

5-14-2019

Estimation of Soil Moisture at Different Soil Levels Using Machine Learning Techniques and Unmanned Aerial Vehicle (UAV) Multispectral Imagery

Mahyar Aboutalebi
Utah State University, mahyar.aboutalebi@gmail.com

L. Niel Allen
Utah State University, n.allen@usu.edu

Alfonso F. Torres-Rua
Utah State University, alfonso.torres@usu.edu

Mac McKee
Utah State University, mac.mckee@usu.edu

Calvin Coopmans
Utah State University, cal.coopmans@usu.edu
Follow this and additional works at: https://digitalcommons.usu.edu/aggieair_pubs

 Part of the [Aviation Commons](#)

Recommended Citation

Mahyar Aboutalebi, L. Niel Allen, Alfonso F. Torres-Rua, Mac McKee, and Calvin Coopmans "Estimation of soil moisture at different soil levels using machine learning techniques and unmanned aerial vehicle (UAV) multispectral imagery", Proc. SPIE 11008, Autonomous Air and Ground Sensing Systems for Agricultural Optimization and Phenotyping IV, 110080S (14 May 2019); <http://dx.doi.org/10.1117/12.2519743>

This Conference Paper is brought to you for free and open access by the AggieAir at DigitalCommons@USU. It has been accepted for inclusion in AggieAir Publications by an authorized administrator of DigitalCommons@USU. For more information, please contact digitalcommons@usu.edu.

PROCEEDINGS OF SPIE

[SPIDigitalLibrary.org/conference-proceedings-of-spie](https://spiedigitallibrary.org/conference-proceedings-of-spie)

Estimation of soil moisture at different soil levels using machine learning techniques and unmanned aerial vehicle (UAV) multispectral imagery

Mahyar Aboutalebi, L. Niel Allen, Alfonso F. Torres-Rua, Mac McKee, Calvin Coopmans

Mahyar Aboutalebi, L. Niel Allen, Alfonso F. Torres-Rua, Mac McKee, Calvin Coopmans, "Estimation of soil moisture at different soil levels using machine learning techniques and unmanned aerial vehicle (UAV) multispectral imagery," Proc. SPIE 11008, Autonomous Air and Ground Sensing Systems for Agricultural Optimization and Phenotyping IV, 110080S (14 May 2019); doi: 10.1117/12.2519743

SPIE.

Event: SPIE Defense + Commercial Sensing, 2019, Baltimore, Maryland, United States

Estimation of soil moisture at different soil levels using machine learning techniques and unmanned aerial vehicle (UAV) multispectral imagery

Mahyar Aboutalebi^{*a}, L. Niel Allen^a, Alfonso F. Torres-Rua^a, Mac McKee^a, and Calvin Coopmans^b

^aUtah Water Research Laboratory, Department of Civil and Environmental Engineering, Utah State University, 8200 Old Main Hill, Logan, UT, USA.

^bElectrical Engineering Department, Utah State University, 8200 Old Main Hill, Logan, UT, USA.

ABSTRACT

Soil moisture is a key component of water balance models. Physically, it is a nonlinear function of parameters that are not easily measured spatially, such as soil texture and soil type. Thus, several studies have been conducted on the estimation of soil moisture using remotely sensed data and data mining techniques such as artificial neural networks (ANNs) and support vector machines (SVMs). However, all models developed based on these techniques are limited to site-specific applications where they are trained and their parameters are tuned. Moreover, since the system of non-linear equations produced by and conducted in the machine learning process are not accessible to researchers, each application of these machine learning approaches must repeat these training steps for any new study area. The fact that the results of this machine learning, black box approach cannot be easily transferred to different locations for extraction of soil moisture estimates is frustrating, and it can lead to inaccurate comparisons between methods or model performance. To overcome the Black-box issue, this study employed a powerful technique called genetic programming (GP), which is a combination of an evolutionary algorithm and artificial intelligence, to simulate soil moisture at different levels using high-resolution, multispectral imagery acquired with an unmanned aerial vehicle (UAV). The output of this approach is either a linear or nonlinear empirical equation that can be used by others. The performance of GP was compared with ANN and SVM modeling results. Several sets of high-resolution aerial imagery captured by the Utah State University AggieAir UAV system over two experimental pasture sites located in northern and southern Utah were used for this soil moisture estimation approach. The inputs used to train these models include the reflectance for the visible, near-infrared (NIR), and thermal bands. The results show (1) the performance of GP versus ANN and SVM and (2) the master equation provided by GP, which can be used in other locations and applications.

Keywords: Soil moisture, Machine learning, High-resolution imagery, Support vector machine, Neural network, Genetic programming, AggieAir, UAV

1. INTRODUCTION

Soil moisture is recognized as a key state variable that affects hydrological responses by controlling infiltration, runoff, and percolation (Ali et al. 2015¹), and it has a variety of applications in drought, flood, and irrigation management (Srivastava et al. 2013²). Sometimes soil moisture is called “memory of precipitation” in meteorological processes or “memory of irrigation” in agricultural studies since it is a variable immediately affected by precipitation and irrigation. In agriculture, accurate estimation of soil water content can help with irrigation scheduling and yield prediction (Gill et al. 2006³). However, producing a soil moisture map using in-situ measurements is a challenging task due to the large spatial and temporal variability. Hence, various empirically and physically-based models have been used to simulate, predict, or downscale soil moisture using remotely-sensed data.

