A Neural Network Modelling for Soil Moisture Prediction

MSc Research Project
Data Analytics

Shubham Patil
Student ID: X17165768

School of Computing
National College of Ireland

Supervisor: Bahman Honari

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Shubham Patil
Student ID:	X17165768
Programme:	Data Analytics
Year:	2019
Module:	MSc Research Project
Supervisor:	Bahman Honari
Submission Due Date:	12/12/2019
Project Title:	A Neural Network Modelling for Soil Moisture Prediction
Word Count:	8537
Page Count:	25

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	
Date:	12th December 2019

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

A Neural Network Modelling for Soil Moisture Prediction

Shubham Patil

X17165768

Abstract

The infield of agricultural, farmers do need to perform various activities throughout the year on land surface as well on subsurface and these activities get affected by soil moisture. That's why it makes soil moisture game-changer in agricultural for farmers to achieve growth in the farming business. So, it is very essential to get trusted and most accurate real-time soil moisture data in order to help crop water demand with managing water resources. This study focused on use of neural network-based models such as $LSTM$(Long short-term memory) and ANN(Artificial Neural Network) to help farmers by prognosticating 24-Hours before soil moisture with help of this farmer would get awareness of crop water needs and help them to improve water resources management in field of agriculture. This developed model trained on past volumetric soil moisture data and able to forecast the next day that is 24-Hours in the future. The neural network model trained on U.S.Geological Survey(USGS) data independently on two sites. The various types of transformation introduced in the field of soil moisture prediction to get high performance while training neural network models without damaging its true sequential nature. The result of each site of developed models is compared against each other and with MSE(Mean Squared Error) as an evaluation parameter. A value of MSE of $LSTM$ model is 0.001552 for the Dana Meadow site and for the Gin Flat, it is 0.002365. Whereas, ANN model is 0.001316 for the Dana Meadow site and the Gin Flat is 0.002022. Results of $LSTM$ model outperformed ANN in terms prediction of volumetric soil moisture for 10cm and 36cm depth for next day that is 24-Hours in future for both sites by catching all oscillation like true values.

1 Introduction

The soil moisture is one of the important ingredients for success in the overall agricultural business process because It plays an important role in the growth of crops. The real-time and accurate soil moisture data can help in various ways such as to improve and manage water resources in the farm, to get help in understanding rainfall and rain-runoff process, irrigation scheduling and also gives acknowledgment about the level of water in the agricultural field. The volumetric soil moisture content is the fuel for many chemicals, physical as a well biological process in agricultural fields(Liu; 2011). It also indicates the energy status of soil water and water irrigation needs for the growth of crops(Shukla; 2013). In the process of proliferation for any crop field in the future is determined by three key factors, first one is content of soil nutrients and their properties, second factor dealing with the quality of seeds selection with regards to their properties and the third factor is soil moisture level(Subir; 2011). The soil moisture also supports an automatic irrigation management system, water-saving, to increase yield, for soil salinity control and also helps in controlling soil erosion(Aniley1 et al.; 2018). Real-time accurate soil moisture can also help for the examination of climatic change and it's the rate of effect on the agricultural field as well in the finding of water needs in the water-limited regions all around the world(Ansari and Deshmukh; 2017). The soil moisture generated in soil totally depends on sufficient rainfall to exceed losses, evapotranspiration of water from the agricultural field, drainage to water irrigation as well on ground level water(Beven; 2019). The soil moisture measured using two units such as Gravimetric and Volumetric. In this research, the volumetric unit measured values are used. The Volumetric is for the volume of water per volume of soil(Bilskie and Scientific; 2001).

In recent years, several researchers developed their models on neural networks for soil moisture prediction as well for estimation. Such as support vector regression, recurrent neural networks, multiple regression and artificial neural network for forecasting of soil moisture(Prasad et al.; 2019)(Chatterjee et al.; 2018). These studies shown good prediction capability for soil moisture, authors also reported their views on the use of the LSTM model to handle time series sequential data(Prasad et al.; 2019). The researchers also noticed that the use of the LSTM model performed well over various other machine learning techniques(Adeyemi et al.; 2018)(Zhanga et al.; 2018). From the literature review, it is evident that a robust and scalable data-driven model for soil moisture prediction developed using the LSTM model. Researchers also noticed that other models struggle to preserve previous data while training on data and this model gives a sub-optimal prediction. The LSTM model can work on minimum pre-process data and able to store information across multiple time steps(Adeyemi et al.; 2018). The LSTM model also very well in another research topic such as classification of chaotic system(Ordóñez and Roggen; 2016), natural language processing(Mikolov et al.; 2011) and speech recognition(Graves et al.; 2013).

In this study, the LSTM model is developed for the prediction of soil moisture for 10cm and 36cm depth for two sites. Each site is trained independently with the proposed model. The performance of the model will get evaluated by comparing the LSTM predicted the soil moisture content of depth 10cm and 36cm with actual true values of soil moisture using graphs as well with predicted soil moisture using ANN model.

1.1 Research Question

- How to prognosticate Dana Meadow and Gin Flat site's volumetric soil moisture of 10cm and 36cm depths one day ahead i.e. 24-Hours in future with help of proposed neural network?

The rest of this research is divided into following sections. In Section 2, Related Work is presented. In Section 3, Methodology is presented. In Section 4, Design Specification is presented. In Section 5, Implementation is presented. In Section 6, Evaluation is presented. In Section 7, Conclusion and Future Work are presented.

2 Related Work

In this Section, previous studies will be reviewed for soil moisture prediction.

2.1 Prediction of soil moisture using deep learning

In agricultural farmers faced lots of challenges in terms of hydrogeological characteristics, boundary conditions and lots of human activities while predicting soil moisture by the physical model. to overcome this problem (Zhanga et al.; 2018) authors developed a time series model based on Long Short-Term Memory (LSTM). This proposed model assembles using the LSTM layer with another LSTM layer on top of it as well as a dropout method applied in the first LSTM layer. For this experiment 14 years of data used from China's Hetao Irrigation District of five sub-areas. This data used to employed model on monthly water diversion, evaporation, precipitation, temperature and soil moisture content. In these 14 years of data 12 years of data used as a training set and rest of data as a validation set for testing. The help of the atop a connected layer of LSTM helped in preserving the previous information, which helped to learn ability of the model for time series data and it worked like PCA. The authors also set dropouts on the LSTM layer to avoid overfitting. (Zhanga et al.; 2018) used the coefficient of determination (R^2) and root mean square error($RMSE$) for evaluation of the model's accuracy between the measured and predicted values. Whereas a study where soil properties were forecasted with soil moisture by (Singh and Kasana; 2019) authors. Dataset used for this experiment taken from (LUCAS) is a project database developed to manages land surface character and data used for this from the years 2009 to 2011.Before passing this data to the proposed model right after preprocessing it passed through Principal Component Analysis (PCA). it used to reduce the dimension of given data which helped (Singh and Kasana; 2019) authors to get efficiency in the prediction of models and authors also aware of the nature of given data due to that they used LSTM in building this prediction model to make proper use of sequential soil dataset. Because this network is suitable for remembering short and long term dependencies making it ideal for hyperspectral data.Unlike previously discuss study here two layers of LSTM used while implementing. The components of PCA chosen in such a way that it can handle variance of a dataset then processed data from PCA were randomly divided into training and testing sets. These sets are to give to the LSTM model to learn and predict the results for these problems. Then the output of trained model is stored and compared to given actual output for evaluation of forecasting capability of the developed trained model.

On the whole, in the research of (Singh and Kasana; 2019) authors shown the use of double LSTM in the prediction of soil moisture and other soil properties and researches also use PCA for dimensional reduction of the database. This double LSTM model outperforms all other machine learning models in terms of results and working robustness and coefficient of determination (R^2) of this model were 0.94. On the other hand authors (Zhanga et al.; 2018), proposed a novel approached rather than double LSTM where the LSTM layer used on top of another LSTM layer with a dropout layer over-applied in the first layer. This proposed model achieved a higher coefficient of determination (R^2) 0.95 in soil moisture prediction than other deep learning models such as a feed-forward neural network(FFNN) and this proves that the proposed model can preserve and learn itself from previous information.

To robust decision making and prediction soil moisture content actually help in forecasting crop water needs. By keeping these minds (Adeyemi et al.; 2018) authors presented a dynamic neural network model for the soil moisture prediction to help predictive irrigation scheduling. Data is gathered from the COSMOS monitoring project in the United Kingdom. This data made up of real-time soil moisture and climate variables. For this study author used three sites of different soil types and the dataset ranges from 2014 to 2017 for all three. Before applying the proposed models on the given dataset were preprocessed with box-cox transformed data and seasonal, trend and residual components. For this study (Adeyemi et al.; 2018) used two models to compared and evaluated in the prediction process of soil moisture. Each model independently trained for three sites. And these forecasted results are used as feedback for predictive irrigation scheduling to predict crop water demands, which then evaluated parallel along with a pre-stimulated model of AQUCAROP in potato growing seasoning. another study where the use of soil moisture active passive (SMAP) mission of NASA's data used to estimate soil moisture. In this research (Fang et al.; 2017) authors, created a model to predicts SMAP level-3 soil moisture with additional atmospheric variables. For this research author focused on SMAP L3 passive radiometer product daily dataset and additional atmospheric data is also gathered from the North American land data assimilation system. Dataset of SMAP photographic in nature by considering this from April 2015 to March 2016 data used for training and April 2016 to March 2017 for testing purposes of the model. For computational efficiency, (Fang et al.; 2017) authors used one pixel from every 4*4 patch. Right after the preprocessing of the dataset, it passed through the LSTM model. The given model also tested three methods such as the least shrinkage and selection operator, autoregressive model and single-layer feedforward neural network with the same set of the dataset. Similarly (Fang et al.; 2019) used SMAP data for estimation of soil moisture but in this researchers try to find out the accuracy of the SMAP dataset by evaluating against situ measurements. Independently both of this dataset used for training and testing purpose to find out the prediction accuracy of LSTM using SMAP and LSTM using situ datasets.

