffe (2011). It can achieve linear projection to a lower-dimensional subspace by using singular value decomposition. It is a linear transformation that rotates the axes of the data along the direction that maximizes its variance, allowing data to be projected into a lower-dimensional subspace Jolliffe (2011). These new axes, or "loadings," are found by calculating the eigenvectors W of the data's covariance matrix, where X is an M × N matrix representing M samples of size N:

$$X^{\mathrm{T}}X = W\lambda \tag{1}$$

The eigenvalue λ represents the importance of the loading in transforming the data. As the loadings are sorted in descending order, W can be truncated to r columns, which can then be used to project data along r dimensions, preserving the dimensions that contribute most to the variance of the distribution. W is often obtained with Singular Value Decomposition (SVD) instead of performing the Eigen decomposition of $X^{\mathrm{T}}X$, as it is more computationally efficient.

PCA can also maximize the variance of the projected data. Therefore, PCA is commonly used in exploratory data analysis and making predictive models.

Linear Discriminant Analysis. Linear Discriminant Analysis (LDA) is a classifier with a linear decision boundary. It is generated by using Bayes' rule to fit class conditional densities to the data. It assumes that all classes share the same covariance matrix. After that, the LDA model can be used to reduce the dimensionality of the input data by projecting it to the most discriminative directions. The model will maximize the distance between means of classes relative to some center point for all classes, while minimizing the variance, or scatter, within each category. In the following equation, C is the number of classes, N_i is the size of class i, μ is the mean of all data points, μ_i is the mean of class i, and x_j is the jth data point in class i:

$$\frac{\sum_{i=1}^{C} N_i(\mu_i - \mu)(\mu_i - \mu)^{\mathrm{T}}}{\sum_{i=1}^{C}\sum_{j=1}^{N_i}(x_j - \mu_i)(x_j - \mu_i)^{\mathrm{T}}} \tag{2}$$

The optimized solution contains eigenvectors, which are descending order of their eigenvalues, which can be used to reduce the dataset similar to PCA. Optimizing for both within and between-class scatter is essential because only maximizing distance between means could lead to scenarios where the variance is high along the axis with great mean distances, increasing the chance that there are points from different classes overlapping. Minimizing the variance ensures that data from each class are grouped tightly along the new axis, increasing separability. Then, the output dimensionality is usually less than the number of classes so that LDA is a very strong dimensionality reduction Friedman et al. (2001).

3. RESULTS AND DISCUSSION

Each radio frequency reflectance image was converted into a 2048-dimension vector for data processing. The data was distributed as 67% for training and 33% for testing. Since the dataset is small for training eleven classifiers, the authors distributed the eleven soil conditions into five different levels from dry to saturation. As shown in Fig. 3 and Fig. 4, Dry means the dry soil. WetTotal stands for the saturated soil. Soil conditions 1, 2, and 3 are included in Wet1. Wet2 contains the soil conditions 4, 5, and 6. Wet3 includes the soil conditions 7, 8, and 9.

Several classifiers in scikit-learn were used for comparison, such as "Nearest Neighbors," "Linear SVM," "RBF SVM," "Gaussian Process," "Decision Tree," "Random Forest," "Neural Net," "AdaBoost," "Naive Bayes," and "QDA". In this soil moisture monitoring problem, the accuracy of these classifiers is shown in Table 2. The best classifiers are "Nearest Neighbors," "Gaussian Process," "Decision Tree," "Random Forest," "Neural Net," and "Naive Bayes" with an accuracy of 95%. The "QDA" is with 90% accuracy. The "Linear SVM" and "AdaBoost" are worst with 55% accuracy.