Further author information: (Send correspondence to Mahyar Aboutalebi)
Mahyar Aboutalebi: E-mail: Mahyar.Aboutalebi@gmail.com

Gill et al. (2006³) applied a support vector machine (SVM) and an artificial neural network (ANN) to predict ground soil moisture on the Little Washita River Experimental Watershed using soil moisture sensors and meteorological data. Results showed that SVM could simulate soil moisture data up to 89% accuracy in terms of correlation coefficient, while ANN's accuracy was up to 74%. Wu et al. (2007⁴) applied SVM and ANN to predict timeseries of soil water content in the Purple Hilly region. The analysis of timeseries generated by SVM and ANN showed that SVM predictions performed better than the ANN model. Liu et al. (2008⁵) combined SVM and ANN and presented a novel hybrid model for soil water content prediction. In this hybrid model, ANN was responsible for classifying input space into different feature clusters for SVM, which is then used as the main regression model. The proposed model showed improved performance compared to individual SVM and ANN models. Liu et al. (2010⁶) introduced a coupled data assimilation model to predict soil water content at different soil layers by merging SVM and the ensemble Kalman filter (EnKF). They found that EnKF can significantly improve the performance of SVM (up to 15% in terms of The correlation coefficient). Zaman et al. (2012⁷) used remotely sensed optical and thermal data to train SVM and a relevant vector machine (RVM) for topsoil moisture estimations. Pasolli et al. (2011⁸) presented an experimental analysis to show the capability of SVM in soil moisture content estimation using remotely-sensed data at field scale. The analysis of the results confirmed the effectiveness of SVM for soil moisture estimation problems. Morellos et al. (2016⁹) evaluated four different regression models to predict soil moisture content by means of a portable VIS-NIR spectrophotometer. The models evaluated were principal component regression, partial least squares regression, least squares-support vector machines (LS-SVM) and Cubist, with LS-SVM outperforming the other three methods. Hassan-Esfahani et al. (2015¹⁰) evaluated the performance of ANN on surface soil moisture estimates using high-resolution imagery captured by AggieAir UAVs. Several vegetation indices, along with visual spectrum (RGB bands), near-infrared, infrared and thermal bands, were used as ANN input variables. Fifty soil moisture samples collected in the field were considered for training and testing the ANN model. In terms of R^2 , ANN accuracy was reported at 77%.

As evidenced in the review, traditional regression models are not able to model soil water content because it is a nonlinear function with several parameters that are not easily measured in the field. Thus, research is mostly focused on evaluation of machine learning models that are developed for both linear and nonlinear problems. Additionally, they can simply handle many inputs by adjusting the weighting vectors assigned for each of the inputs. Machine learning models have been a priority, but producing soil water content at the spatial scale requires spatial remotely-sensed input datasets. However, most studies have focused on the application of network-based machine learning models (e.g., SVM and ANN) for soil moisture prediction, and the performance of evolutionary-based models such genetic programming (GP) is still untouched. One of the disadvantages of network-based methods is that the trained network is not transferable, particularly between two platforms or two programming languages. In addition, studies in which soil moisture content is simulated or predicted using UAV high-resolution imagery are mostly limited to one flight due to flight costs. In these studies, the variation of soil moisture is too low, and machine learning performances is valid solely for a specific small range of soil moisture variation.

To overcome this limitation and show the performance of an evolutionary-based model, GP, which is a combination of an optimization algorithm and artificial intelligence, is employed in this study to simulate soil moisture using several high-resolution multi-spectral UAV imagery capture by the Utah State University AggieAir UAV system over two experimental pasture sites located in northern and southern Utah. The training inputs were reflectance for the visible, near-infrared (NIR), and thermal bands, and the output of the models was the soil moisture measured by accurate sensors at three different soil depths (near surface, medium, and deep layers). The results show the performance of GP versus ANN and SVM models for soil water content simulation.

2. METHODS AND MATERIALS

2.1 Study Areas

To evaluate the performance of three different machine learning techniques (SVM, ANN, and GP) on simulation of soil water content at three different soil layers using UAV imagery, two different Utah State University experiment pasture sites in northern and southern Utah are considered as the study area (Fig 1). The pasture sites contained primarily tall fescue. The field experiments included 20 50 x 20 m plots, with 4 replications of 5 irrigation levels.

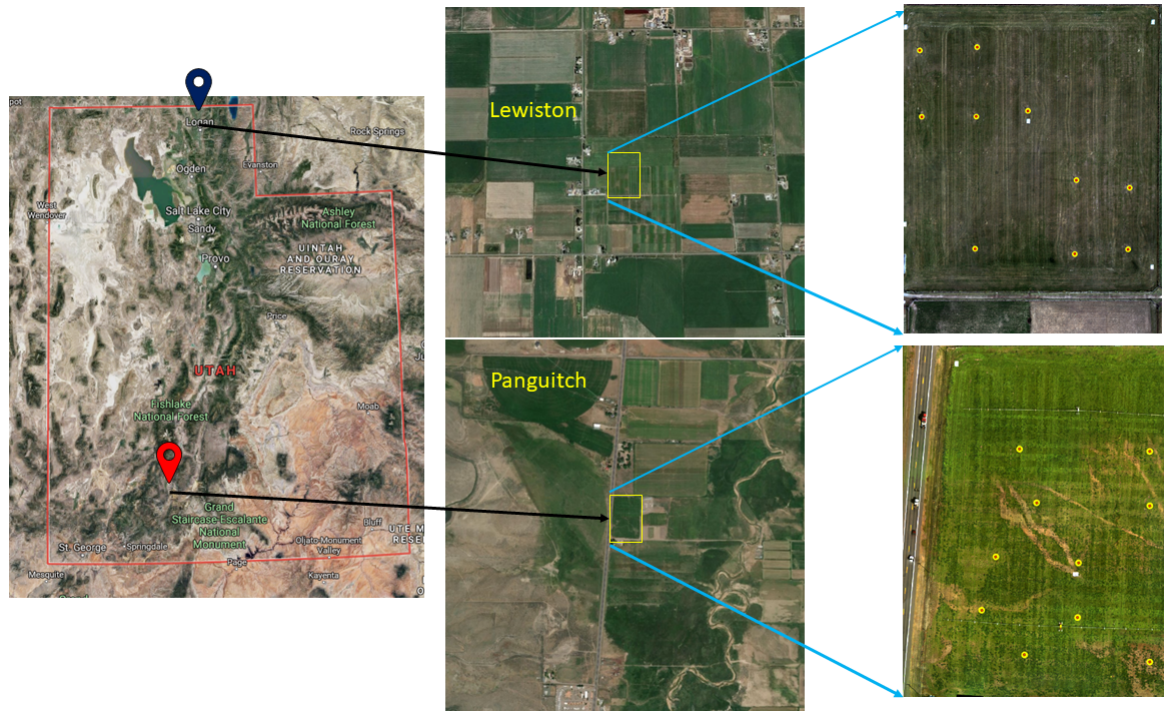


Figure 1. Two USU field experiment pasture sites located in northern (close to Lewiston city) and southern (close to Panguitch city) Utah. Red dots and yellow shape files are showing the location of soil moisture sensors along with their 1-m buffer, respectively.