In the end (Adeyemi et al.; 2018) researchers focused on forecasting soil moisture to help predictive irrigation scheduling of agricultural area to help crop water demands. For all three sites, authors got the same coefficient of determination (R^2) i.e. above the 0.94 for model evaluation and this helped in saving water from 20 to 40 percent overall. This proves usefulness over other deep learning models. In opposite to this (Fang et al.; 2017) authors, used SMAP data for modeling to prediction of soil moisture. during the implementation, researchers faced the problem with not well optimization of the model and they also not addressed anything regarding the accuracy of results. whereas, (Fang et al.; 2019) shown the accuracy of LSTM for prediction of soil moisture using SMAP dataset over long term prediction. researchers also noticed that LSTM trained with SMAP as the target and it adopted discrepancies in between situ and SMAP data.

Using a deep learning regression model with the help of soil moisture data and meteorological data (Cai et al.; 2019) authors predicted soil moisture. In this study, the authors tried to overcome problems from existing research such as prediction accuracy, generalization, and multi-feature processing capability to get optimum prediction performance. For this dataset are integrated while doing preprocessing acquired data then time-series parameters are analyzed of given set variables and their relationship between while predicting soil moisture with help of Taylor diagram.(Cai et al.; 2019) The authors also noticed that the meteorological variables impact lot soil moisture variables in this process of forecasting. Time series data gathered from the Beijing Meteorological Bureau and it contains soil moisture data with environmental variables. The period covered by this data from 2012 to 2016 with 10cm and 20cm depth soil moisture content of farmland. In preprocessing missing values are removed and correlation analysis performed to see the correlation between variables. This indicated strong correlation characteristics between the variables. To check the proposed model accuracy it got compared with linear regression(LR), support vector regression(SVR) and artificial neural network(ANN). In another study where authors used linear regression(LR), support vector regression(SVR) and recurrent neural network(RNN) for forecasting soil moisture in range of 1day, 2days and 7days ahead from given date(Prakash et al.; 2018). Each of these models applied independently to a different dataset collected from various data repositories of soil moisture. The first dataset gathered from Braggs farm in Alabama of Unites states department of agriculture from June 2015 to December 2016. The second data set acquired from the TAMU North American soil moisture database from January 2000 to September 2012. Similarly, the third dataset took from theOz Net Hydrological monitoring Network of the Australian monitoring network from March 2016 to May 2016. Then this collected data is normalized using the standard scaling technique and a total of nine variables were selected for the implementation of the soil moisture prediction model. Dataset divided into 80 percent for training and 20 percent for independent test.(Prakash et al.; 2018) authors noticed that with given models multiple linear regression(MLR) is superior to SVM and RNN. In another case of (Chatterjee et al.; 2018) to give sustainable help in agricultural application authors used a modified flower pollination algorithm(MFPA) to train artificial neural networks (ANN) to predict soil moisture quantity. In this study soil moisture quantity forecasted in terms of soil temperature, air temperature, and relative humidity. Multilayer perceptron feed-forward network(MLP-FFN) used for ANN, which determined how much adjustment is required to the weights of the network. Dataset of this study gathered from a publicly available repository of the year 2010 which also consists of environmental variables such as air temperature, relative humidity, and soil temperature. Results shown by this model in terms of root mean square error($RMSE$) have proved ingenuity of the proposed model.

on the whole, the research of (Cai et al.; 2019), shows the superiority of deep learning regression model over SVM, ANN and LR in compared and evaluated results with given performance measures. whereas, (Chatterjee et al.; 2018) models shown and established the ingenuity cause of its highly immune to changes in the presence of varying weather conditions. And (Prakash et al.; 2018), shows that in between implemented models given multiple linear regression(MLR) performed superior. But (Prakash et al.; 2018) authors also suggested LSTM is better suited for long term dependencies and it will more purposeful over regression models for predicting soil moisture.

In dry sub-humid environments having soil moisture-holding capacity data is essential for sustainable crop growth and agrohydrological functions. That's why (Kaingo et al.; 2018) authors proposed a novel approach. Authors also noticed that predicting and measuring soil moisture physically hard and costly. hence they used pedotransfer functions(PTFs) who use physicochemical properties of soil as input for prediction of soil moisture. In this study, the authors used the support vector machines method for the implementation of PTFs for dry sub-humid tropics. For this researcher used a soil dataset made up of 296 samples collected from June 2014 to July 2015 and took from a hundred different locations. Afterward, data get split into a 2:1 ratio for training and testing set. Correlation analysis, descriptive statistics and normality test performed on the datasets. To check the performance of the proposed model (Kaingo et al.; 2018) authors made another model of multiple linear regression(MLR)-PTFs developed and tested along with SVR-PTFs. Both of these models were evaluated using root mean square error($RMSE$), mean error(ME) and coefficient of determination (R^2) as indicators. Similarly, using PTFs to forecast soil hydraulic properties from available datasets to avoid cost and time consuming of physical work (D'Emilio et al.; 2018) authors presented an artificial neural network(ANN) for prediction of water holding capacity of Sicilian agricultural soil. Data collected from 359 soil horizons and it contains data from 21 different sites of the Sicilian territory. ANN presented in this study composed of one input layer, two hidden layers with 15 neurons in each and one output layer. With different combinations of input variables different ANN tested independently to see the predictive performance of model. (D'Emilio et al.; 2018) authors used normalized root mean square error($NRMSE$), normalized mean error(NME) and coefficient of determination (R^2). Using PTFs function is formed using the properties of saturated hydraulic conductivity(Ks) for prediction of the fate of water in soils(Araya and Ghezzehei; 2019). in this study authors compared four different models to check performance accuracy. The main aim of this study is to find out soil variables who impact a lot while predicting (Ks) and developed a robust model. In this (Araya and Ghezzehei; 2019) authors implemented four model such as k-nearest neighbors(KNN), support vector regression(SVR), random forest(RF) and boosted regression tree(BRT). Dataset gathered from the USKSAT for training and testing of all four models. In preprocessing of dataset all those records having missing values and more than one value were removed. To achieve optimal performance and model tuning purposes, the selection of optimal parameters done by using the k-fold cross-validation method. performance of all four models was evaluated using root mean square error($RMSE$), mean error(ME) and coefficient of determination (R^2).

To sum from the above studies, (Kaingo et al.; 2018) authors proposed model SVR-PTF performed to some degree better than previously developed MLR-PTFs and this study used PTFs while the prediction of soil moisture. In opposite to this, (Araya and Ghezzehei; 2019) authors developed four different models to maximize the use of the acquired dataset by developing this model which truly based on the PTFs model for estimation of Ks to forecast soil moisture. whereas, (D'Emilio et al.; 2018) authors implemented different five different variations of ANN with a different set of variables and shown ANN with all the environmental variables performed well then the rest of four ANN variations.

Soil moisture is not feasible as always in vast areas of agricultural regions cause of this farmer's face lots of difficulties while forecasting soil moisture in those areas. Hence (Khedri et al.; 2017) researchers proposed an imaging-based model. To develop this model author used a polarimetric synthetic aperture(PolSAR) imaging system used to implement the support vector regression model. In order to optimize the whole working of the present model (Khedri et al.; 2017) authors selected optimal parameters from the available sets of inputs using a feature selection method and it also helped in getting core base model prediction of soil moisture. feature selection is done using two methods such as sequential forward selection and sequential backward selection. For this study author used the AIRSAR dataset and proposed model performance evaluated using root mean square error($RMSE$) and coefficient of determination (R^2). In opposite to (Khedri et al.; 2017), (Zhaoa et al.; 2018) developed random forest-based model to disaggregate the soil moisture's active and passive with synergistic use of optical infrared observation from the moderate resolution imaging spectroradiometer onboard the terra and aqua satellites. Data for this model gathered from the study site of Iberian Penisula and dataset range from 2015 to 2016. Random forest model implemented on this set of data. Because of dataset consist of different overpass times in the daytime, hence four different combinations considered with different surface temperature are Terra and Aqua, soil moisture proxies are AM and PM.