Table 2. Classifiers accuracy

Classifiers	Accuracy
Nearest Neighbors	0.95
'Linear SVM'	0.40
'RBF SVM'	0.95
'Gaussian Process'	0.95
'Decision Tree'	0.95
'Random Forest'	0.95
'Neural Net'	0.95
'AdaBoost'	0.55
'Naive Bayes'	0.95
'QDA'	0.90

Scikit-learn's accuracy classification score function evaluated the performance of the classifiers. This function computes the subset accuracy, in which the labels predicted for sampling must exactly match the corresponding true labels. Estimators use this score method as the evaluation criterion for the classification problems. All scorer objects follow the convention that higher return values are better than lower return values.

3.1 LDA

Several LDA methods are used for soil moisture classification, as shown in Table 3. decision_function(X) is for predicting confidence scores of soil samples. fit(X, y[, store_covariance, tol]) is for fitting the LDA model according to the given soil images training data and parameters. fit_transform(X[, y]) is for fitting data and transform it. get_params([deep]) is used for setting parameters for the estimator. Then, predict(X) can predict the class labels for soil samples. predict_log_proba(X) and predict_proba(X) can estimate the probability. Finally, score(X, y[, sample_weight]) can return the mean accuracy on the given soil images test data and labels. set_params(**params) is for setting estimator parameters. transform(X) is for projecting data to maximize soil class separation.

Table 3. LDA Method

Methods
decision_function(X)
fit(X, y[, store_covariance, tol])
fit_transform(X[, y])
get_params([deep])
predict(X)
predict_log_proba(X)
predict_proba(X)
score(X, y[, sample_weight])
set_params(**params)
transform(X)

The performance of the LDA for soil moisture monitoring was shown in Fig. 3. There are five different soil moisture levels with different colors, Dry, Wet1, Wet2, Wet3, and WetTotal, which means the soil sampling is saturated. LDA classifiers firstly reduced the original dimension to 2 components. As seen from Fig. 3, different colors mean different soil moisture levels and the axes of the figure are dimensionless., the LDA can classify the five different levels of soil moisture in different areas of the coordinate, so that the LDA can classify the soil moisture with an accuracy of 100 %.

3.2 PCA

In PCA methods, fit(self, X[, y]) is to fit the model with the input soil images data X. fit_transform(self, X[, y]) is for fitting the model with X and apply the dimensionality reduction on X. get_covariance(self) is for computing the data covariance with the generative model. get_params(self[,deep]) is to get parameters for the estimator. get_precision(self) is for computing the data precision matrix with the generative model. Then, inverse_transform(self, X) can transform the data back to its original space. Finally, score(self, X[, y]) can return the average log-likelihood of all samples. score_samples(self, X) can return the log-likelihood of each sample. set_params(self, params) can help set the parameters of the estimator. transform(self, X) is being used for applying dimensionality reduction to soil images input.

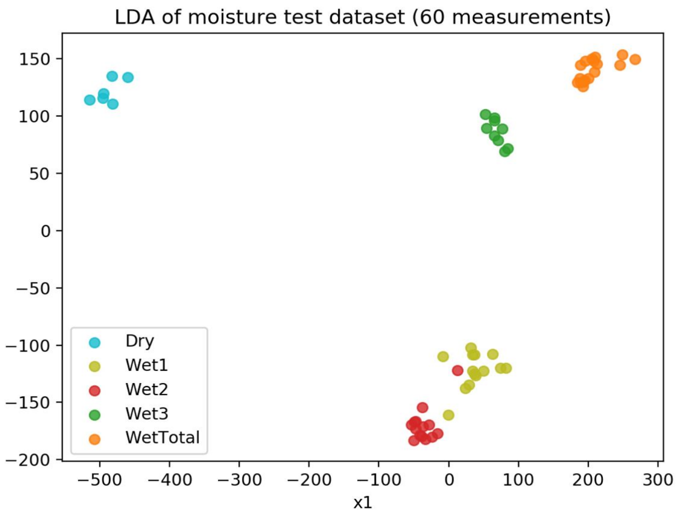

Fig. 3. LDA results for soil moisture measurement

Table 4. PCA Method

Methods
fit(self, X[, y])
fit_transform(self, X[, y])
get_covariance(self)
get_params(self[,deep])
get_precision(self)
inverse_transform(self, X)
score(self, X[, y])
score_samples(self, X)
set_params(self, params)
transform(self, X)