Fig 1 shows the study areas along with the locations of Acclima TDR-315L soil moisture sensors installed in the experimental field. The TDR-315L is a true waveform digitizing Time Domain Reflectometer that estimates soil permittivity and water content from the propagation time of an electromagnetic impulse conveyed along its waveguide. It has an ultra-fast rise time step function generator, a waveform digitizer (200 GSPS) and a picosecond-resolution time base (5ps). Since it is a true time domain device, its readings are not impacted by soil electrical conductivity and compaction. It uses the standard SDI-12 interface and is compatible with all data recorders that are version 1.3 compliant. Red dots indicate soil moisture sensors and yellow shape files show a 1-m buffer around each sensor. The mean values of each multi-spectral band from UAV imagery for each buffer are extracted using a Python script. These values are considered as the input matrix into the machine learning methods.

Five sets of high-resolution imagery were acquired over the experimental fields (Table 1) using UAVs developed by the AggieAir team at USU. All of the images have 5 bands: Red, Green, Blue, near-infrared (NIR), and Thermal, but the last one has a virtual Red-Edge band as well. Also, the point cloud of each image was provided by the AggieAir team. The elevation of the first two flights was 400 m above the ground, and the images were captured using a fixed wing UAV (Fig 2). The elevation of the last three flights was 120 m, and the images were captured using a multi-rotor vertical take-off and landing (VTOL) UAV (Fig 2). For all flights, LAI was measured with a LAI-2200C sensor, elevation of ground control points was collected using high-accuracy real-time kinematic (RTK) GPS. Image resolution was 0.1 m for the first and second flights and 0.02 m for the last three flights.

Table 1. Date, times, locations, UAV elevations, and the bands' name of AggieAir multispectral UAV imagery

| Date | Flight Time | Location | UAV elevation (agl) meters | Bands |
|----------|-------------|-----------|----------------------------|-------------------------|
| 20170618 | 12:07-12:40 | Lewiston | 400 | RGBNIR Thermal |
| 20170822 | 10:20-10:41 | Lewiston | 400 | RGBNIR Thermal |
| 20171017 | 10:33-10:45 | Lewiston | 120 | RGBNIR Thermal |
| 20180814 | 13:26-13:36 | Lewiston | 120 | RGBNIR Thermal |
| 20180912 | 11:17-11:22 | Panguitch | 120 | RGBNIR Red Edge Thermal |

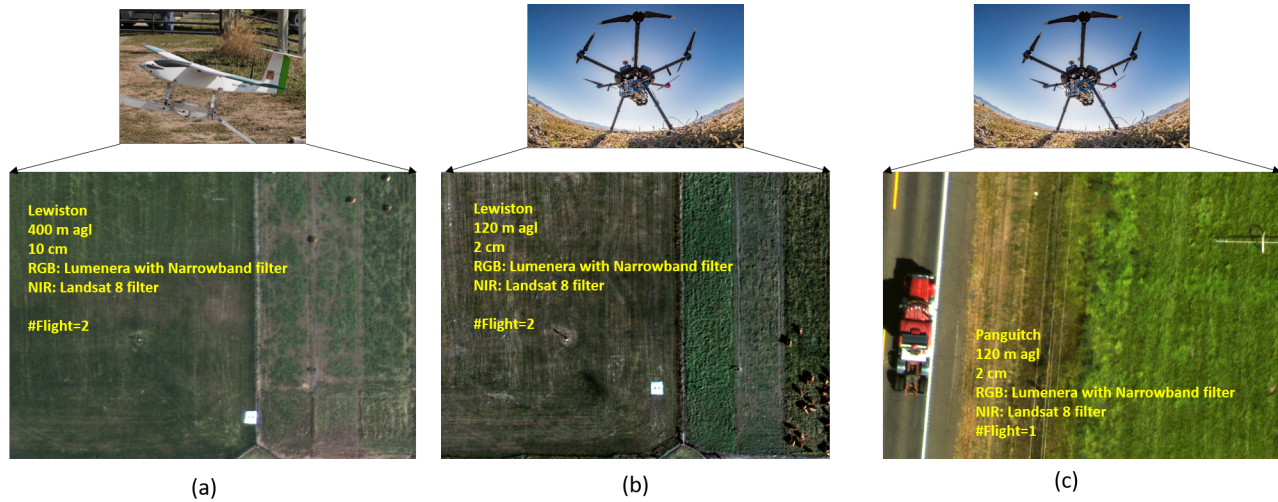


Figure 2. (a) Lewiston flights with a fixed-wing AggieAir UAV, (b) Lewiston flights with a VTOL AggieAir UAV, and (c) Panguitch flight with a VTOL AggieAir UAV

2.2 Methods

In this section, the theory of SVM, ANN, and GP as used to simulate soil moisture at different soil layers is described briefly.