To summarize the above two studies, (Khedri et al.; 2017) and (Zhaoa et al.; 2018) based on imaging dataset available from the satellite-based system as well both these models used a regression model for estimation of soil moisture. In (Khedri et al.; 2017) authors developed a model with SVR and used features utilizing feature selection technique to get more accuracy as compared to the SVR without optimized. This tells the importance of parameter optimization while implementing and developing models for getting better results. On another hand, (Zhaoa et al.; 2018) authors presented model missing with features like optimized parameters but (Zhaoa et al.; 2018) authors implemented a new method for downscaling passive microwave low-resolution surface soil moisture product with help of gathered data based on random forest regression method.

3 Methodology

For this study, CRISP-DM(Cross Industry Standard Process for Data Mining) approach have considered and each stage of CRISP-DM is describe as below below(Wirth; 2000)(vision Europe; 2018)(Jensen; 2012),

- **Business Understanding**: In this stage through the business perspective objectives of the study are understand. The main goal of this stage is to light on important factors that could influence the results of the project.

- **Data Understanding**: It deals with data acquisition of study resources. This stage handles data loading, data description, explore and visualize given data.

- **Data Preparation**: The datasets are clean, transformed and loaded for model implementation.

- **Modeling**: A modeling techniques are selected for problem statements with respective modeling assumptions. This stage also handles test generation, building the model and assessing the model.

- **Evaluation**: This stage checks the accuracy and generality of implemented models and it also seeks to determine reasons why the implemented model not performed well in terms of its results. After the assessment of model results and then models are approved with respect to business criteria.

- **Deployment**: This is the last stage of CRISP-DM, where the results of the evaluation stage help to determine a strategy for model deployment. This stage handles plan deployment, plan monitoring and maintenance as well as produce final reports and review projects.

3.1 Data Description

3.1.1 Study area

The data used for this research acquired from The U.S. Geological Survey of two different sites[1]. This data made up of an hourly time series record of soil temperature, volumetric soil moisture content. This data collected by the U.S. Geological Survey(USGS)[2] in cooperation with the California Department of Water Resources, National Park Service and Pepperwood Preserve at five different locations across Yosemite National Park. At each location, soil probe are installed from their total soil profile data collected. This dataset developed for the understanding of soil relation to climate and it also helps in the contribution of long term hourly time series soil moisture and temperature dataset. From an available set of dataset Dana Meadows, Gin Flat site's dataset is selected to developed and to test models. As shown in below table1, details of the sites datasets and in table2, data definition of both sites.

Site No.	Site Name	Data Range	Count of Records
1	Dana Meadows	09/11/2005 to 07/26/2017	100,420
2	Gin Flat	09/01/2005 to 07/26/2017	99,836

Table 1: Details of the sites data

Column Name	Data Type	Description
Date/Time	datetime64	Date wise records for hours
ST_10cm	float64	Soil temperature of depth 10cm in Celsius
ST_36cm	float64	Soil temperature of depth 36cm in Celsius
Soil_VWC10cm	float64	volumetric water content(soil moisture) of 10cm
Soil_VWC36cm	float64	volumetric water content(soil moisture) of 36cm
Soil_TWC	float64	Soil total water content

Table 2: Data Definition of both sites

3.2 Exploratory Data Analysis

In this section, exploratory data analysis took place on acquired datasets to understand important aspects of datasets before applying to actual models. Using pandas package datasets are loaded into python environment for further exploration. The data types of variables of raw data converted to required formats such as date time column data types converted to datetime64 and all remain variables converted to float. Data type conversion helped in maintaining integrity and uniformity across data. whole prediction will be depended on the date-time column by keeping this in mind it set as an index for both of the dataset using the available function of python. In the following subsection different comprehensive aspects as well tests on given time series data are visualized and tested to understand the true nature of data,

[1]USGS-Data `https://pubs.usgs.gov/ds/1083/ds1083_tables12-15_17.zip`.
[2]USGS `https://doi.org/10.3133/ds1083`.

3.2.1 Statistical Details of Data

The table3 below gives statistical description of Dana Meadows site variables,

Variables	Min	Max	Mean	Std	Skew-ness	Kurtosis
ST_10cm	-8.70	22.19	4.01	5.67	0.98	-0.38
ST_36cm	-3.96	13.66	4.07	4.28	0.74	-1.05
Soil_VWC10cm	0.02	0.57	0.11	0.13	2.33	4.76
Soil_VWC36cm	0.03	0.52	0.13	0.13	1.78	2.35
Soil_TWC	2.43	41.67	10.45	10.53	1.82	2.29

Table 3: Statistical Details of Dana Meadows's variables

The table4 below gives a statistical description of Gin Flat site variables,

Variables	Min	Max	Mean	Std	Skew-ness	Kurtosis
ST_10cm	-2.82	29.54	7.55	8.48	0.68	1.10
ST_36cm	-0.94	21.61	7.39	7.02	0.56	-1.32
Soil_VWC10cm	0.02	0.60	0.12	0.06	0.73	1.66
Soil_VWC36cm	0.05	0.49	0.17	0.09	0.93	1.18
Soil_TWC	3.14	36.77	11.07	6.23	0.95	1.00

Table 4: Statistical Details of Gin Flat's variables

3.2.2 Time Series Seasonality Plot

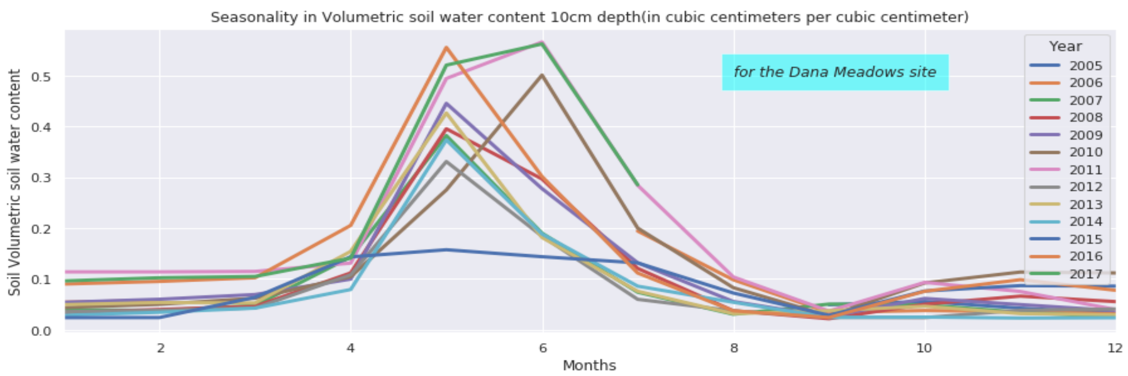


Figure 1: Seasonality Plot: Volumetric Water Content(VWC) of 10cm depth: Dana Meadows Site

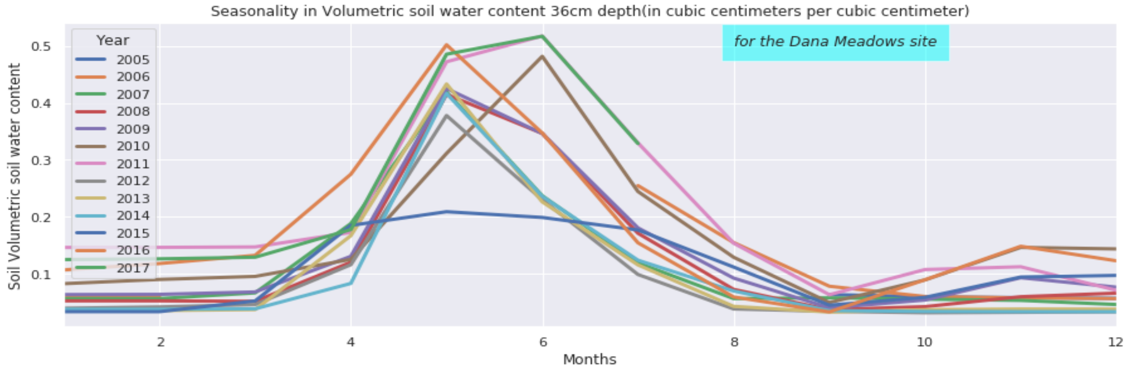


Figure 2: Seasonality Plot: Volumetric Water Content(VWC) of 36cm depth: Dana Meadows Site

From above figure1 and figure2 illustrate seasonal distribution of volumetric water content(VWC) from years 2005 to 2017 with months range from 1 to 12 for Dana Meadows site. In figure1, it depicts about volumetric water content (VWC) of depth 10cm for given years and months. From the figure1, it observed that for the first six months there is fluctuation in volumetric water content(VWC) for all given years. For the months 7 to 8 volumetric water content (VWC) is 0.05. whereas for months 9 to 12 volumetric water content(VWC) gradually climbing up from 0.05 to 0.20. In figure2, it shows volumetric water content (VWC) of depth 36cm for given years and months. From the figure2, it observed that for the first four-month there is fluctuation in volumetric water content(VWC) for all given years. For the months 7 to 10 volumetric water content (VWC) is 0.07. whereas for months 9 to 12 volumetric water content(VWC) steadily climbing up from 0.10 to 0.30.