In Fig. 4, PCA can also classify the soil moisture successfully but not entirely. As shown in Fig. 4, WetTotal points are on the left and right sides of the image. The Wet3 and Wet2 data points did not drop in the same area. The reason might be that the PCA can not detect the features difference from the data. Similar to LDA, the PCA classifiers firstly reduced the original dimension to 2 components. Then, each classifier was tested against reduced dimensionality data with the component as 2. Results showed that LDA performs much better than the PCA method.

4. CONCLUSION

Soil moisture monitoring is essential in precision agriculture, which has a significant effect on crop evapotranspiration, the exchange of water, and energy fluxes. Soil moisture can be measured by many remote sensing or proximate sensing techniques, such as thermal, optical, and microwave measurements. However, there are limiting factors for the applications of these methods, such as low spatial resolution, limited surface penetration and vegetation. In this study, the authors used a portable sensor to classify different soil moisture successfully. By using the PCA and LDA machine learning methods, the Walabot can recognize small changes in different levels of

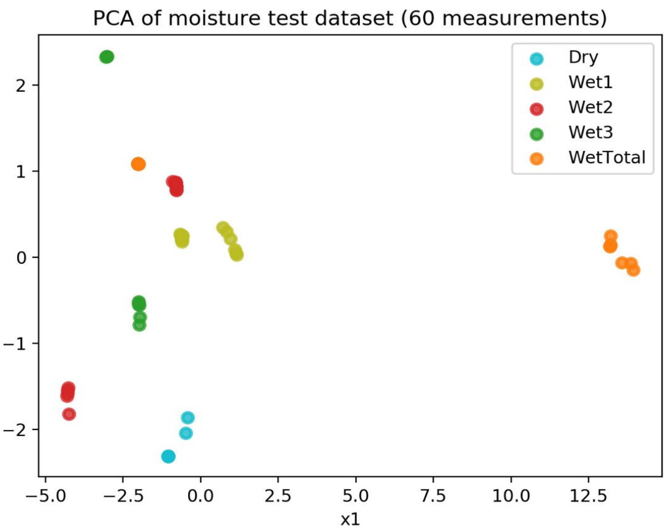

Fig. 4. PCA results for soil moisture measurement

soil moisture and can detect soil moisture difference with an accuracy of 100 % by LDA algorithms, which shows a state-of-art performance in estimating soil moisture. As a pocket-sized device cutting-edge technology for radio frequency tridimensional sensing, we believe the sensor can work flexibly in the field and provide data information more promptly and accurately than traditional remote sensing or proximate sensing method.

So far, the Walabot can only detect the difference in soil moisture. In the future, we will compare it with different soil moisture sensors to see if we can find the regression model and quantify the soil moisture measurements by using Walabot. With the development of wireless technology and microelectromechanical systems and computer vision, it might be able to be mounted on unmanned ground vehicles (UGVs) for proximate sensing Niu et al. (2019a); Tian et al. (2019); thus we can use the sensor to recognize real-time soil moisture monitoring.

ACKNOWLEDGEMENTS

Thanks go to Sebastien Sadlo for programming the Walabot detection method during the project.

REFERENCES

Balakrishnama, S. and Ganapathiraju, A. (1998). Linear discriminant analysis: A brief tutorial. *Institute for Signal and Information Processing*, 18, 1–8.

Engman, E.T. (1990). Progress in microwave remote sensing of soil moisture. *Canadian Journal of Remote Sensing*, 16(3), 6–14.

Engman, E.T. (1991). Applications of microwave remote sensing of soil moisture for water resources and agriculture. *Remote Sensing of Environment*, 35(2-3), 213–226.

Entekhabi, D., Nakamura, J., and Njoku, E. (1994). Retrieval of soil moisture profile by combined remote sensing and modeling. *JPL TRS*.