2.2.1 Support Vector Machine: SVM

Vapnik (1995¹¹) presented a new machine learning method called support vector machine (SVM) for classification, regression, and clustering analysis. Since the regression version of the SVM model is used in this study, the support vector regression (SVR) model is described. In a linear-fitted problem, SVR, tries to fit Eq. (1) over training datasets.

$$y = f(x) = w^T \times x + b, \quad (1)$$

in which, y = a vector of output, $f(x)$ = a linear regression function of input vectors, w = weighting vector, x = vector of input variables, T = the transpose sign, and b = bias. In Eq. (1), output and inputs are known based on the training datasets separated from entire datasets. Therefore, the goal is to find w vector and b by minimizing an error function. Vapnik (1998¹²) presented Eq. (2) as an objective function and named it the epsilon intensive (e-intensive) function.

$$|y - f(x)| = \begin{cases} 0 & \text{if } |y - f(x)| \leq \kappa \\ |y - f(x)| - \kappa = \xi & \text{otherwise} \end{cases}, \quad (2)$$

in which, κ = permitted error threshold so that an error less than κ is zero and an error greater than κ is actual absolute error minus permitted error (ξ). Therefore, the e-intensive function is an absolute error function that is

intensive to a range of bias. This range is defined by users based on a κ threshold. The magnitude of κ depends on the magnitude of output and the object of the study. For instance, it cannot be 1 km when the maximum of output value is 10 m. In addition to the e-intensive function, Vapnik (1998¹²) added a regularization term ($\frac{1}{2}\|w\|^2$) to avoid over-fitting (Aboutalebi, 2018¹³) and a penalty coefficient (C) to adjust the weight of the e-intensive function in the SVM objective function (Eq. (3)).

$$\begin{aligned} & \text{Minimize } \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m (\xi_i) \\ & \text{subject to } |y_i - f(x_i)| - \kappa < \xi_i. \end{aligned} \quad (3)$$

Although SVM was first presented for linear regression problems, it can be used for non-linear problems by means of a non-linear kernel function ($K(x, x_i)$). Kernel functions are developed to transfer inputs in a space where inputs are linearly related to outputs. Radial Basis Function (RBF) and logarithm functions are popular kernel functions used in most machine learning methods. Therefore, Eq. (1) can change to Eq. (4) if outputs are non-linearly modeled with a vector of inputs.

$$y = f(x) = w^T \times K(x, x_i) + b. \quad (4)$$

The regression form of SVM has three parameters that need to be defined by users before executing: κ , C , and kernel function parameter (such as γ in RBF). Most researchers use a trial-error approach to tune SVR parameters. In complex problems, specifically when there are a high number of possible input parameters, the speed of SVR decreases significantly. Several approaches to decrease the size of the input matrix can help overcome this issue. For instance, Aboutalebi et al. (2015¹⁴) coupled SVM with an evolutionary optimization algorithm called a non-dominated sorting genetic algorithm (NSGAI) to automatically tune SVR parameters and to add an input selection feature. The superiority of SVR-NSGAI compared to the original SVR form has been demonstrated in several case such as hydrograph routing (Ref.¹⁵), water quality monitoring (Ref.¹⁶), pollution in a river-reservoir system (Ref.¹⁷). However, since the input matrix to simulate soil moisture content is set to multi-spectral imagery band values, and having a fair comparison between SVR and ANN, the traditional form of SVR is employed in the current study.

2.2.2 Artificial Neural Network: ANN

ANN is a machine learning method that, similar to SVM, relates inputs to output using a linear function (Eq. 1), and by substituting kernel functions in Eq. 1, ANN can be used as a nonlinear regression model (Eq. 4). However, two major differences exist between SVM and ANN. The first is related to the objective function defined for ANN. As opposed to SVM and its specific objective function (e-intensive function), the ANN objective function is defined by users. RMSE, R^2 , and absolute error are popular objective functions considered in ANN to find weighting vectors (w) and bias (b). The second difference is related to the network structure of ANN versus SVM. In SVM, only one middle layer (usually called the hidden layer) with one neuron exist. In contrast, the number of hidden layers along with their neurons in ANN is a user-defined parameter (Fig 3). Increasing the number of hidden layers and neurons leads to increasing computation time and network complexity. Despite recommendations for the number of hidden-layers and neurons, most ANNs have only one hidden layer, and the number of neurons is defined by a trial and error process. In this study, an ANN with one hidden layer is used to simulate soil moisture records in different layers of soil.

2.2.3 Genetic Programming: GP

Genetic Programming (GP) is a machine learning method inspired by the genetic algorithm (GA). In contrast to ANN and SVM output, which is a trained network, the output of GP is a trained equation that researchers can simply use. Similar to GA (Fig 4), GP uses a searching process to solve optimization problems. It starts with many possible solutions in the form of chromosomes, in which each gen could be a function (*sin*, *log*, *cos*, and *exp*), an operator (+, -, /), an input variable (x_1, \dots, x_n) or a number (1, 2, 3, ..., n).

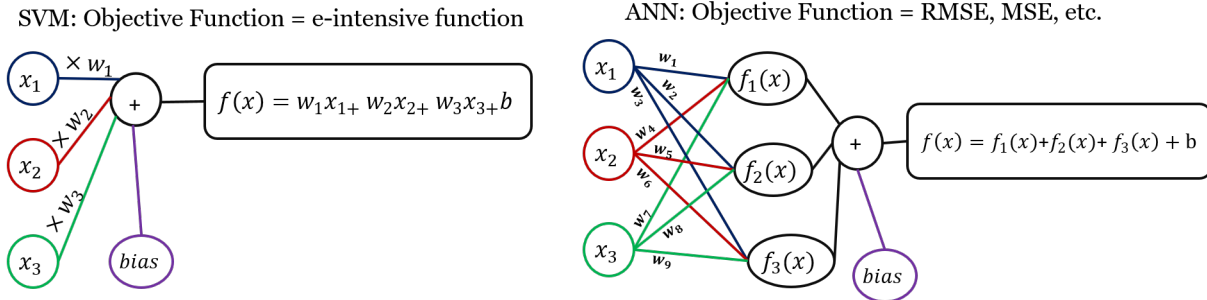


Figure 3. Schematic components of SVM (one hidden layer and one neuron) and ANN (one hidden layer and 3 neurons)