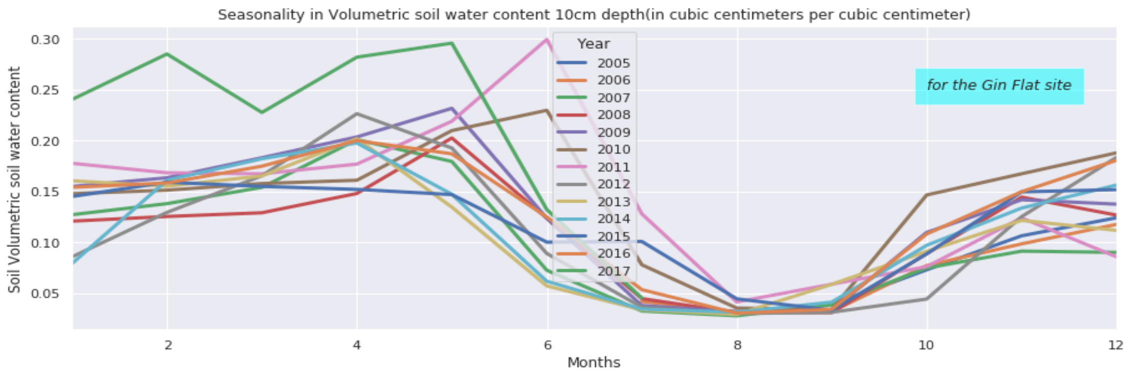


Figure 3: Seasonality Plot: Volumetric Water Content(VWC) of 10cm depth: Gin Flat site

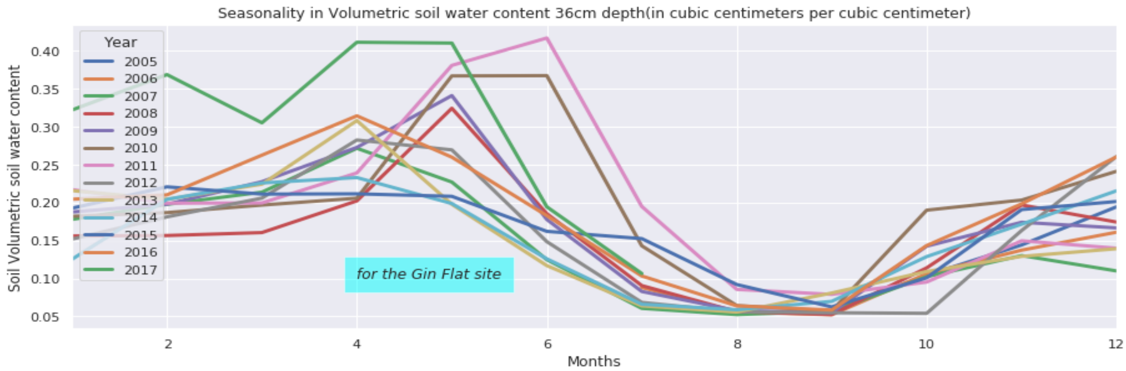


Figure 4: Seasonality Plot: Volumetric Water Content(VWC) of 36cm depth: Gin Flat site

From above figure3 and figure4 illustrate seasonal distribution of volumetric water content(VWC) from years 2005 to 2017 with months range from 1 to 12 for Gin Flat site. In figure3, it depicts about volumetric water content (VWC) of depth 10cm for given years and months. For the months 1 to 3 volumetric water content (VWC) is in between 0.0 to 0.1. whereas for months 4 to 7 significantly increased in between 0.2 to 0.5. In figure4, it depicts about volumetric water content (VWC) of depth 36cm for given years and months. For the months 1 to 3 volumetric water content (VWC) is in between 0.1 to 0.2. whereas for months start of 4 to end of 6 significantly increased in between 0.2 to 0.5.

3.2.3 Time series Trend Plot

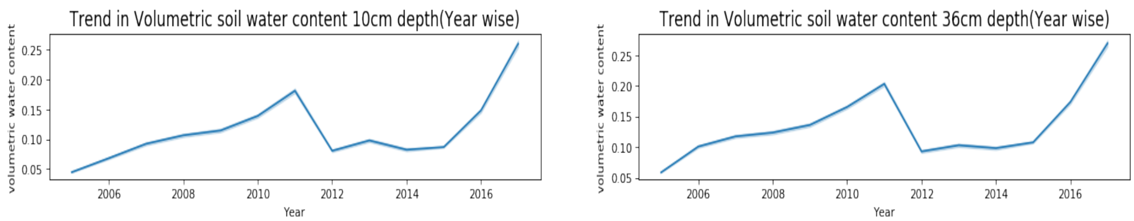


Figure 5: Trend Plot: Volumetric Water Content(Year wise): Dana Meadows site

The above figure5 illustrate year wise trend in soil Volumetric Water Content for depth 10cm and 36cm in Dana Meadows site. The chart has shown that Volumetric Water Content for both depth values is gradually increased in the given years.

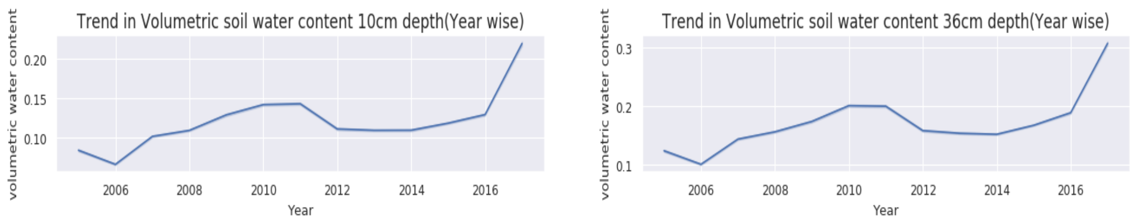


Figure 6: Trend Plot: Volumetric Water Content(Year wise): Gin Flat site

The above figure6 illustrate year wise trend in soil Volumetric Water Content for depth 10cm and 36cm in Gin Flat site. The chart has shown that Volumetric Water Content for both depth values is gradually increased in the given years.

3.2.4 Test for Stationary series Test

It is important to check the given data is stationary or non-stationary before applying to models. Stationary series data make predictions relatively easy and it will more reliable. To check the Stationarity of given data series two tests are used such as ADF (Augmented Dickey-Fuller) Test and KPSS (Kwiatkowski-Phillips-Schmidt-Shin) Test,

- **ADF(Augmented Dickey Fuller)**:
  - **Null Hypothesis**: The series has a unit root.
  - **Alternate Hypothesis**: The series has no unit root.

  if the test statistic is less than the critical value, then the null hypothesis is rejected it means series is Stationary. when the test statistics are greater than the critical value then alternate hypothesis is rejected which means the series is not Stationary.

- **KPSS(Kwiatkowski-Phillips-Schmidt-Shin)**:
  - **Null Hypothesis**: The series has a unit root.
  - **Alternate Hypothesis**: The series has no unit root.

  if the test statistics are greater than the critical value, then the null hypothesis is rejected it means series is not Stationary. When the test statistics are less than the critical value, then the alternate hypothesis is rejected which means series is Stationary.

```
ADF Statistic: -5.288347382095587
p-value: 5.788329523680391e-06
Critial Values:
   1%, -3.430415154253313
Critial Values:
   5%, -2.8615687971603614
Critial Values:
   10%, -2.5667853277205785
```

```
KPSS Statistic: 2.659656
p-value: 0.010000
Critial Values:
   10%, 0.347
Critial Values:
   5%, 0.463
Critial Values:
   2.5%, 0.574
Critial Values:
   1%, 0.739
```

Stationarity test of soil Volumetric Water Content 10 cm depth

```
ADF Statistic: -4.239077931365423
p-value: 0.0005654195434078626
Critial Values:
   1%, -3.430415154253313
Critial Values:
   5%, -2.8615687971603614
Critial Values:
   10%, -2.5667853277205785
```

```
KPSS Statistic: 2.399483
p-value: 0.010000
Critial Values:
   10%, 0.347
Critial Values:
   5%, 0.463
Critial Values:
   2.5%, 0.574
Critial Values:
   1%, 0.739
```

Stationarity test of soil Volumetric Water Content 36cm depth

Figure 7: Stationary Test: Soil volumetric water content of 10cm and 36cm depth: Dana Meadows site

The Stationary test is conducted on volumetric soil moisture, in the figure above test results of volumetric soil moisture are shown for the Dana Meadow site. The Figure7 are the test results of ADF(Augmented Dickey-Fuller) and KPSS (Kwiatkowski Phillips Schmidt Shin) tests for Volumetric soil moisture of 10cm and 36cm depth. The results of ADF (Augmented Dickey-Fuller) indicate that test statistics for volumetric soil moisture for given depths are less than critical value. Hence the null hypothesis is rejected and it means given Soil volumetric water content series of 10cm and 36cm depths are stationary.