Fast, J.D. and McCorcle, M.D. (1991). The effect of heterogeneous soil moisture on a summer baroclinic circulation in the central united states. *Monthly Weather Review*, 119(9), 2140–2167.

Friedl, M. and Davis, F. (1994). Sources of variation in radiometric surface temperature over a tallgrass prairie. *Remote Sensing of Environment*, 48(1), 1–17.

Friedman, J., Hastie, T., and Tibshirani, R. (2001). *The elements of statistical learning*, volume 1. Springer Series in Statistics New York.

Jolliffe, I. (2011). *Principal component analysis*. Springer.

Niu, H., Zhao, T., and Chen, Y. (2019a). Intelligent bugs mapping and wiping (iBMW): An affordable robot-driven robot for farmers. In *2019 IEEE International Conference on Mechatronics and Automation (ICMA)*, 397–402. IEEE.

Niu, H., Zhao, T., Wang, D., and Chen, Y. (2019b). Estimating evapotranspiration with UAVs in agriculture: A review. In *2019 ASABE Annual International Meeting*. American Society of Agricultural and Biological Engineers.

Niu, H., Zhao, T., Wang, D., and Chen, Y. (2019c). A UAV resolution and waveband aware path planning for onion irrigation treatments inference. In *2019 International Conference on Unmanned Aircraft Systems (ICUAS)*, 808–812. IEEE.

Niu, H., Zhao, T., Westphal, A., and Chen, Y. (2020). A low-cost proximate sensing method for early detection of nematodes in walnut using walabot and scikit-learn classification algorithms. In *Autonomous Air and Ground Sensing Systems for Agricultural Optimization and Phenotyping V*. International Society for Optics and Photonics.

Powell, L., Berg, A.A., Johnson, D., and Warland, J. (2007). Relationships of pest grasshopper populations in Alberta, Canada to soil moisture and climate variables. *Agricultural and Forest Meteorology*, 144(1-2), 73–84.

Tian, J., Niu, H., Wang, P., and Chen, Y. (2019). Smart and autonomous farm field scouting service robot as an edge device under $1000: Challenges and opportunities. In *ASME 2019 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. American Society of Mechanical Engineers Digital Collection.

Walabot (2018). Walabot technical brief. See also URL `https://walabot.com/docs/walabot-tech-brief-416?type=pdf`.

Walker, J.P. (1999). *Estimating soil moisture profile dynamics from near-surface soil moisture measurements and standard meteorological data*. Ph.D. thesis, University of Newcastle.

Wang, L. and Qu, J.J. (2007). NMDI: A normalized multi-band drought index for monitoring soil and vegetation moisture with satellite remote sensing. *Geophysical Research Letters*, 34(20).

Wang, L. and Qu, J.J. (2009). Satellite remote sensing applications for surface soil moisture monitoring: A review. *Frontiers of Earth Science in China*, 3(2), 237–247.

Wang, Y., Niu, H., Zhao, T., Liao, X., Dong, L., and Chen, Y. (2019). Contactless Li-Ion battery voltage detection by using walabot and machine learning. In *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. American Society of Mechanical Engineers.

Wigneron, J.P., Calvet, J.C., Pellarin, T., Van de Griend, A., Berger, M., and Ferrazzoli, P. (2003). Retrieving near-surface soil moisture from microwave radiometric observations: Current status and future plans. *Remote Sensing of Environment*, 85(4), 489–506.

Wold, S., Esbensen, K., and Geladi, P. (1987). Principal component analysis. *Chemometrics and Intelligent Laboratory Systems*, 2(1-3), 37–52.

Zhao, T., Koumis, A., Niu, H., Wang, D., and Chen, Y. (2018). Onion irrigation treatment inference using a low-cost hyperspectral scanner. In *Multispectral, Hyperspectral, and Ultraspectral Remote Sensing Technology, Techniques and Applications VII*. International Society for Optics and Photonics.