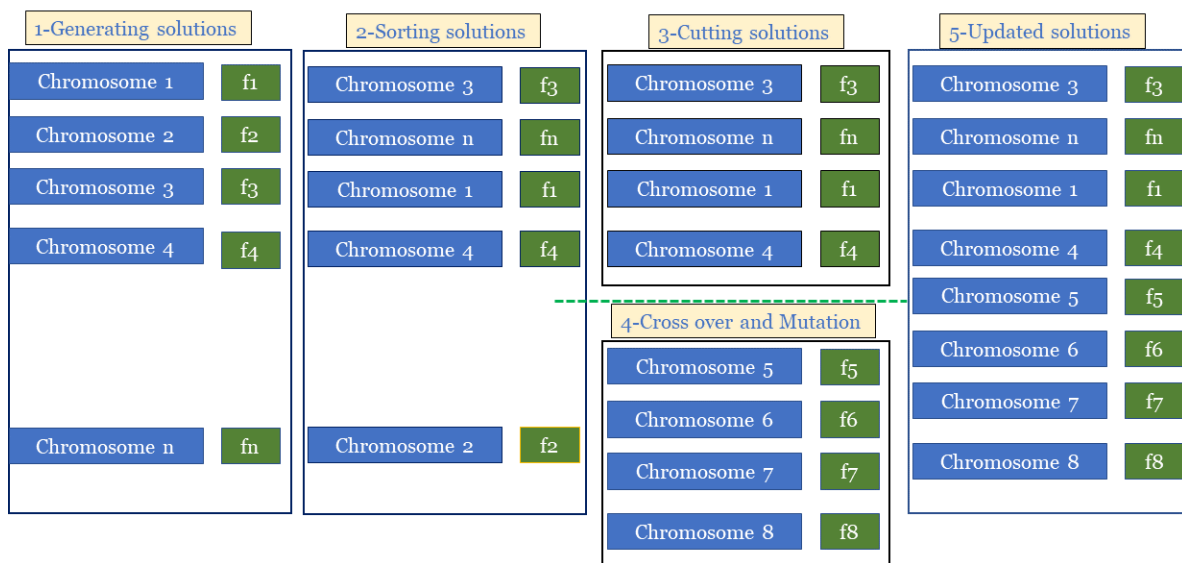


Figure 4. A graphical visualization of the various stages of GA

In iteration 1, chromosomes are generated by a random initial solution. Then an objective function (e.g. RMSE) for each chromosome is calculated based on the output of chromosomes and training datasets. These chromosomes are sorted for the next iterations. In subsequent iterations, chromosomes must be updated. Each chromosome can be modified in each iteration of the search process using cross-over and mutation functions. Cross-over is responsible for interpolation between two chromosomes, and mutation is designed for extrapolation. In each iteration, if the stopping criteria (e.g. number of iterations $< 1e6$) is satisfied, GP will stop, and the first among the sorted chromosomes, which is a fitted linear or nonlinear equation, is reported. Fig 5 shows the evolving process for one chromosome after one iteration using mutation and cross-over functions.

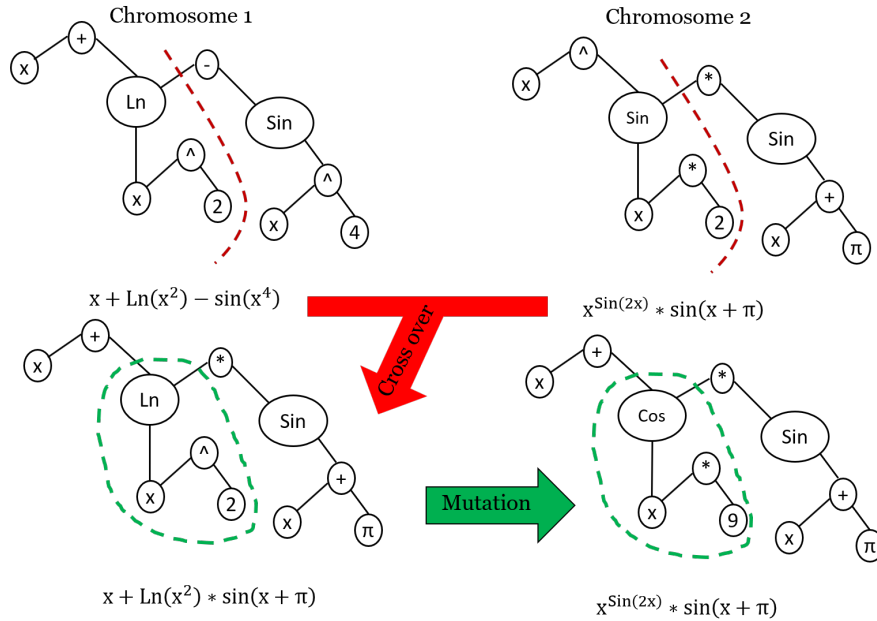


Figure 5. A graphical visualization of the various stages of GP

3. RESULTS AND DISCUSSION

3.1 Soil Moisture Records

As shown in Fig 1, Acclima TDR-315L (true waveform digitizing Time Domain Reflectometer) soil moisture sensors are installed at 10 locations in each pasture site. These sensors provide hourly soil moisture content (% by volume) at three different soil depths: 15 cm, 45 cm, and 76 cm. A timeseries of soil moisture content recorded by one of these sensors in the Lewiston field area is shown in Fig 6.