The results of KPSS (Kwiatkowski Phillips Schmidt Shin) also show that test statistics are greater than the critical value for both of depths. which indicates that the alternative hypothesis is rejected and it means given volumetric soil moisture content series of 10cm and 36cm depths are stationary.

```
ADF Statistic: -5.661834836316565
p-value: 9.341536581888261e-07
Critial Values:
   1%, -3.4304155474080607
Critial Values:
   5%, -2.8615689709266885
Critial Values:
   10%, -2.566785420210647
```

```
KPSS Statistic: 4.931417
p-value: 0.010000
Critial Values:
   10%, 0.347
Critial Values:
   5%, 0.463
Critial Values:
   2.5%, 0.574
Critial Values:
   1%, 0.739
```

Stationarity Test of Volumetric Water Content 10cm depth

```
ADF Statistic: -5.257781597177048
p-value: 6.693053825558189e-06
Critial Values:
   1%, -3.430415548065089
Critial Values:
   5%, -2.8615689712170815
Critial Values:
   10%, -2.5667854203652136
```

```
KPSS Statistic: 5.606717
p-value: 0.010000
Critial Values:
   10%, 0.347
Critial Values:
   5%, 0.463
Critial Values:
   2.5%, 0.574
Critial Values:
   1%, 0.739
```

Stationarity Test of Volumetric Water Content 36cm depth

Figure 8: Stationary Test: Soil volumetric water content of 10cm and 36cm depth: Gin Flat site

The Figure8 are the test results of ADF(Augmented Dickey-Fuller) and KPSS (Kwiatkowski Phillips Schmidt Shin) tests for volumetric soil moisture content of 10cm and 36cm depth of Gin Flat site. The results of ADF (Augmented Dickey-Fuller) indicate that test statistics for both depths are less than the critical value. Hence the null hypothesis is rejected and it means given Volumetric soil moisture content series of 10cm and 36cm depths are stationary. The results of KPSS (Kwiatkowski Phillips Schmidt Shin) also show that test statistics are greater than the critical value for both depths. which indicates that the alternative hypothesis is rejected and it means given Soil volumetric water content series of 10cm and 36cm depths are stationary.

3.3 Data Cleaning and Pre-Processing

3.3.1 Replacing missing values

This acquired datasets having some missing values and it is important to replace missing values of the dataset to get better results from models. Firstly, the total number of missing values from both datasets are calculated with the help of function presents in python. That given general idea about the count of missing values from each variable of both selected sites datasets. After that missing values replaced by previous year data of same-day that is from previous year value of the same day using user-defined function.

3.3.2 Removing Seasonality and Trend

With the help of the STL method(Seasonal and Trend decomposition using Loess) given datasets are divided by seasonal index obtained from the STL method of decomposition that removed seasonality from datasets. Similarly, after getting a trend component from given STL decomposition, it used to subtract trends from given datasets. For all the given variables of datasets trend and seasonality are removed. As a result of some negative temperature, values present in dataset Multiplicative Time-series modeled used to remove trend and seasonality from given datasets. The multiplicative model used to removed trend and seasonality from given datasets for all variables. Multiplicative explained using following equation such as,

Multiplicative Time Series: $Value = BaseLevel\ \times$Trend$\times$Seasonality$\times$Error.

4 Design Specification

This section focused on neural networks models used for this study,

4.1 Artificial Neural Network (ANN) Structure

This neural network inspired by the human nervous systems like massive interconnection and parallel execution approach. ANN having the ability to solve problems with machine learning neurons with the help of a data-based mathematical model. ANN is very good at identifying the relationships between inputs and outputs without direct knowledge. The relationship between input and output for the artificial neural network explain as follow:

$$Y = f(X^n). \tag{1}$$

for above equation, X^n is for n-dimensional input vector made up of variables like $X_1,\cdots,X_i,\cdots,\ X_n$; and where Y represents output vector. The above equation no really presents the ANN model structure rather it will show network parameters. The most commonly used ANN structure is assembled with the help of the input layer, hidden layer, and output layer. In this structure input and output, layers are totally independent of each other and between these two layers, one or more hidden layers may present. As shown in Figure9 below the structure of typical ANN architecture,

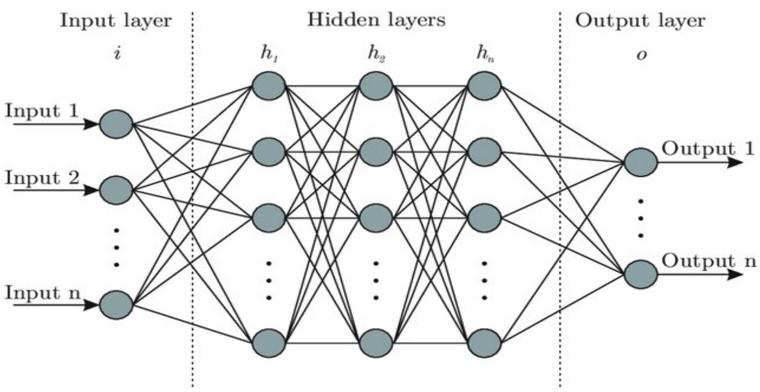


Figure 9: Architecture of typical Artificial Neural Network(ANN)(Facundo et al.; 2017)

The various types of ANN are distinguished by their network architecture. The complexity of ANN is closely dependant on a number of hidden layers used as well a number of neurons in each hidden layer. And the use of a number of hidden layers and number of neurons in each hidden layer depends on characteristics of the input data such as data types and sometimes the size of the dataset.

4.2 The LSTM Structure

To overcome the drawbacks of RNN by adding more additional interactions per cell this study delivered by (Hochreiter and Schmidhuber; 1997) and proposed the LSTM model. So, LSTM is a special form of RNN which is capable to grasp and learn long term dependencies for any given time periods. It works like a chain structure but instead of using a single form of a neural network like RNN, it has four different interacting layers with a unique method of communication. The structure of the LSTM model presented in the following Figure10,

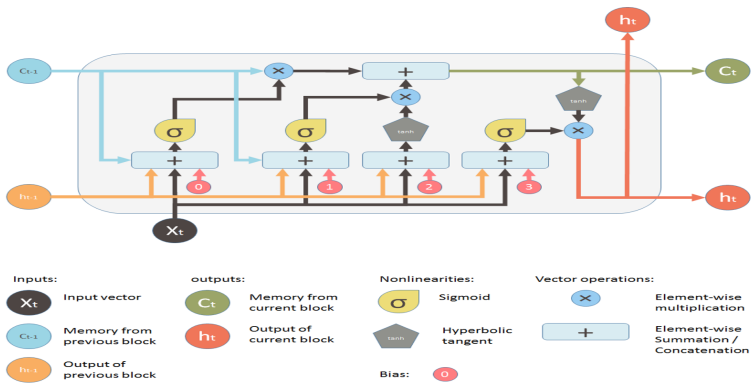


Figure 10: Structure of LSTM Memory Block(Yan; 2016)

As shown in above Figure10 a typical memory block of LSTM cell. An LSTM network is developed with the help of several memory blocks like above. Each cell of the LSTM network transformed next is present cell state and hidden state. Where cell state plays an important role in keep data flowing through cells and this makes data flow in a forward manner without unchanged but some linear transformation may occur while forwarding data. With the help of Sigmoid gates, data can be added or removed from given cell state. A gate is nothing but a series of matrix operations that store different individual weights. Cause of this structure LSTM able to avoid long term dependency problems due it uses the gate for controlling the memorizing process. The LSTM computes the mapping from an input sequence X to the output by looping through equations 1 to 6 with initial values C_o and h_o as described by (Goodfellow et al.; 2016),

Firstly, the LSTM network identifies information which not required and that information will be omitted from the cell in that given step. As discussed before the Sigmoid function used to identify and exclude data, this function takes inputs from last LSTM unit output($h_t - 1$) at time $t - 1$ and the current input (X_t) at time t. The Sigmoid function also helps in determining which part from the old output should be eliminated. This gate is called the forget gate (or f_t); where f_t is nothing but vector with values ranging from 0 to 1, corresponding to each number in the cell state, $C_t - 1$.

$$f_t = \sigma(W_f[h_{t-1}, X_t] + b_f). \tag{2}$$

Here, σ is function, and W_f and b_f are the weight matrices and bias, respectively, of the forget gate. The following step is for deciding and storing information from the new input (X_t) in the cell state as well as to update the cell state. This step made up of two parts, the Sigmoid layers and second the tanh layer. Firstly, the Sigmoid layer decides whether the new information should be updated or ignored ($0 or 1$), and second, the tanh function gives weight to the values which passed by, deciding their level of importance ($-1 to 1$). The two values are multiplied to update the new cell state. This new memory is then added to old memory (C_{t-1}) resulting in C_t.

$$i_t = \sigma(W_i[h_{t-1}, X_t] + b_i), \tag{3}$$

$$N_t = \tanh(W_n[h_{t-1}, X_t] + b_n), \tag{4}$$

$$C_t = C_{t-1}, f_t + N_t i_t. \tag{5}$$

Here, C_{t-1} and C_t are the cell states at time $t-1$ and t, while W and b are the weight matrices and bias, respectively, of the cell state. In the final step, the output values (h_t) is based on the output cell state (O_t) but is filtered version. First, a Sigmoid layer decides which parts of the cell state make it to the output. Next, the output of the Sigmoid gate (O_t) is multiplied by the new values created by the tanh layer from the cell state (C_t) with a value ranging between -1 and 1.

$$O_t = \sigma(W_o[h_{t-1}, X_t] + b_o), \tag{6}$$

$$h_t = O_t \tanh(C_t). \tag{7}$$

Here, W_o and b_o are the weight matrices and bias, respectively, of the output gate.

5 Implementation

5.1 Data Transformation

whole process data transformation explained as below,

- **Changing Data Types of Variables:**
  - All Variables are converted to float data type.
  - Date and time variable is converted to datetime data type.
- **Making Date Time Variable as Index:**
  - Date time variable set as index for given time series datasets.
- **Replacing the missing values:**
  - Missing values in both sites are replaced using from previous year value of same day using user-defined function.
- **Stabilize Variance:**
  - To stabilize the variance of variables, yeo-johnson transform is used(Adeyemi et al.; 2018).
- **Standardization:**
  - Datasets are standardized by using z-score of data points of given variables(Adeyemi et al.; 2018).
- **Deseasonalized & Detrend:**
  - Using Multiplicative time series model the given datasets deseasonalized by using dividing each variable with its seasonal index which obtained using STL method.
  - Using Multiplicative time series model the given datasets detrend using subtracting each variable with its trend component which obtained using STL method.

- **Adding Two Input Signals:**
  - With help of index variable two new variables are added to dataset.
  - The day of year variable is added which shows value of day between 1 to 366.
  - The hour of day variable is added which shows value of hour between 0 to 23.
- **formation of Train & Test Set:**
  - The 90:10 ratio is selected for dividing the dataset of each sites. Each model is trained on 90 percent of data and 10 percent data is selected for testing the implemented models.
- **Scaling of Train & Test Set:**
  - While training the models for each site, train and test data is scaled using sckit-learn's MinMaxScaler.
  - Reason for doing scaling on dataset is neural networks works well on values between -1 to 1. So both site's train and test datasets are scaled using this to achieve good performance and results as well.

5.2 Batch formation

Instead of training the models on whole train sets, the training dataset used in batch-wise by using the generator function. The user-defined function used is batch_generator to defined two new arrays for a batch of input signals as well for a batch of output signals and these batches of input and output. This two array is filled by taking the random indexes to the training set. part of data is copied into these batches which starting at this index. For this defined function, the first parameter is batch size is passed of value 256 to make use of full GPU in order to perform this modeling for a given problem. The second parameter is sequence length, in this problem 8-week time-steps are chosen which means each random sequence gives 8 weeks observations. Total sequence length is formed like this, one time-step formed by using one hour which means in one day a total 24 observation. To count it for one week it is like 24×7. So, to covered required observation periods that is 8 weeks, it calculated using $24 \times 7 \times 8$ which equal to 1344. So sequence length and batch size for batch generator function are 1344 and 256 respectively. This function used for both models as well as to train each model on each individual site.

5.3 Validation set

To avoid problems of overfitting of purposed models that is the models perform well on training sets so that it does not generalize well on testing data. By keeping this point in mind hence the performance of each model is observed for after each epoch on the test set. If the performance of the model seen improving on the test set then only model weights are stored. As compared to batches for training the purposed model using batch_generator function but for the testing the whole sequence of data is passed from the given test set and then after prediction accuracy measured on that whole testing set.

5.4 Modeling

5.4.1 Model creation

```
model.summary()

Model: "sequential_20"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
lstm_32 (LSTM)               (None, None, 512)         1064960
_________________________________________________________________
dropout_29 (Dropout)         (None, None, 512)         0
_________________________________________________________________
dense_41 (Dense)             (None, None, 2)           1026
=================================================================
Total params: 1,065,986
Trainable params: 1,065,986
Non-trainable params: 0
```

Figure 11: Model Summary: LSTM model: Site 1

```
model.summary()

Model: "sequential_21"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
dense_42 (Dense)             (None, None, 100)         800
_________________________________________________________________
dense_43 (Dense)             (None, None, 100)         10100
_________________________________________________________________
dense_44 (Dense)             (None, None, 2)           202
=================================================================
Total params: 11,102
Trainable params: 11,102
Non-trainable params: 0
```

Figure 12: Model Summary: ANN model: Site 1

```
model.summary()

Model: "sequential_22"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
lstm_33 (LSTM)               (None, None, 512)         1064960
_________________________________________________________________
dropout_30 (Dropout)         (None, None, 512)         0
_________________________________________________________________
dense_45 (Dense)             (None, None, 2)           1026
=================================================================
Total params: 1,065,986
Trainable params: 1,065,986
Non-trainable params: 0
```

Figure 13: Model Summary: LSTM model: Site 2

```
model.summary()

Model: "sequential_23"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
dense_46 (Dense)             (None, None, 100)         800
_________________________________________________________________
dense_47 (Dense)             (None, None, 100)         10100
_________________________________________________________________
dense_48 (Dense)             (None, None, 2)           202
=================================================================
Total params: 11,102
Trainable params: 11,102
Non-trainable params: 0
```

Figure 14: Model Summary: ANN model: Site 2

As shown in above Figure11 and Figure13 model summary of LSTM model for site 1 and site 2 respectively. There is a first layer of LSTM with 512 output for each time-step in the given sequence. Because this is the first layer of models due to that Keras needs aware about the shape of the input sequence and which is batch of a sequence of arbitrary length which shown by None as well each observation has 6 input-signals. After that, there is one dropout layer with a dropout of 0.5. Next to drop out layer their fully connected layer i.e. dense layer which maps 512 values down to only 2 values because this study focused on prediction on 2 output signals from given input signals. In the previous stage of data transformation where scaler object used to limit dataset value between 0 and 1. due to this output of LSTM also needed to be in between 0 and 1. hence Sigmoid activation function used in dense layer to get the output of LSTM model in the required format. As shown in above Figure12 model for Site 1 and Figure14 model for Site 2 model summary of ANN model for site 1 and site 2 respectively. For ANN first layer is fully connected layer with 100 output for each time-step in the given sequence. It is the first layer of ANN due to Keras needed to know about the shape of the input sequence and sequence arbitrary length. Each observation having 6 input signals. there is one more dense layer is connected with the same output for each time step similarly first layer. last dense layer which used to maps 100 values down to only 2 values as per the prediction purpose of two output signals.

For both the model such as LSTM and ANN, RMSprop optimizer used with initial learning rate at $1e-3$ for prediction of volumetric soil moisture of site 1 and site 2 for depth $10cm$ and $36cm$.

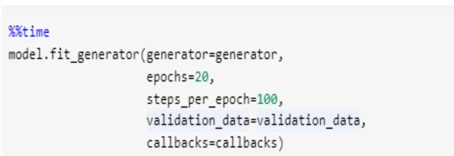
5.4.2 Loss Function

To check how well the model performed in the prediction of values and how close the prediction values to true values by using loss function. For this proposed study Mean Squared Error(MSE) is used to check loss in the prediction process. As proposed models are using input signals for a few time steps at the very beginning. so, it may possible that output is inaccurate. Hence then models would try to use loss-value in early time-steps to get accurate output but this may twist out the shape of the model. By keeping this thing in mind proposed model made by giving a warmup-period of 50-time steps to get accuracy in output in later time-steps.

5.4.3 Callback Function

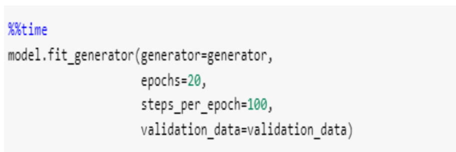
While training the proposed models all callbacks i.e. values of checkpoints are saved for Keras using various callback functions. In this study total, four callback function is implemented for the various proposed use. The first callback used to save all checkpoints during the training of the model. The second callback used to stop the optimization of the model when the performance of the proposed model getting exacerbate on the given validation set. Third callback function used for saving the TensorBoard log while training the model. The fourth callback function used to monitor validation loss, basically, it used to reduce the given learning rate for optimizer function when the loss of validation is not improving since the last epoch. fourth callback function contains factor value 0.1 and when this function needed to reduce learning rate then it multiplies by factor value and the learning rate doesn't get reduce below the $1e - 4$ as defined in a fourth callback function.

5.4.4 Training model



```
%%time
model.fit_generator(generator=generator,
                    epochs=20,
                    steps_per_epoch=100,
                    validation_data=validation_data,
                    callbacks=callbacks)
```

Figure 15: Training of LSTM model



```
%%time
model.fit_generator(generator=generator,
                    epochs=20,
                    steps_per_epoch=100,
                    validation_data=validation_data)
```

Figure 16: Training of ANN model

Whole proposed models are implemented in google cloud service with 13Gb NVIDIA's Tesla K80 GPU. As shown in above Figure15 and Figure16 training of LSTM and ANN.

6 Evaluation

In this section predictive performance of neural network models presented,

6.1 Case Study 1: LSTM

The table5 below, shows 24-Hour before prediction results of Volumetric Water Content for depth 10cm and 36cm for given sites using the LSTM model.