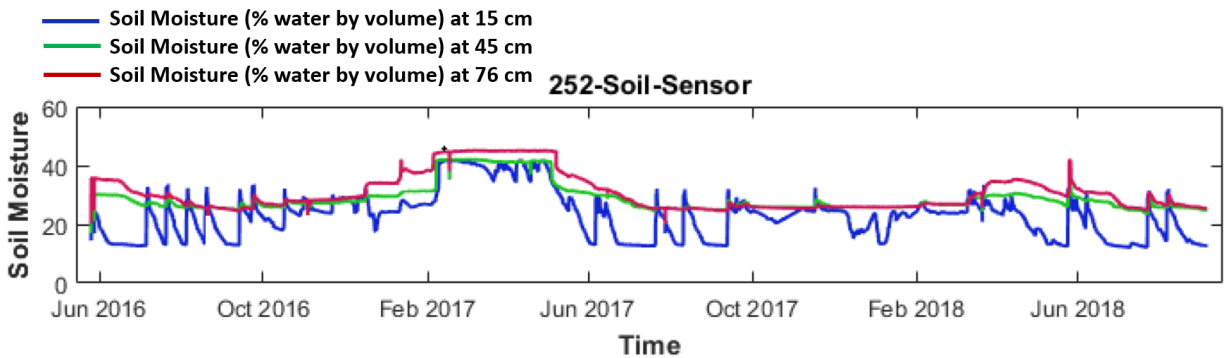


Figure 6. A timeseries of soil moisture content at three different soil depths recorded by sensor 252 in the Lewiston pasture site

At the UAV flight time (Table 1), we found no missing/outlier record except one that is removed from the entire dataset. Therefore, at each flight time, the values for soil moisture at the different soil layers for each sensor is extracted and considered as an output matrix in the machine learning techniques.

3.2 Evaluation of soil water content versus each band of UAV imagery

Fig 7 allows us to assess possible correlations between soil water content from installed sensors and the mean values of each band from multi-spectral UAV imagery (R, G, B, etc.) for each 1-m buffer.

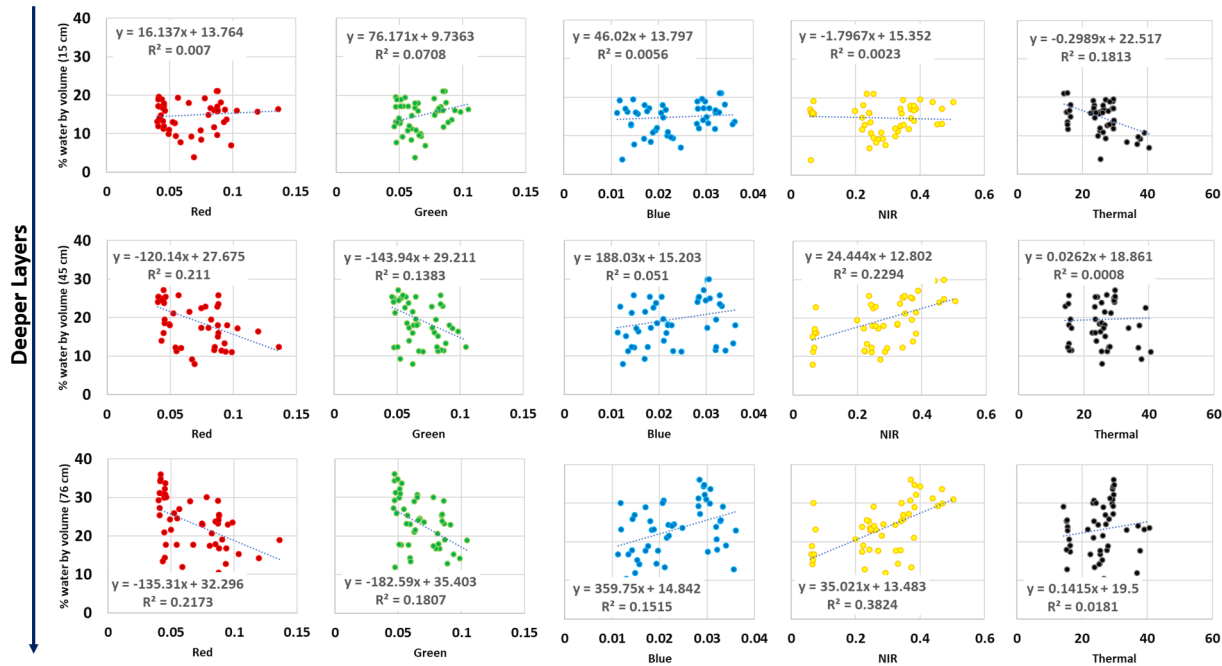


Figure 7. Scatter plots showing the relationship between soil water content at different soil layers and each band of AggieAir UAV imagery

As shown in Fig 7, soil moisture cannot be modeled with the information from one band. In terms of R^2 , the strongest relationship in near surface, middle, and deepest layer are for thermal, Red, and NIR bands, respectively. However, the maximum R^2 is less than 40%, which is not enough to present a reliable model. Increasing the soil layer depth increases the relationship between soil moisture and visible bands; however, the correlation between soil moisture and the thermal band decreases in deeper soil layers. The low values reported for R^2 show that simple linear regression models cannot be useful for simulating or predicting soil water content at different soil layers. Therefore, machine learning methods, which are able to find a linear and non-linear function between several inputs and an outputs, are employed to overcome the issues related to low R^2 values. It should be mentioned that the red edge information is removed from the analysis due to insufficient red edge values and soil moisture records (9 records only in the last flight).

3.3 The performance of SVM, ANN, and GP in simulation of soil moisture records

In this study, original versions of SVM and ANN that are built-in models in MATLAB software are used to estimate soil water content with information from multi-spectral AggieAir UAV imagery. To assess the performance of GP, we use Eureqa software (Ref. ¹⁸ and Ref. ¹⁹) and execute GP with the same inputs for ANN and SVM. 70% of the records are randomly selected for training machine learning models and 30% are considered for testing the models. SVM and ANN are executed using default parameters in MATLAB's built-in function. For GP, in addition to constant, addition, subtraction, multiplication, and division, sin, cos, power, exponential, and square root function are considered in formula building-blocks. Fig 8 shows the simulated soil moisture from SVM, ANN, and GP versus the observed ones.