Site	Epoch	MSE
Dana Meadow	16	0.001552
Gin Flat	19	0.002365

Table 5: Result Evaluation of LSTM Model

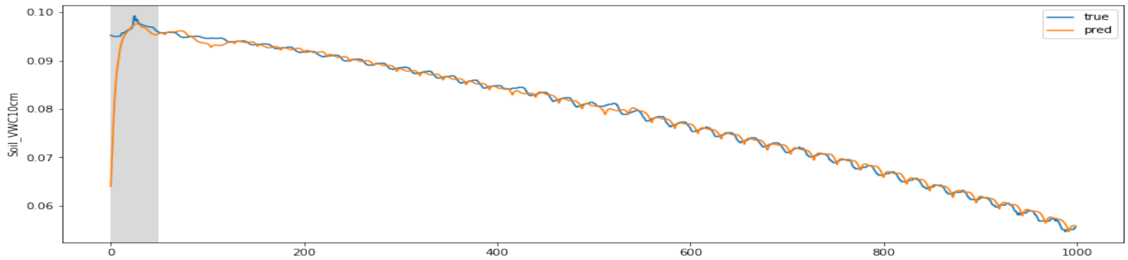


Figure 17: True Value and LSTM Predicted Plot: Volumetric Water Content(VWC) 10cm depth: Dana Meadow

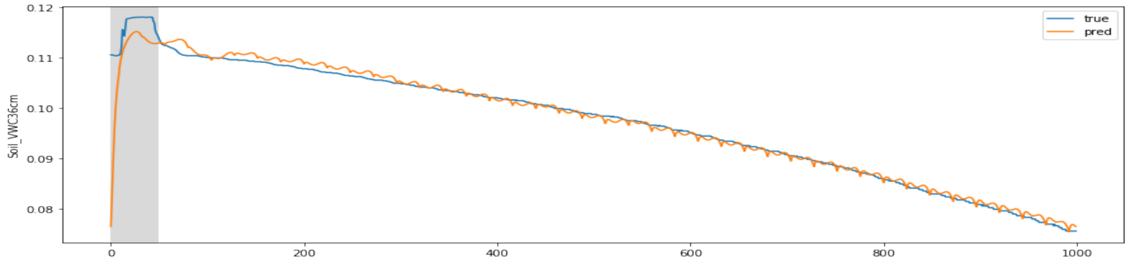


Figure 18: True Value and LSTM Predicted Plot: Volumetric Water Content(VWC) 36cm depth: Dana Meadow

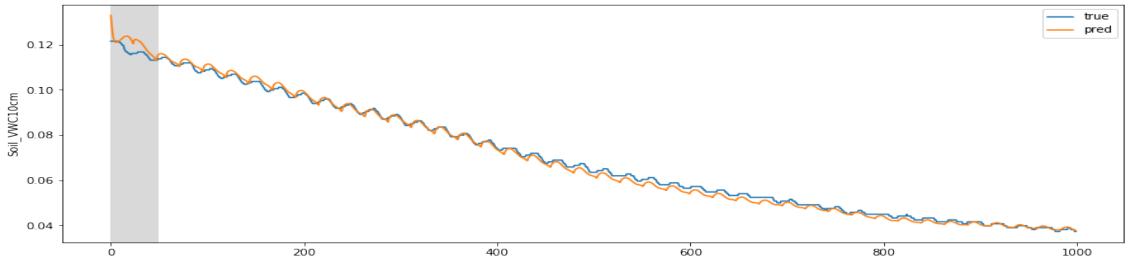


Figure 19: True Value and LSTM Predicted Plot: Volumetric Water Content(VWC) 10cm depth: Gin Flat

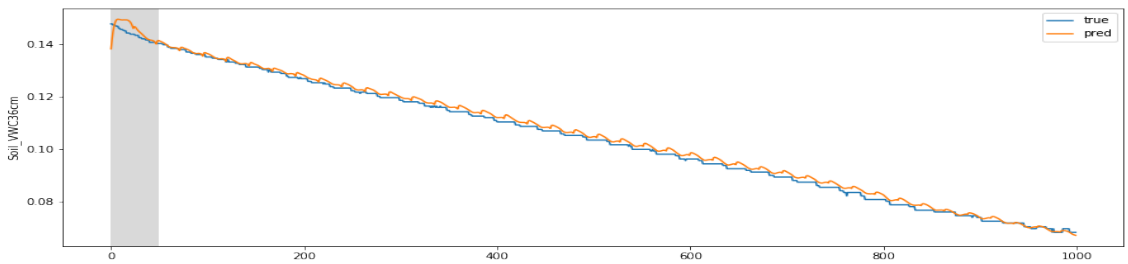


Figure 20: True Value and LSTM Predicted Plot: Volumetric Water Content(VWC) 36cm: Gin Flat

In conclusion to the above LSTM model results, each site trained separately to predict 24-hour before Volumetric Water Content of 10cm and 36cm depth. As shown in the above Figures of Dana Meadow site results, the Volumetric Water Content of 10cm depth predicted very well with catching all peaks like true values. Whereas for volumetric soil moisture of 36cm depth forecasted very well but sometimes inaccurate peaks. As a result figure of the Gin Flat site, the LSTM model learned well for forecasting Volumetric Water Content for 10cm depth as well as 36cm depth but prediction somewhat struggles to catch center level peaks.

6.2 Case Study 2: ANN

The table6 below, shows 24-Hour before prediction results of Volumetric Water Content for depth 10cm and 36cm for given sites using ANN model.

Site	Epoch	MSE
Dana Meadow	20	0.001316
Gin Flat	20	0.002022

Table 6: Result Evaluation of ANN Model of Transformed Data

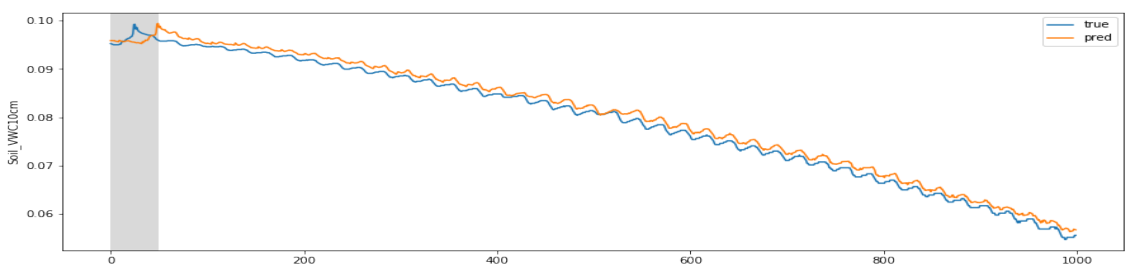


Figure 21: True Value and ANN Predicted Plot: Volumetric Water Content(VWC) 10cm depth: Dana Meadow

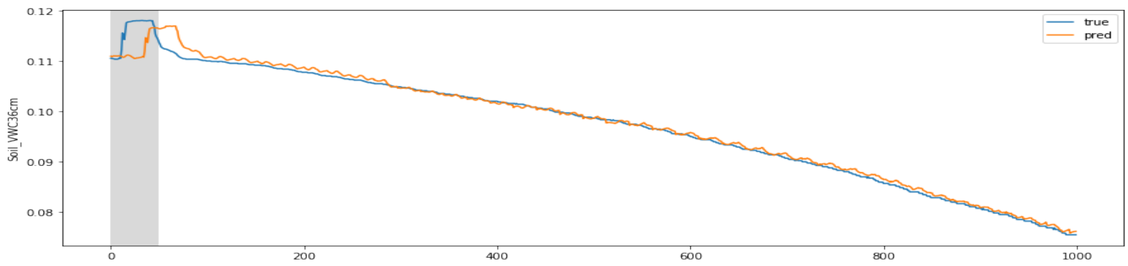


Figure 22: True Value and ANN Predicted Plot: Volumetric Water Content(VWC) 36cm depth: Dana Meadow

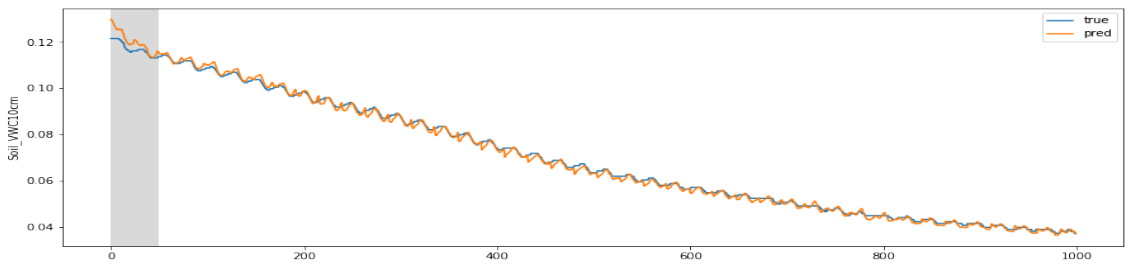


Figure 23: True Value and ANN Predicted Plot: Volumetric Water Content(VWC) 10cm depth: Gin Flat

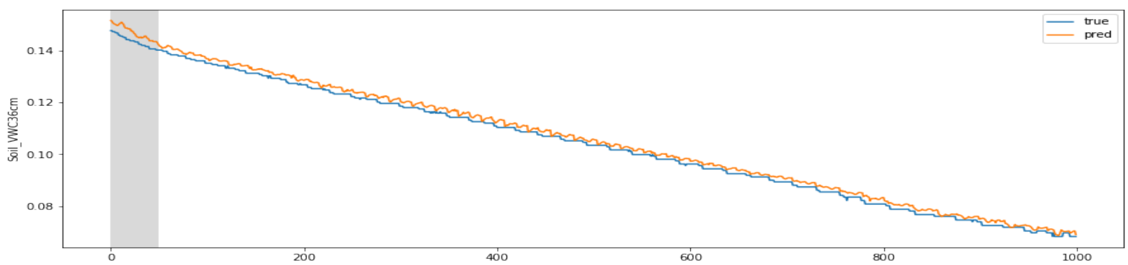


Figure 24: True Value and ANN Predicted Plot: Volumetric Water Content(VWC) 36cm depth: Gin Flat

To sum up to the above ANN model results, each site trained separately to predict 24-hour before Volumetric Water Content of 10cm and 36cm depth. As shown in the above Figures of Dana Meadow site results, the Volumetric Water Content of 10cm depth not predicted so well. Whereas for Volumetric Water Content of 36cm depth forecasted very well but sometimes inaccurate peaks. As a result figure of the Gin Flat site ANN model merely learned to predict, Volumetric Water Content of both 10cm and 36cm depth.