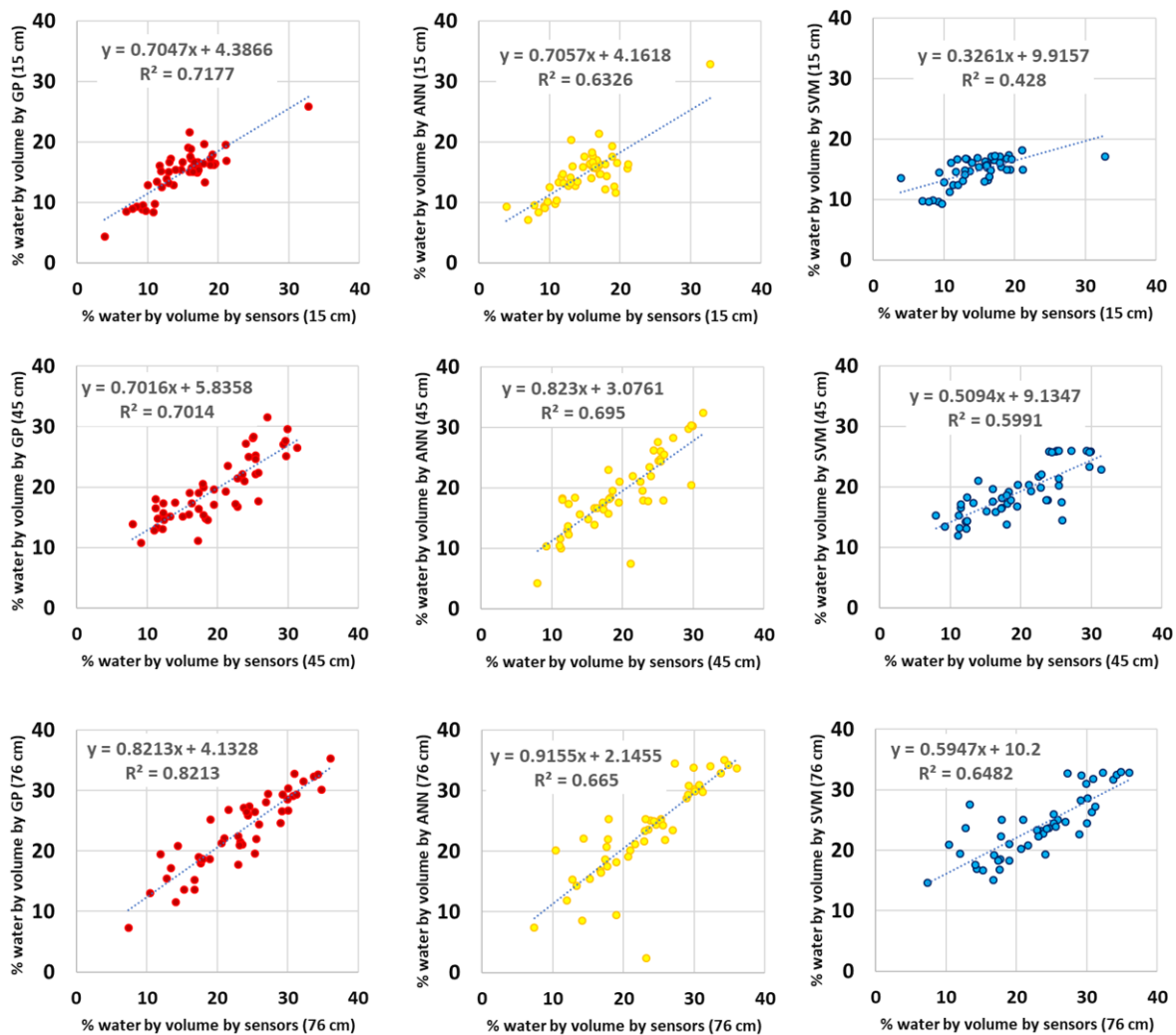


Figure 8. Scatter plots showing the simulated soil water content from machine learning methods versus that observed from soil moisture sensors

As shown in Fig 8, GP outperformed ANN and SVM in estimation of soil moisture in all soil layers except the middle layer. At the middle soil depth, the performance of ANN and GP is similar. In the deepest soil layer, GP performed very well compared to ANN and SVM, and the agreement of both GP and SVM with observed records increases in deeper soil layers. In terms of R^2 , GP, ANN, and SVM can estimate soil water content on average at the different soil layers with 74%, 66%, and 55%, respectively, which shows the superiority of GP over ANN and SVM. According to the results and due to the low accuracy of SVM in terms of R^2 in the surface layer, employing SVM to produce surface soil moisture at the spatial scale is not recommended. For the middle layer, although the accuracy of GP is higher than ANN when considering R^2 , the uncertainty of ANN results appears to be lower than for GP. For the deepest layer, the GP performance ($R^2=82\%$) is significantly higher than ANN ($R^2=66\%$) and SVM ($R^2=64\%$).

As discussed, the trained ANN and SVM networks are not transferable in a practical mode, which prevents reproducible results. However, GP provides a master equation that can be easily transferable among researchers. The GP master equations for the current study are shown in Eq. (5) – Eq. (7).

$$SM_1 = 1390.5544 * G + 84.7259 * N + 0.9102 * Thermal + 889.1360/Thermal + 94546.6195 * B^2 - 114.8553 - 57592.6512 * G * B - 78026.9360 * N * B^2 \quad (5)$$

$$SM_2 = 16.3772 + 38.9839 * N - 0.7653 * N * Thermal - 3.4282 * \sin(202210.9482 * N) - 6.0288 * \sin(\sin(5.8336 + 212321.8258 * N) - 9.5490 * B * Thermal) \quad (6)$$

$$SM_3 = 38.0639 + 1388.4115 * B * N + -0.0570 * \sin(634.0062 * G + 3.4110 * N)/(R * N) - 29.9377 * N - 253.0933 * G - 4.7097 * \sin(2744257.0988 * B) \quad (7)$$

in which, SM_1 , SM_2 , and SM_3 are soil water content at 15 cm, 45 cm and 76 cm below the surface soil layer. although the empirical equations found by GP are complex, they can be easily evaluated in different case studies using a simple spread sheet.

4. CONCLUSION

In this study, the performance of three different machine learning methods (GP, ANN, and SVM) for estimating soil moisture at different soil layers using AggieAir UAV multi-spectral imagery is presented. We found that, among five UAV imagery bands, the NIR and thermal bands were more correlated to soil moisture content. With increasing soil depth, the performance of machine learning technique increases, particularly the GP and SVM accuracy. The performance of GP and ANN were similar at surface and middle soil depths. The agreement between SVM and observed results was very low in the surface soil layer. The accuracy of GP at 76 cm below the soil surface layer was significantly higher than ANN and SVM. In addition to superiority of GP compared to ANN and SVM in terms of R^2 , the output of GP is a transferable equation that can be used and tested by researchers in different case studies.

ACKNOWLEDGMENTS

The authors would like to thank Carri Richards for editing the manuscript and the Utah Water Research Laboratory for financially support of this project.

REFERENCES

- [1] Ali, I., Greifeneder, F., Stamenkovic, J., Neumann, M., and Notarnicola, C., "Review of machine learning approaches for biomass and soil moisture retrievals from remote sensing data," *Remote Sensing*. **7**, 16398–16421 (2015).
- [2] Srivastava, P.K., Han, D., Ramirez, M.R., and Islam, T., "Machine learning techniques for downscaling SMOS satellite soil moisture using MODIS land surface temperature for hydrological application," *Water Resources Management*. **27**, 3127-3144 (2013).
- [3] Gill, MK., Asefa, T., Kemblowski, M., and McKee, M., "Soil moisture prediction using support vector machines," *JAWRA Journal of the American Water Resources Association*. **42**(4), 1033-1046 (2006).
- [4] Wu, W., Wang, X., Xie, D., and Liu, H., "Soil water content forecasting by support vector machine in Purple Hilly region," In: Li D. (eds) *Computer And Computing Technologies In Agriculture, Volume I. CCTA 2007. The International Federation for Information Processing*. **258**. Springer, Boston, MA, (2008).
- [5] Liu, H., Xie, D. and Wu, W., "Soil water content forecasting by ANN and SVM hybrid architecture," *Environmental Monitoring and Assessment*. **143**(1), 187–193 (2008). <https://doi.org/10.1007/s10661-007-9967-9>.
- [6] Liu, D., Yu, Z., and L, H. "Data assimilation using support vector machines and ensemble Kalman filter for multi-layer soil moisture prediction," *Water Science and Engineering*. **3**, 361-377 (2010).

- [7] Zaman, B., McKee, M., and Neale, C.M.U. "Fusion of remotely sensed data for soil moisture estimation using relevance vector and support vector machines". *International Journal of Remote Sensing*. **33**, 6516-6552 (2012).
- [8] Pasolli, L., Notarnicola, C., Bruzzone, L. "Estimating soil moisture with the support vector regression technique," *IEEE Geoscience and Remote Sensing Letters*. **8**, 1080-1084 (2011).
- [9] Morellos, A., Pantazi, X.E., Moshou, D. Alexandridis, T., Whetton, R., Tziotziou, G. Wiebensohn, J., Bill, R., and Mouazen, A.M., "Machine learning based prediction of soil total nitrogen, organic carbon and moisture content by using VIS-NIR spectroscopy," *Biosystems Engineering*, **152**, 1-13, (2016).
- [10] Hassan-Esfahani, L., Torres-Rua, A., Jensen, A., and McKee, M. "Assessment of surface soil moisture using high-resolution multi-spectral imagery and artificial neural networks. *Remote Sensing*. **7**, 2627–2646 (2015).
- [11] Vapnik, V. "The nature of statistical learning theory," Springer, New York (1995).
- [12] Vapnik, V. N. "Statistical learning theory," Wiley, New York (1998).
- [13] Aboutalebi, M. "Discussion of Equation to Predict Riverine Transport of Suddenly Discharged Pollutants by Mostafa Farhadian, Omid Bozorg-Haddad, Samaneh Seifollahi-Aghmiuini, and Hugo A. Loaiciga," *Journal of Irrigation and Drainage Engineering*, **144**(4) (2018). DOI: 10.1061/(ASCE)IR.1943-4774.0001269.
- [14] Aboutalebi, M., Bozorg-Haddad, O., and Loaiciga, H. "Optimal Monthly Reservoir Operation Rules for Hydropower Generation Derived with SVR-NSGAI", *Journal of Water Resources Planning and Management*, **141**(11) (2015). DOI:10.1061/(ASCE)WR.1943-5452.0000553.
- [15] Aboutalebi, M., Bozorg Haddad, O., and Loaiciga, H.A., "Application of the SVR-NSGAI to hydrograph routing in open channels," *Journal of Irrigation and Drainage Engineering*, **142**(3) (2015). DOI : 10.1061/(ASCE)IR.1943-4774.0000969.
- [16] Aboutalebi, M., Bozorg Haddad, O., and Loaiciga, H.A., "Multi-objective design of water-quality monitoring networks in river-reservoir systems," *Journal of Environmental Engineering*.**143**(1) (2016).DOI: 10.1061/(ASCE)EE.1943-7870.0001155.
- [17] Aboutalebi, M., Bozorg Haddad, O., and Loaiciga, H.A., "Simulation of methyl tertiary butyl ether concentrations in river-reservoir systems using support vector regression," *Journal of Irrigation and Drainage Engineering*. **142**(6) (2016). DOI: 10.1061/(ASCE)IR.1943-4774.0001007.
- [18] Schmidt, M., Lipson, H., "Distilling free-form natural laws from experimental data," *Science*. **324**(5923), 81–85 (2009).
- [19] Schmidt, M., Lipson, H., "Eureqa (Version 0.98 beta) [Software]," Available from www.nutonian.com, (2014).