7 Conclusion and Future Work

The implemented neural network model such as LSTM performed well for prediction of soil moisture for 10cm and 36cm depths as compared to ANN for given sites.

The LSTM model performed excellently in terms of prediction Volumetric Water Content of 10cm depth for Dana Meadow as well as Gin Flat sites by catching all oscillation cycles like true values of Volumetric Water Content of 10cm depths. In opposite to this, in results of 36cm depth, Volumetric Water Content peaks are inaccurate and are excessive somewhere. LSTM performed with MSE value of 0.001552 and 0.002365 for Dana Meadow and Gin Flat sites respectively.

Whereas, results of ANN for both sites are inconsistent with each other. For instance, in the prediction of volumetric soil moisture of 10cm depth, ANN did a fine job for the Gin Flat site as compared to the results of the Dana Meadow site. Whereas, prediction of volumetric soil moisture of 36cm depth mostly accurate for the Dana Meadow site as against to Gin Flat site results for the same depth. ANN performed with MSE value of 0.001316 and 0.002022 for Dana Meadow and Gin Flat sites respectively.

To conclude from the above finding, LSTM surpasses in the learning of Volumetric Water Content for 10cm depth for both sites as compared to 36cm depth. For ANN, for each site that is for both sites sometimes it performed excellent either for 10cm depth or 36cm depth. So, ANN results are different for different sites.

The future work will contain five things firstly, as LSTM model prediction is accurate for both sites for 10cm depth then the possibility of prediction of Volumetric Water Content of 10cm depth of Dana Meadow from trained model of LSTM on Gin Flat site will be check and Vise Versa. Secondly, this prognostic system will be expanded to forecast rainfall in both areas to improve irrigation facilities of farms to achieve growth in business. Third, this prediction results will be passed as inputs to the predictive irrigation system for the prediction of water irrigation scheduling. Fourth, additional soil properties data and environmental data will be included checking their effect on given predictions. Fifth, timestamp of prediction that is one day ahead(24-Hours) in future it will be extended to fifteen days(360-Hours) and thirty days(720-Hours) in the future for the same depths.

References

Adeyemi, O., Grove, I., Peets, S., Domun, Y. and Norton, T. (2018). Dynamic neural network modelling of soil moisture content for predictive irrigation scheduling, *MDPI* **18**(3408). doi:10.3390/s18103408.

Aniley1, A. A., S.K, N. K. and A, A. K. (2018). Soil moisture sensors in agriculture and the possible application of nanomaterials in soil moisture sensors fabrication, **6**. www.ijaert.org/wp-content/uploads/2018/01/131469.pdf.

Ansari, S. and Deshmukh, R. R. (2017). Estimation of soil moisture content: A review, *International Journal of Theoretical and Applied Mechanics* **12**: 571–577. https://www.ripublication.com/ijtam17/ijtamv12n3_18.pdf.

Araya, S. N. and Ghezzehei, T. A. (2019). Using machine learning for prediction of saturated hydraulic conductivity and its sensitivity to soil structural perturbations, *Water Resources Research* **55**: 5715—-5737. doi:10.1029/2018WR024357.

Beven, K. J. (2019). Hydrologic sciences.

Bilskie, J. and Scientific, C. (2001). Soil water status: content and potential.

Cai, Y., Zheng, W., Zhang, X., Zhangzhong, L. and Xue, X. (2019). Research on soil moisture prediction model based on deep learning, *PLOS ONE* . doi:10.1371/journal.pone.0214508 .

Chatterjee, S., Dey, N. and Sen, S. (2018). Soil moisture quantity prediction using optimized neural supported model for sustainable agricultural applications, *Sustainable Computing: Informatics and Systems* . doi:10.1016/j.suscom.2018.09.002.

D'Emilio, A., Aiello, R., Consoli, S., Vanella, D. and Iovino, M. (2018). Artificial neural networks for predicting the water retention curve of sicilian agricultural soils, *Water* **10**(10). doi:10.3390/w10101431.
URL: *https://www.mdpi.com/2073-4441/10/10/1431*

Facundo, B., Juan, G. and Víctor, F. (2017). Prediction of wind pressure coefficients on building surfaces using artificial neural networks, *Energy and Buildings* **158**.

Fang, K., Pan, M. and Shen, C. (2019). The value of smap for long-term soil moisture estimation with the help of deep learning, *IEEE* **57**(4): 2221–2233. doi:10.1109/TGRS.2018.2872131.

Fang, K., Shen, C., Kifer, D. and Yang, X. (2017). Prolongation of smap to spatio-temporally seamless coverage of continental us using a deep learning neural network, *Geophysical Research Letters* **44**(5): 11030–11039. doi:10.1002/2017GL075619.

Goodfellow, I., Bengio, Y. and Courville, A. (2016). *Deep Learning*, MIT Press. `http://www.deeplearningbook.org`.

Graves, A., rahman Mohamed, A. and Hinton, G. E. (2013). Speech recognition with deep recurrent neural networks, *2013 IEEE International Conference on Acoustics, Speech and Signal Processing* pp. 6645–6649.

Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory, *Neural Computation* **9**(8): 1735–1780.

Jensen, K. (2012). A diagram showing the relationship between the different phases of crisp-dm, `ftp://public.dhe.ibm.com/software/analytics/spss/documentation/modeler/18.0/en/ModelerCRISPDM.pdf`.

Kaingo, J., Tumbo, S. D., Kihupi, N. I. and Mbilinyi, B. P. (2018). Prediction of soil moisture-holding capacity with support vector machines in dry subhumid tropics, *Hindawi* pp. 1–10. doi:10.1155/2018/9263296.

Khedri, E., Hasanlou, M. and Tabatabaeenejad, A. (2017). Estimating soil moisture using polsar data: A machine learning approach, *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **XLII-4/W4**: 133–137.
URL: *https://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XLII-4-W4/133/2017/*

Liu, X. (2011). Evaluation of the heat-pulse technique for measuring soil water content with thermo-tdr sensor, *Procedia Environmental Sciences* **11**: 1234 – 1239. 2011 2nd International Conference on Challenges in Environmental Science and Computer Engineering (CESCE 2011).
URL: *http://www.sciencedirect.com/science/article/pii/S1878029611010085*

Mikolov, T., Kombrink, S., Burget, L., Cernocký, J. and Khudanpur, S. (2011). Extensions of recurrent neural network language model, *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* pp. 5528–5531.

Ordóñez, F. J. and Roggen, D. (2016). Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition, *Sensors* **16**(1).
URL: *https://www.mdpi.com/1424-8220/16/1/115*

Prakash, S., Sharma, A. and Sahu, S. S. (2018). Soil moisture prediction using machine learning, *2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT)* pp. 1–6. doi:10.1109/icicct.2018.8473260.

Prasad, R., Deo, R. C., Li, Y. and Maraseni, T. (2019). Weekly soil moisture forecasting with multivariate sequential ensemble empirical mode decomposition and boruta-random forest hybridizer algorithm approach, *Catena* pp. 149–166. doi:10.1016/j.catena.2019.02.012.

Shukla, M. K. (2013). *Soil Physics:*.

Singh, S. and Kasana, S. S. (2019). Estimation of soil properties from the eu spectral library using long short-term memory networks, *Geoderma Regional* . doi:10.1016/j.geodrs.2019.e00233.

Subir, D. (2011). Design and fabrication of a soil moisture meter using thermal conductivity properties of soil, *Sensor Transducers* **132**: 8.

vision Europe, S. (2018). What is the crisp-dm methodology?, `https://www.sv-europe.com/crisp-dm-methodology/`.

Wirth, R. (2000). Crisp-dm: towards a standard process modell for data mining.

Yan, S. (2016). Understanding lstm and its diagrams, `https://medium.com/mlreview/understanding-lstm-and-its-diagrams-37e2f46f1714`.

Zhanga, J., Zhub, Y., Zhanga, X., Yec, M. and Yang, J. (2018). Developing a long short-term memory (lstm) based model for predicting water table depth in agricultural areas, *Journal of Hydrology* pp. 918–929. doi:10.1016/j.jhydrol.2018.04.065.

Zhaoa, W., Sánchezb, N., Luc, H. and Li, A. (2018). A spatial downscaling approach for the smap passive surface soil moisture product using random forest regression, *Journal of Hydrology* pp. 1009–1024. doi:10.1016/j.jhydrol.2018.06.081.